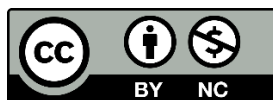




UNIVERSITAT<sub>DE</sub>  
BARCELONA

**El cultiu del blat a l'ambient mediterrani davant la crisi climàtica:  
Perspectives geogràfiques des de la teledetecció i l'ecofisiologia.**

Joel Segarra



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial 4.0. Espanya de Creative Commons**.

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial 4.0. Spain License**.



# **EL CULTIU DEL BLAT A L'AMBIENT MEDITERRANI DAVANT LA CRISI CLIMÀTICA:**

PERSPECTIVES GEOGRÀFIQUES  
DES DE LA TELEDETECCIÓ I  
L'ECOFISIOLOGIA.

**JOEL SEGARRA**









Barcelona, juny 2024



UNIVERSITAT DE  
BARCELONA

**El cultiu del blat a l'ambient mediterrani davant la crisi climàtica:  
Perspectives geogràfiques des de la teledetecció i l'ecofisiologia.**

Memòria presentada per Joel Segarra per optar al grau de Doctor per la  
Universitat de Barcelona.

Programa de doctorat d'Ecologia, Ciències Ambientals i Fisiologia Vegetal de la  
Universitat de Barcelona

El present treball s'ha fet al Departament de Biologia Evolutiva, Ecologia i  
Ciències Ambientals de la Universitat de Barcelona sota la direcció del  
Dr. J.L. Araus (tutor) i del Dr. S. C. Kefauver.

Doctorand:

Joel Segarra

Una signatura manuscrita en negre que sembla dir "Joel", amb una línia fluida que s'estén cap a l'esquerra.







Als meus pares i avis

¿En quina època de la vinya t'ha tocat l'herència de treballar-la? ¿A l'hora de cavar? ¿A l'hora de veremar? ¿De celebrar la festa? Tot és una sola cosa.

*Ascesi.* N. Kazantzakis

Ves per lo món i meravella't.

*Llibre de meravelles.* R. Llull



## Agraïments

D'ençà que era petit he tingut la sort de viure el ritme de l'horta, les oliveres i els ametllers. El cicle de la vida i el suport mutu en el qual bastim la nostra existència.

Gràcies als aprenentatges, a base d'experiències i observació, que vaig viure amb els avis Josep i Conxita he pogut cultivar un gran interès per la biologia i, particularment, el món vegetal. Treballar al tros, collir olives o preparar conserves ha estat una gran escola i per això estic profundament agraït als avis. Ells no van poder estudiar més enllà de la infantesa, però em van transmetre sempre un profund coneixement. Tot sovint els més savis són els que menys oportunitats han tingut. Ells em van ensenyar a valorar la fortuna de poder estudiar sense haver de carregar-me, com ells havien fet d'ençà que eren vailets, de les cures excessives dels familiars o de collir i batre el blat sense tenir més trajectòria que servir la malaurada subsidència del no tenir gairebé res (perquè tot s'ho queden uns pocs).

Estic també profundament agraït a la iaia Encarnita perquè em va transmetre els valors de l'escola republicana i la consideració de l'aprenentatge com a única eina que una filla d'obriers del tèxtil podia tenir per tirar endavant. També a l'atzar genealògic dels besavis i rebesavis que m'ha permès viure entre parets de còdols davant d'una eixida que és mitja vida: aigua, terra, fruita i Sol. A l'avi Anselm, que no vaig poder conèixer, per entendre que som d'allà on anem i que l'herència d'on hem vingut, algeriana i de més enllà en el seu cas, és com la picada que lliga el guisat.

Voldria agrair al meu pare per haver-me transmès la seva passió per la lectura, l'escriptura i el coneixement. També per haver-me donat suport en desenvolupar la meva curiositat més enllà de no compartir els mateixos interessos. Agraeixo a la mare haver-me guiat en els estudis, acompanyat en els deures, haver-me transmès la fermesa i compromís de fer les coses i alhora l'amor i la consideració de les cures. Per emmirallar-me en la seva trajectòria professional apassionada i interessant. Per haver-me dut, d'ençà que era un nen, a viatjar pel món i per haver-me tractat amb franquesa i respecte dotant-me de veu i consideració per dir el que creia o pensava. Per haver-me permès, amb valentia, viatjar des dels tretze anys gairebé sol i aprendre llengües. Estic profundament agraït a la mare i a l'avi Josep per haver-me fet entendre que la crítica o l'assenyalament d'un error no és un atac a la meva personalitat sinó una oportunitat d'aprenentatge, sempre que es transmeti des de l'amor, com en el seu cas. També voldria agrair a la Maria, la meva companya, pel seu suport i el nostre bell vincle com un camí d'aprenentatges.



Estic agraït a tots els companys de l'escola i institut, perquè el suport mutu entre companys és l'únic que tenim. Estic feliç d'haver tingut la sort de poder gaudir de grans companys i amics i espero que aquells engegats per l'egoisme puguin trobar algun dia la pau de la solidaritat. Voldria agrair a tots els professors de l'escola pública el Turó i de l'institut Montserrat Miró i Vilà de Montcada i Reixac. Perquè els joves dels barris perifèrics també tenim coses a dir i amb la meitat d'oportunitats que d'altres tenen segur que molts més tirarien endavant amb estudis superiors.

Estic agraït també a tots els companys del grup de recerca d'ecofisiologia integrativa de cultius perquè en aquests cinc anys hem passat per molts moments diferents, però sempre hem mantingut algun espai per l'aprenentatge, el debat i el suport entre nosaltres que sempre és necessari en qualsevol feina: l'Omar, la Rut, la Cristina, en Thomas, en Jose Armando, la Fàdia, l'Adrià, la Maria Luisa, la Fàtima Zahra, la Melissa, en Vàlter, la Jara, en Yassine, l'Elena, en Mário, l'Ayesha i la Mariam, entre d'altres. També vull agrair als companys i professors del màster d'agrobiologia ambiental de la Universitat de Barcelona, del Postgrau en dinamització agroecològica de la UAB i del grau en Biotecnologia i estudis ambientals de la UAB i la Universitat Politècnica d'Hamburg respectivament.

Voldria donar les gràcies al professor Quim Azcón que per un curiós atzar em va il·lustrar en la comprensió de la fotosíntesi i el funcionament de les plantes quan era adolescent, uns aprenentatges que em van il·luminar el camí posteriorment. També al professor Josep Espluga que em va donar suport per poder rebre una beca a l'AMB i iniciar-me en la recerca amb l'estudi de l'agricultura de muntanya al Baix Llobregat, fet que em permeté treballar amb l'urbanista Teresa Gómez-Fabra i fascinar-me per la geografia, camp d'estudi del qual ara me'n sento part. Agraeixo també al professor J. L. Araus la seva proposta de fer el doctorat en el seu grup després de sentir-me presentar la tesi de màster sense gairebé conèixer-nos i que penso que ha resultat en un reeixit vincle professional. Així com al professor S. C. Kefauver especialment pels moments de debat i reflexió al voltant de l'ecofisiologia i la teledetecció que han catalitzat un gran aprenentatge.

Més enllà de l'educació reglada, voldria agrair als companys del col·lectiu de científics ConCiencias que des del 2016 hem estat co-creant coneixement i debatent al voltant de la ciència i la societat, especialment els primers anys a l'Ateneu Llibertari. Ha estat un espai que m'ha permès veure la realitat polièdrica de la ciència i poder transitar un camí d'investigació que tingués sentit des de la meva curiositat i inquietuds. També als companys de l'enyorat casal el Brot de Montcada i de l'actiu Ateneu la Malgirbada de Granollers, perquè construir des d'un horitzó col·lectiu és feixuc, però dona un sentit de solidaritat i aprenentatges que alleugereix la

malaurança individualista. També als companys de la plataforma antiincineració de Montcada i Reixac en aquesta lluita contra la multinacional del ciment que ens menja l'entorn i la salut. Vull agrair també als companys del grup de lectura del decreixement, principalment de la UAB, per totes les experiències i aprenentatges compartits que m'han permès comprendre millor el metabolisme planetari i establir fecundes amistats.

He tingut la fortuna de poder fer dues estades de dos mesos en cada cas a Mèxic i a Xile durant el meu doctorat. Estic molt agraït als companys de l'institut de geografia de la UNAM (ciutat de Mèxic) per haver-me acollit i donat suport en moments complicats. Les experiències en el treball de camp i les tècniques emprades així com els paisatges transitats formen part d'un bellíssim record que també ha quedat cristal·litzat en treballs acadèmics compromesos. Voldria agrair als amics i investigadors Andrea González, Javier Osorno i Don André Couturier la seva generositat i bonhomia. També voldria donar les gràcies als companys de la Universitat de Talca que em van acollir de manera excel·lent i em van permetre treballar en els paisatges mediterranis de la zona central de Xile així com conèixer la serralada dels Andes i d'altres geografies australs. Moltes gràcies a la Garazi pel seu suport i la seva llum, a en Pancho pel seu tarannà acollidor i generós i al professor Alejandro del Pozo per tots els rics debats i excursions compartides.

Gràcies a tothom que ha compartit el seu temps, coneixement i escolta amb mi en el camí d'aquesta tesi. Totes aquestes persones han endolcit i dotat de sentit aquest esforç a voltes difícil.



## **Abstract**

The Mediterranean Basin is a climate change hotspot, so finding sustainable agricultural management strategies that can both mitigate and adapt to the climate crisis is critical. One way to better adapt to the climate crisis is to use remote sensing techniques, which allow for precise monitoring of crop fields and improved management. Furthermore, in addition to monitoring individual fields, it provides a better spatial understanding of the agricultural mosaic and presents novel options for shaping sustainable landscapes. Satellite data has become more widely used in agricultural monitoring due to increased image availability and resolution improvements over the last decade. This thesis investigated the use of agronomic, physiological, environmental, and remotely sensed data to monitor and map wheat production, namely grain yield quantity and area, and quality, grain nitrogen (protein) content, in the Mediterranean environment. I have developed several modeling approaches using multispectral and geographically referenced landscape datasets. Machine learning approaches outperform simpler linear approaches because they capture the complexities of spectral data and better relate it to agronomic outputs. In my thesis, I argue that freely accessible Sentinel-2 imagery and publicly available georeferenced data can help assess wheat performance and define sustainable management approaches. In the current technological milieu, the sustainability of agroecosystems depends on an efficient and equitable understanding and distribution of the satellite data infrastructure. At the margins, particularly in less considered croplands, there are clear opportunities to improve yields and better integrate agricultural production into the landscape. The use of remote sensing with an agroecological perspective could improve the sustainability of Mediterranean cereal landscapes.

## Índex

<b>Introducció general</b>	1
1. El medi	3
1.1. Els cultius, el clima mediterrani i els seus canvis	3
1.2. La tècnica i l'espai	4
1.3. Els paisatges i les terres marginals	6
2. L'ecofisiologia i la singularitat del blat	8
2.1. Característiques	8
2.2. Rendiment i qualitat	9
2.3. Agroecosistemes cerealístics i el seus manejos	11
3. Comprensió del medi i teledetecció	13
3.1. Els satèl·lits, els drons i altres nivells de mesura	13
3.2. Missió Copernicus i el Sentinel-2	15
3.3. La fisiologia amb la teledetecció: els models	16
<b>Objectius</b>	20
<b>Informe del director de la tesi sobre l'impacte dels articles publicats</b>	21
<b>Resultats</b>	25
<b>Capítol 1.</b> Trade-offs between wheat agroecological cultivation and farming recovery on abandoned cropland in the Euro-Mediterranean region	25
<b>Capítol 2.</b> Suitability mapping and management monitoring in Castilian organic and conventional wheat fields with Sentinel-2 and spatial data	71
<b>Capítol 3.</b> Estimating wheat grain yield using Sentinel-2 imagery and exploring topographic features and rainfall effects on wheat performance in Navarre	84
<b>Capítol 4.</b> Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content	110
<b>Capítol 5.</b> Farming and Earth Observation: Sentinel-2 data to estimate within-field wheat grain yield	131
<b>Discussió general</b>	145
1. Estimació de rendiment del blat amb teledetecció	145



2. Estimació de qualitat del blat amb teledetecció.....	148
3. Efectes ambientals i del paisatge en el rendiment del blat.....	152
4. Mapatge de cultius amb teledetecció i sistemes d'informació geogràfica.....	154
5. Un maneig sostenible del blat.....	157
6. La importància geogràfica.....	160
7. Línies a seguir per establir agrosistemes de blat sostenibles.....	162
<b>Conclusions.....</b>	<b>165</b>
<b>Bibliografia.....</b>	<b>168</b>
<b>Annexos.....</b>	<b>177</b>
<b>Annex 1.</b> Remote sensing for precision agriculture: Sentinel-2 improved features and Applications.....	177
<b>Annex 2.</b> Forest cover loss in the Nevado de Toluca volcano protected area (Mexico) after the change to a less restrictive category in 2013.....	197
<b>Annex 3.</b> Managing abandoned Mediterranean mountain landscapes: The effects of donkey grazing on biomass control and floral diversity in pastures.....	224

## Introducció general

El coneixement és un riu alimentat de molts afluents. En aquest aiguabarreig el cabal es fa més substanciós. El riu avança i duu rics sediments que alimenten les assedegades lleres on broten cultius de tota mena, si l'home hi intervé, o frondosos boscos de ribera si l'aigua hi brolla bé. En arribar a la desembocadura el riu s'abraça amb la mar. Els colors blavissos reflecteixen tota la seva complexitat en un retrat bellíssim. De tot aquest impuls en surten deltes que prenen a la mar un tros del seu espai, en diàleg amb l'aigua i tots els seus habitants.

Com un riu, el coneixement porta un ric aliment a la nostra assedegada existència, impulsa la curiositat i l'afany per aprendre. Un exercici profundament humà; tanmateix, poc freqüent. Contemplar, pensar, assajar, interpretar resultats, o ésser creatiu és un exercici profundament escàs. La parcel·lació del coneixement entesa des d'un reduccionisme rígid nega també qualsevol opció o vocació il·lustrada. Una acadèmia que sovint s'encotilla i una il·lustració trobada al marge de l'imperi (Graeber, 2023). El generalista és sospitós. Sempre cal conèixer només d'un tema en molta profunditat. Com en un engranatge industrial. Un canal rodejat de ciment sense gaire més. Ni afluents ni aiguabarreig ni rics sediments de zones diverses, una il·lustració insuficient (Subirats, 1981).

Tot i la circumstància limitada del nostre temps, aquesta tesi naix des de l'afany interdisciplinari que la complexitat de la realitat demana. L'actualitat planetària configura un home desencisat, poruc i abocat a un soroll eixordador constant. La crisi climàtica i les poques perspectives de canviar el metabolisme planetari industrial, lligat certament a uns interessos destructors, ens empeny a una gran desil·lusió. De l'anàlisi dels límits planetaris proposada per Rockström et al. (2009), actualment sis dels nou límits planetaris han superat el llindar de regeneració. En particular, això inclou el canvi climàtic i la pèrdua de biodiversitat, que contribueixen a un estat d'extinció massiva a la Terra (Ceballos et al., 2015).

En aquesta circumstància la dificulta d'intentar copsar la realitat des de les ciències i trobar-hi sentit des d'alguna banda és majúscula. La resposta és, potser, en la mateixa saviesa d'allò que anomenem naturalesa, tan propera als estudis del meu departament, o bé en allò que anomenà ecosofia (Panikkar, 2021) el savi català Raimon Panikkar, doctor en química i filosofia. En aquest cas, per trobar-hi sentit, m'he submergit en els darrers anys en l'estudi del blat en l'ambient mediterrani davant la crisi climàtica. En aquesta afortunada casualitat he recorregut la geografia ibèrica i mediterrània tot estudiant-la des del mateix territori o a través de dades percebudes remotament.



Aquest viatge de recerca que he viscut en els darrers anys queda cristal·litzat en les paraules d'aquest text. *El cultiu del blat a l'ambient mediterrani davant la crisi climàtica: perspectives geogràfiques des de la teledetecció i l'ecofisiologia* és el títol d'aquesta tesi. Com es pot imaginar el lector, el cultiu estudiat és el blat al Mediterrani i la teledetecció i l'ecofisiologia són els camps diversos, però recíprocs, des d'on s'estudia. Tot des d'una perspectiva geogràfica doncs els elements que es produeixen en l'espai no es troben aïllats entre si sinó que formen part de la matriu, el medi, que és també tècnica com veurem en la primera part de la introducció.

Aquesta introducció s'organitza en tres punts. Per una banda, el primer punt parla sobre el medi, les característiques del Mediterrani, les dificultats d'aquesta zona davant la crisi climàtica així com els seus paisatges. En aquest primer punt també es desenvolupa epistemològicament una aproximació sobre què és l'espai i com es produeix partint dels treballs del geògraf M. Santos sobre la tècnica i l'espai, així com d'altres autors. Crec que aquesta aproximació és importantíssima perquè actualment les dades, enteses com a infraestructura, produeixen l'espai. Els sensors o els sistemes de monitoratge com a elements materials, però també les dades que recullen com a guia per configurar els paisatges.

Per altra banda, el punt dos de la introducció transcorre sobre les característiques del blat, des de la seva gènesi fins al seu maneig o com es configuren els agro-ecosistemes. A més a més, descriu l'ecofisiologia d'aquest cultiu, és a dir quins efectes té l'ambient en el cultiu i com afecta el seu funcionament. Això és cabdal per poder entendre el rendiment, és a dir la quantitat de blat que es cull, o bé la qualitat, com en el cas del contingut de proteïna en el gra, importantíssima pels processos d'elaboració d'aliments derivats d'aquest cereal mediterrani.

Finalment, el tercer punt és la integració del primer i el segon punt. És a dir, l'ús en aquest cas de tecnologies de la informació geogràfica, principalment de teledetecció satel·lital i altres nivells de mesura, per relacionar-los amb les característiques fisiològiques i agronòmiques del blat. A més a més, també descriu l'ús de bases de dades diverses en l'ambient mediterrani: des de característiques del paisatge, com ara paràmetres del sòl, o bé dades climàtiques. També fa especial èmfasi en el cas del satèl·lit Sentinel-2 i el programa Copernicus de la Unió Europea com a font de dades emprada en els estudis presentats en aquesta tesi. Finalment, aquest darrer punt de la tesi descriu els models avançats, en certa manera, que s'han usat, com el cas de l'aprenentatge automàtic o, en anglès, *Machine Learning* (ML) i altres aproximacions.

Son cinc els capítols de la tesi presentats als resultats. Quatre que ja s'han publicat i un que està en procés d'avaluació i correcció. A part dels cinc capítols al voltant del blat, també presento en

els annexos tres articles més que he fet durant la meva tesi i que considero que han estat rellevants a l'hora d'aprofundir en els coneixements i tècniques treballats en la tesi. A més, ajuden a comprendre el marc de recerca proposat i a compartir millor l'aspecte interdisciplinari que he desenvolupat en els darrers anys. Tanmateix, com es tracten d'articles de revisió (en el cas de l'ús del Sentinel-2 en agricultura, annex 1), o bé de monitoratge en altres cobertures vegetals que no són blat (pastures i bosc, annex 2 i 3) no els he inclòs en els resultats.

Trobo que les iniciatives interdisciplinàries son cabdals per redreçar la crisi climàtica actual. Tanmateix, és sempre un desafiament i, certament, hi ha certes limitacions en la recerca que vol ser interdisciplinària. És sobretot, per mi, un repte perquè la meva recerca parteix d'una feina principalment individual i tot sovint solitària. Com a biòleg de formació, la majoria de coneixements trenats en els estudis d'aquesta tesi provenen de l'autoaprenentatge i del meu propi tarannà curiós, però no pas d'un procés reglat en tots els àmbits de coneixement aquí treballats. Des de l'enginy i aquest afany de conèixer, penso que he pogut desenvolupar, amb els seus encerts i errors, un camp d'estudi suggeridor, un riu regat per la geografia, l'ecofisiologia, l'agroecologia, la teledetecció i la modelització de cultius. Espero que amb les seves limitacions i virtuts aquest riu de rius pugui inspirar i contribuir a enriquir aquest delta, és a dir el camp del coneixement del qual tinc la sort de formar-ne part.

## 1. El medi

### 1.1. Els cultius, el clima mediterrani i els seus canvis

El Mediterrani és un subtipus de clima temperat caracteritzat per estius secs i hiverns suaus i humits. També té unes precipitacions irregulars (Lionello et al., 2006). El blat és un dels conreus mediterranis tradicionals i ha estat el cultiu bàsic de la conca mediterrània des que es va produir la domesticació d'aquest cereal (Peña, 2002). Les diferents condicions ambientals mediterrànies amb què s'ha trobat el cultiu del blat al llarg dels anys han modelat la seva progressiva adaptació genètica a diferents entorns, juntament amb les pràctiques agrícoles i els coneixements bioculturals.

La regió mediterrània ha despertat un interès acadèmic considerable en relació amb la producció, el consum, la sostenibilitat i la diversitat d'aliments que s'hi dona (Blas et al., 2019; García-Martín et al., 2021; Ponti et al., 2016). L'impacte del canvi climàtic a la regió mediterrània s'espera que sigui de gran magnitud, cosa que les dades dels darrers anys ja corroboren, a causa del descens substancial previst de les precipitacions i l'augment de la temperatura durant les pròximes dècades (Cos et al., 2022). Es preveu que aquesta reducció de la precipitació produirà un fort

estrès als cultius i ecosistemes mediterranis que ja es troben actualment en perllongats períodes de sequera (Tuel i Eltahir, 2020). A més, com a conseqüència del canvi climàtic, el clima mediterrani pot ampliar el seu abast geogràfic i estendre's en altres zones temperades europees (Beck et al., 2018) amb un increment de zones àrides a la península Ibèrica i d'altres regions.

En aquesta tesi el treball de camp s'ha centrat en la península Ibèrica, però també s'han utilitzat dades en l'àmbit euromediterrani. La distribució geogràfica del cultiu del blat i els canvis en els paisatges agrícoles mediterranis s'han emprat per poder avaluar formes de maneig més sostenibles dels agro-ecosistemes entesos a nivell d'eco-regió. És a dir, àrees específiques amb una distribució de flora (de la qual prenen el nom, Figura 1) que contenen un conjunt diferenciat de comunitats naturals i espècies dins d'un límit que s'aproxima al seu estat original abans de canvis d'ús del sòl significatius (Olson et al., 2001).



**Figura 1.** Les regions euro-mediterrànies, basat en Olson et al. (2001).

## 1.2. La tècnica i l'espai

Des que l'agricultura va començar fa milers d'anys, els humans han afaïçonat la Terra tot produint paisatges agrícoles a la majoria dels hàbitats on vivim avui o on hem viscut en l'antigor. Les feixes d'arrossars a l'Àsia, les illes hidropòniques de blat de moro d'alt rendiment a Mèxic (*chinampas*)

o els sistemes agro-forestals mediterranis han articulats tradicionalment els paisatges agrícoles en les seves respectives regions. La tècnica i el medi geogràfic canvien simultàniament amb l'acció humana (Santos, 2000) en diàleg amb l'ambient, i així passa actualment amb una tècnica que ha anat més enllà de l'ús de recursos i espècies locals per a l'agricultura i s'ha convertit en una activitat global altament avançada tecnològicament i productiva, com moltes altres a la societat.

L'espai i la naturalesa es modifiquen a través de l'evolució de la tècnica, el concepte de tècnica va més enllà del de tecnologia i es defineix com la principal forma de relació entre la humanitat i el medi. En l'actualitat, el medi es caracteritza per una dimensió global fruit del seu aspecte tècnic, científic i d'informació. En aquest sentit, la capacitat actual d'observació de la Terra (teledetecció) genera quantitats massives de dades obtingudes amb satèl·lit (Lang et al., 2019; Sudmanns et al., 2020), que s'utilitzen per monitorar i aprofitar els recursos naturals en multitud de dominis ambientals com ara els boscos (Lechner et al., 2020) o conreus (Segarra et al., 2020; Weiss et al., 2020), així com altres processos físics i socials. Aquest ús de la teledetecció està inscrit en el territori perquè els paisatges agrícoles, com tants d'altres, es configuren, cada cop més habitualment, a través de les dades.

Les dades s'han convertit en un aspecte central de la política ambiental (Nost i Goldstein, 2022), així com en l'ús dels recursos. L'enorme capacitat dels sistemes de teledetecció d'adquirir dades fa prioritari el seu ús en la sostenibilitat agrícola i ambiental. En aquest sentit, seguint el flux de dades, es pot aplicar en el maneig agrícola i ambiental un enfocament mediat digitalment. Per exemple, les dades obtingudes amb teledetecció s'utilitzen cada cop més en activitats econòmiques i en el monitoratge dels recursos de la Terra, però també en la sostenibilitat dels agro-ecosistemes (Segarra, 2024). Així doncs, es pot afirmar que la infraestructura de dades satel·lital és avui dia inseparable del medi (Bakker and Ritts, 2018) i, per tant, d'una producció sostenible dels paisatges agrícoles.

El lliure accés a les dades fa que la informació estigui disponible per un ús molt extens. Per exemple, aquestes dades es poden fer servir potencialment per controlar els impactes ambientals i fer estimacions per un benestar ecològic i social més equilibrat, com en el cas que ens ocupa. No obstant això, en aquesta dialèctica les dades satel·litals d'accés obert s'usen també per incrementar asimètricament llocs d'extracció de recursos, com en el cas de la mineria d'or (Nagaraj, 2022), perquè les dades de lliure accés fan augmentar les perspectives d'ingressos tot reduint-ne el risc d'inversió en exploració.



A part de les dades de satèl·lit d'accés públic, algunes empreses tenen sistemes de teledetecció com ara eixams de satèl·lits (petits satèl·lits de l'empresa Planet Lab), que proporcionen imatges multispectrals d'alta resolució diàriament. En aquest context, les imatges diàries i molt detallades de la superfície de la Terra solen estar disponibles a preus elevats a través d'empreses privades. En aquest sentit, empreses com Orbital Insight Analytics utilitzen imatges d'alta resolució de les corporacions Planet Lab i Maxar Technologies per oferir informació avantatjosa a les empreses. Mukherjee et al. (2021) va observar que les dades de teledetecció analitzades per Orbital Insight Analytics oferien estimacions millorades per a les perspectives del mercat de la indústria petroliera abans de la publicació d'altres fonts de dades. Això també succeeix amb empreses que gestionen, processen i venen dades agrícoles per poder millorar fertilitzacions i anticipar rendiments a l'agricultor. Aquestes iniciatives de teledetecció contribueixen a generar activitats orientades a l'ús dels recursos de la Terra (Alvarez Leon, 2022) i, per tant, configuren els paisatges d'una manera concreta. És aquí on s'inscriu la tècnica i el medi com a elements fosos en el paisatge, en aquest cas una tècnica que és infraestructura de dades, tangible i intangible, però productora de l'espai en ambdós casos.

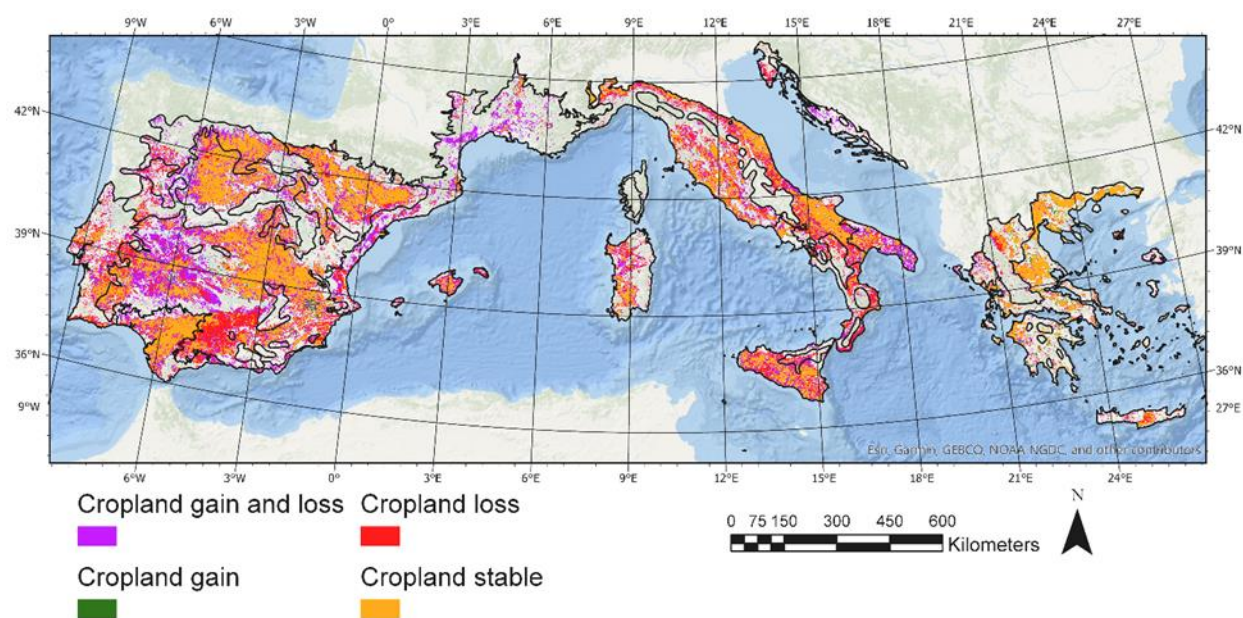
### 1.3. Els paisatges i les terres marginals

El cultiu del blat (*Triticum* spp.) ha configurat els paisatges mediterranis durant els darrers deu mil anys i ha sostingut la nutrició de la seva població. Un paisatge agrícola mediterrani típic es caracteritza tradicionalment per parcel·les de cereals de secà, oliveres i vinyes, entre d'altres. Les feixes a les terres marginals muntanyoses i uns paisatges agrícoles heterogenis han caracteritzat la regió durant els darrers segles (Blondel, 2006).

D'ençà mitjan segle XX, les observacions amb teledetecció dels canvis en els cultius o els canvis en el paisatge poden orientar la implementació de polítiques per a la recuperació dels paisatges tradicionals o la implementació de nous usos més sustentables, com s'ha fet en diferents casos (González-Fernández et al., 2022; Osorno-Covarrubias et al., 2018). Captar imatges amb drons, avions o satèl·lits al llarg del temps permet controlar els canvis temporals dels paisatges mediterranis (Lasanta and Vicente-Serrano, 2012; Zomeni et al., 2008) així com el seu abandonament (Segarra et al., 2023). A més, les dades de teledetecció permeten monitorar la cobertura vegetal i comprendre l'evolució del conreu a través de paràmetres biofísics, de biomassa, d'estat hídric o fertilitat del sòl, entre d'altres. Aquesta informació permet tenir una millor gestió dels cultius (von Keyserlingk et al., 2021) tot definint el maneig més adequat segons les característiques espaciotemporals del conreu.

La Unió Europea té previst assolir el 25% del total de les terres de conreu europees en producció ecològica per a la propera dècada (2030), tal com s'indica a la declaració de la Comissió Europea titulada "Pla d'acció per al desenvolupament de la producció ecològica a la UE" (2021). Per tant, trobar terres de conreu marginals pot ajudar a definir àrees amb menys diferència de producció quan es comparen l'agricultura ecològica amb la convencional. Les zones agrícoles més òptimes es troben generalment en terres de conreu amb una agricultura més intensiva. També s'hi troba una concentració de les millors tecnològiques més avançades, mentre que les terres de conreu marginals es troben sovint en entorns amb condicions agroecològiques que limiten la seva productivitat (Ceccarelli, 1994).

A les zones marginals, l'erosió del sòl, les pluges i la temperatura no compleixen les condicions òptimes per al desenvolupament dels cultius, la qual cosa resulta en rendiments més baixos (Zampieri et al., 2020). Des de la implantació de l'agricultura industrial a mitjan segle XX (aplicació d'avenços en les zones agrícoles més productives amb rendiments millorats), aquestes terres de conreu han estat progressivament abandonades a diverses regions mediterrànies (Jiménez-Olivencia et al., 2021) tal com s'observa al mapa global de canvis en l'ús de les terres de conreu elaborat per (Winkler et al., 2021) i readaptat per mi en l'àmbit de l'euromediterrània a la Figura 2.



**Figura 2.** Dinàmica de les terres de conreu 1960-2019 basat en el conjunt de dades HILDA+ (Winkler et al. 2021) per a la regió euromediterrània. El guany i la pèrdua de terres de conreu (en violeta) implica múltiples esdeveniments de canvi dins del període estudiat, el guany de les terres de conreu (en verd) es

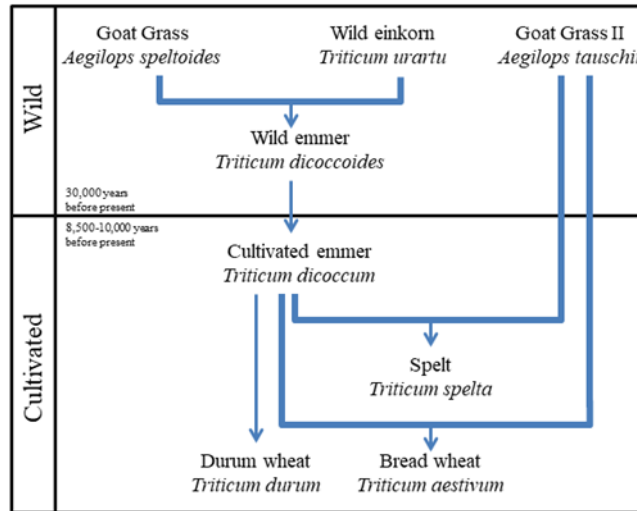
refereix a l'expansió espacial de les terres de conreu a zones no agrícoles, la pèrdua de terres de conreu (en vermell) es refereix a la pèrdua d'àrea agrícola, mentre que les terres de conreu estable (en taronja) fa referència a les terres agrícoles estables.

Per tant, definir zones marginals juntament amb la idoneïtat geogràfica del cultiu de blat, és a dir, l'avaluació climàtica, edàfica o topogràfica permet aplicar manejos agrícoles adequats a entorns específics. Aquest enfocament es pot aplicar a altres regions mediterrànies i parts del món, adaptant la metodologia als conjunts de dades ambientals i de teledetecció rellevants disponibles. A més, aquest enfocament pot jugar un paper en evitar desajustos en la producció i entendre el paisatge com una unitat geogràfica produïda de manera més o menys sostenible i més o menys equilibrada. Tanmateix, cal abordar algunes limitacions pel que fa a la integració de diversos conjunts de dades geogràfiques i la governança de les pròpies dades, és a dir qui hi té accés i qui en pot fer ús.

## 2. L'ecofisiologia i la singularitat del blat

### 2.1. Característiques

El blat (*Triticum* spp.) és un cereal actualment àmpliament conreat a tot el món que es va originar al Creixent Fèrtil, fa més de 10.000 anys. Posteriorment, es va estendre a la regió euromediterrània arribant primer als Balcans i a Grècia (fa uns 8.000 anys) i, més tard (fa uns 7.000 anys), a les penínsules Itàlica i Ibèrica (Royo et al., 2017). Diverses espècies de blat, tant silvestres com domesticades, conformen el gènere *Triticum*. Els dos tipus més comuns són el blat comú hexaploide (*Triticum aestivum* L.) i el blat dur tetraploide (*Triticum durum*). A més d'aquests dos, es conreen d'altres tipus de blat com l'espelta (*Triticum spelta*) i la pisana diploide (*Triticum monococcum*) en menor mesura. L'evolució del blat comú es va formar a partir d'un híbrid entre el tetraploide *Triticum turgidum* i el diploide *Aegilops tauchii* (Figura 3). El blat comú és, amb diferència, el blat més conreat a nivell mundial i ocupa zones importants a l'euromediterrani; el blat dur és el tipus de blat més important al sud i l'est de la conca mediterrània i també és molt important al nord de la Mediterrània, només superat pel blat comú.



**Figura 3.** Evolució genètica del blat, espècies silvestres (wild) i cultivades (cultivated).

Actualment el blat comú (*Triticum aestivum* L.) ocupa una part important dels conreus a la península Ibèrica i representa una part significativa de la producció europea (d'Andrimont et al., 2021). El blat dur, és també molt important quant a superfície i producció. La major part de la producció de blat es localitza al centre (blat comú) i sud (blat dur) de la península Ibèrica, que correspon a les regions de Castella i Andalusia.

## 2.2. Rendiment i qualitat

El rendiment, la quantitat de gra collit, i la qualitat del blat, en aquest cas centrada en el percentatge de proteïna (nitrogen) del gra, depenen de factors genètics i ambientals i tenen una gran importància socioeconòmica. A partir de tècniques de teledetecció o d'anàlisi d'isòtops estables es pot estimar tan el rendiment com els paràmetres de qualitat en el blat. Pel que fa al rendiment, hi ha dues maneres principals d'estimar-lo a partir de dades de teledetecció. D'una banda, models empírics basats en dades de camp (quantitat recol·lectada) i índexs de vegetació (combinacions de bandes concretes de l'espectre de reflectància) s'han utilitzat amb èxit amb dades de Sentinel-2 i estimació de rendiment en blat (Cavalaris et al., 2021).

Per l'altra banda, models de creixement (Spitters, 1989; Weiss et al., 2020) que impliquen estimacions de paràmetres fisiològics basats en la complexitat de la reflectància de les estructures vegetals i modelitzat inversament en *radiative transfer models* (RTM). La disponibilitat de dades (Burke i Lobell, 2017) i els requisits de processament informàtic necessaris per a la modelització del creixement dels cultius fan que els models empírics siguin més versàtils per



utilitzar-los a la pràctica per a conjunts de dades limitats. Tot i tenir més variabilitat interanual i menys robustesa, en certa manera, que els RTM.

Per tal d'escalar les prediccions de rendiment i qualitat de nivell de camp a nivell regional, el mapatge de tipus de cultius és essencial, ja que pot definir els camps als quals aplicar els models de predicció de rendiment o qualitat. Tant les classificacions com les estimacions de superfície de conreu de blat són de gran interès per millorar la gestió agrícola. L'estudi del rendiment a escala regional també permet analitzar els llocs més adequats per al cultiu de blat tot explorant quins factors espacials afecten el rendiment. Tot i això, molt sovint, estimar el rendiment a nivell de camp és laboriós i complex. Un factor és la disponibilitat d'informació de rendiments agrícoles per parcel·la o de la pròpia variabilitat dins la parcel·la. Això es pot considerar informació sensible i, fins i tot, difícil d'obtenir a causa dels interessos del mercat i altres factors socioeconòmics. A més, un altre factor limitant és fer mesures de camp per part dels investigadors (segar, batre o pesar seccions de la parcel·la) amb el vistiplau del pagès. Això pot comprometre la disponibilitat de dades d'entrenament i validació per als models d'estimació de rendiment.

A més de les condicions de creixement dels cultius, aquests models depenen dels contextos geomorfològics locals (és a dir, les característiques topogràfiques) (Matsushita et al., 2007). Defourny et al. (2019) també van observar, amb dades de teledetecció, que un gradient agroclimàtic que abasti un territori ampli fa canviar gradualment el calendari de desenvolupament del cultiu, així com la distribució del tipus de cultiu, i en conseqüència afecta el rendiment dels cultius. Així doncs, la capacitat de la teledetecció de monitorar la fenologia del blat, per exemple, és un factor clau a l'hora d'entendre el rendiment potencial i el desenvolupament del cultiu. Això, alhora, implica definir millor el maneig més adequat en una localitat específica. Poder estimar el rendiment en etapes fenològiques primerenques permet aplicar manejos específics per part del pagès que incrementin el rendiment final. A més, poden millorar l'aplicació de tractaments tot definint geogràficament les zones dins de la parcel·la amb més mancances. Allò que es coneix com agricultura de precisió.

Quant al percentatge de proteïna en el gra de blat, aquest està determinat per factors genètics i ambientals (sobretot la disponibilitat de fertilitzant i aigua (Daniel i Triboí, 2002; Lloveras et al., 2001; Serrano et al., 2002)) i està directament relacionat amb el percentatge de nitrogen al gra. És possible obtenir el percentatge de nitrogen a través de l'anàlisi d'isòtops de nitrogen amb un analitzador elemental. L'acumulació del nitrogen en els grans de blat, en forma de proteïna, comença a produir-se un cop el sembrat ja ha guaijat i espigat i va entrant a la maduresa tot omplint el gra de l'espiga. És a dir, la concentració de nitrogen de les plantes comença a baixar

en l'etapa d'ompliment del gra. En el blat, entre el 60 i el 95% del nitrogen es re-mobilitza de les fulles i les tiges cap als grans, sent la font més important de nitrogen en els grans (Palta i Fillery, 1995; Papakosta i Gagianas, 1991) i, per tant, de proteïna. Poder captar les fases de remobilització del nitrogen en la planta i entendre l'estat del cultiu permet poder adreçar manejos que assegurin uns percentatges òptims de proteïna en el gra.

### 2.3. Agrosistemes cerealístics i el seus manejos

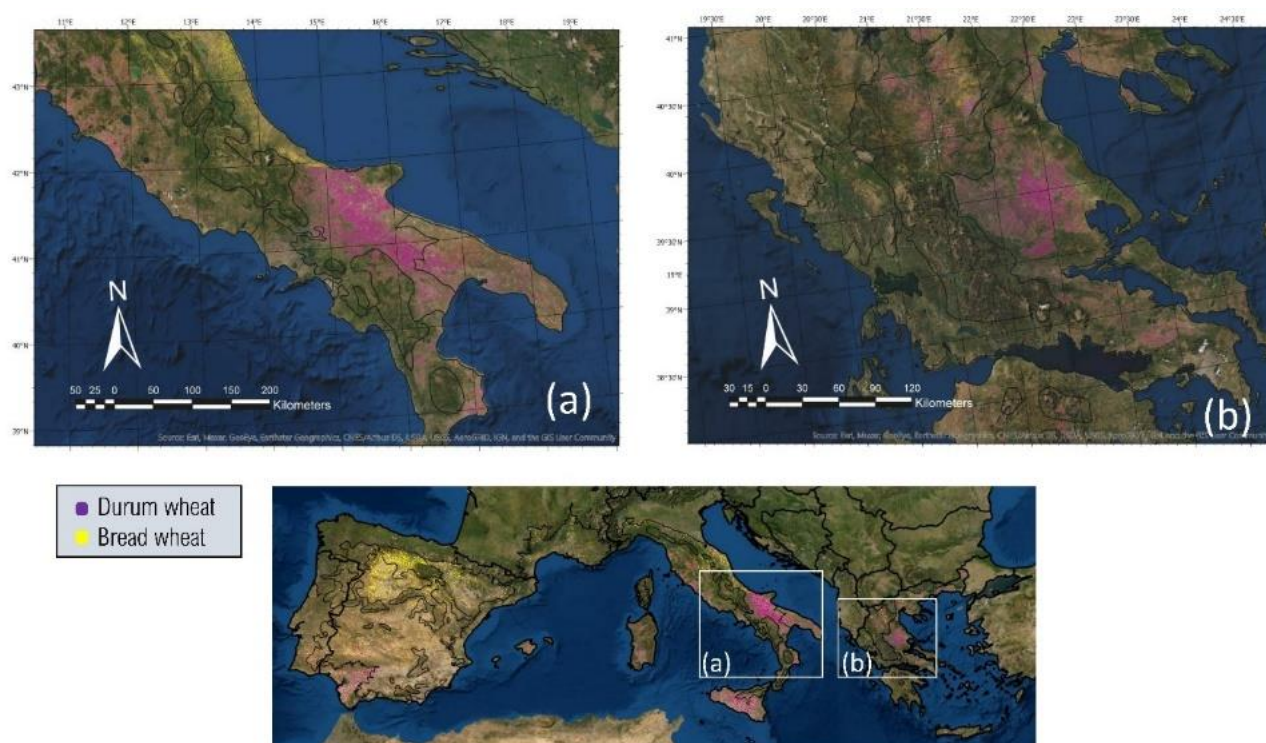
En el paradigma existent de l'agricultura industrial, assumint que no hi hagi alteracions radicals en les tecnologies agronòmiques i de millora genètica, es pot inferir que els rendiments agrícoles a la regió mediterrània disminuiran com a conseqüència de les condicions ambientals. Això implica que es requerirà una ampliació de la superfície de les terres agrícoles conreades. A més, caldran canvis en la dieta actual, o bé una recuperació de la vella dieta mediterrània que ja es troba malauradament desdibuixada del seu origen. En aquest sentit, s'espera que hi hagi una disminució del 15 al 30% de rendiment del blat com a conseqüència d'un augment de 2 °C de la temperatura (Bento et al., 2021) prevista per a l'any 2040, sinó abans. Uns rendiments esquistats ja es van veure en la temporada 2022-2023 on a molts llocs ni es va recol·lectar el gra. A part del factor ambiental, un dels factors que ha fet que les collites s'hagin estancat en els darrers temps (Le Gouis et al., 2020) és parcialment atribuït a l'absència d'un progrés genètic significatiu, especialment en el blat dur (Chairi et al., 2018).

A la Mediterrània europea, així com a d'altres regions del món, els sistemes de producció de cultius han experimentat importants transformacions des de mitjan segle XX, passant a un sistema industrialitzat caracteritzat per una major productivitat agrícola. Aquests avenços han tingut un paper crucial per combatre la manca d'aliments (Evenson and Gollin, 2003). No obstant això, el sistema alimentari actual i les pràctiques predominants de l'agricultura industrial, que es caracteritzen per un mercat globalitzat i una utilització generalitzada d'inputs químics en la producció agrícola, produeixen efectes negatius sobre la sostenibilitat ambiental (Rounsevell et al., 2003). Aquests efectes es manifesten en alteracions negatives dels cicles dels nutrients (Cameron et al., 2013), esgotament dels recursos hídrics (Rosegrant et al., 2009) i pèrdua de diversitat biològica (Lark et al., 2020). Aquestes conseqüències ambientals representen una amenaça per als recursos globals, amb efectes adversos sobre els ecosistemes i el benestar humà (Willett et al., 2019).

En relació amb aquesta qüestió, el sector agrícola és responsable de més del 20% de les emissions antropogèniques de gasos d'efecte hivernacle a tot el món, tenint en compte els canvis d'ús del sòl (Tubiello et al., 2015). A més, ocupa al voltant del 40% de la superfície terrestre

accessible (Foley et al., 2005) i requereix l'ús del 70% del subministrament global d'aigua dolça (Rosegrant et al., 2009). Per tant, la producció d'aliments és àmpliament reconeguda com un dels principals catalitzadors de la transformació ambiental global i alhora es veu profundament afectada pel canvi climàtic com a conseqüència de la disminució de la disponibilitat de l'aigua i les altes temperatures, entre d'altres.

Com s'observa en la Figura 4, la superfície de blat, tan comú com dur, és molt extensa en l'euromediterrània. A la península Ibèrica destaca l'estepa cerealística a Castella, principalment de blat comú, i importants extensions de blat dur a Andalusia. Agrosistemes de blat dur també es troben extensament al sud d'Itàlia, principal productor mundial de blat dur, i a Grècia. En si el mapa de la Figura 4 il·lustra els agrosistemes cerealístics a l'euromediterrània i alhora la importància d'aquestes extenses geografies i, per tant, del seu rol en preservar i impulsar la sostenibilitat ambiental en l'euromediterrània.



**Figura 4.** Àrees de cultiu de blat a les regions euromediterrànies a partir de (d'Andrimont et al. 2021). La imatge A mostra el sud i centre italians i, mentre que la imatge B zones de Grècia. Els píxels grocs corresponen a blat comú, mentre que els píxels morats corresponen a píxels de blat dur.


### 3. Comprensió del medi i teledetecció




#### 3.1. Satèl·lits, drons i tècniques de mesura fisiològica vegetal (creixement i estat hídric)

La tècnica de teledetecció consisteix en capturar la llum reflectida de la superfície terrestre mitjançant sensors instal·lats en satèl·lits que orbiten la Terra, en avions o en drons. També amb càmeres instal·lades en extensors o imatges preses directament de forma manual (Taula 1). Durant les últimes tres dècades, la capacitat de captar imatges i extreure'n dades ha augmentat moltíssim, com per exemple pel que fa a la quantitat d'imatges de satèl·lit de resolució millorada disponibles públicament (Union of Concerned Scientists, 2023). En comparació a fa unes dècades, avui dia la teledetecció de la superfície terrestre permet distingir objectes més petits (resolució espacial millorada) amb una major freqüència. Actualment es disposa de sensors dissenyats per detectar la llum en dominis concrets i ben definits de l'espectre electromagnètic. Les dades de satèl·lit més habituals són sensors en el rang espectral del visible (llum reflectida que és visible per l'ull humà), com podria ser una càmera digital. També és comú disposar de dades del rang visible i infraroig proper, el rang visible i infraroig, que inclou ones curtes, les bandes infraroges i infrarojos tèrmics, així com la gamma de microones.

Les imatges en el visible es poden adquirir amb alta resolució i són útils per identificar i delimitar amb precisió els objectes de la superfície terrestre (categories d'ús del sòl, extensió de masses d'aigua, etc.). Les imatges en el rang del infraroig solen estar disponibles amb menys resolució espacial però permeten, en combinació amb les dades adquirides en el rang del visible, el càlcul d'índexs de vegetació, útils per a la detecció de canvis en la superfície terrestre i canvis en la vegetació al llarg del temps (Adams and Gillespie, 2006). A més, el senyal reflectit en l'espectre vermell i infraroig proper sobre les superfícies de l'aigua s'incrementa amb partícules en suspensió, la qual cosa pot permetre l'estimació de la concentració de contaminants (Doxaran et al., 2002).

**Taula 1.** Nivells de mesura i instruments.

Level	Spectral data	Sensor
Ground	Spectral data	<b>GreenSeeker</b> (60 cm spatial resolution)
	Multispectral	
	RGB	<b>Sony ILCE-QX1</b> (0.01 mm/pixel)

		
		<b>Lumix GX7</b> (5 cm/pixel)
	RGB	
UAV		<b>Tetracam micro-MCA</b> (5 cm/pixel)
	Multispectral	
Satellite	Multispectral	<b>Sentinel-2 a + b</b> (10 and 20 m/pixel)

Un altre eina de teledetecció és el radar d'obertura sintètica (SAR de les sigles en anglès) en el rang espectral de les microones, en contrast amb els sensors en el visible i infraroig, el SAR funciona en gairebé totes les condicions atmosfèriques perquè la llum de microones pràcticament no es veu afectada pels núvols i el fum. En aquesta tesi he treballat amb dades derivades de SAR pel mapeig de tipus de cultius però no he emprat directament les dades de radar en els models d'estimació de rendiment o qualitat del blat.

A part de la teledetecció al nivell de satèl·lit i dron amb dades multiespectrals, també s'han emprat mesures a nivell de planta (Taula 1), com imatges amb càmeres digitals i mesures complementàries com el Greenseeker (mesura d'un índex de vegetació sensible a les superfícies fotosintèticament actives). A més a més, també s'han emprat isòtops estables. He estudiat el rendiment i la qualitat del gra de blat en règims hídrics i condicions de fertilització contrastades, caracteritzades mitjançant composicions isotòpiques de C i N dels grans. La composició d'isòtops de C ( $\delta^{13}\text{C}$ ), quan s'analitza en els grans de blat madurs, és un indicador de l'estat hídric durant el creixement del cultiu (Araus et al., 2013, 2003), especialment singular pel cas de la Mediterrània. Pel que fa a l'abundància natural de la composició d'isòtops de N ( $\delta^{15}\text{N}$ ) en grans madurs, les dades indiquen la font de fertilització nitrogenada del cultiu (Serret et al., 2008).



### 3.2. Missió Copernicus i el Sentinel-2

Durant els últims 50 anys d'evolució digital, s'han posat en òrbita nombrosos satèl·lits amb sensors multiespectrals amb el propòsit d'observar remotament la Terra. Actualment, un dels programes d'observació de la Terra més rellevants centrats en la detecció i avaluació de la vegetació és el Programa Copernicus. Aquest programa està liderat per l'agència espacial europea (ESA en anglès). És un programa que té com a objectiu augmentar les capacitats d'observació amb satèl·lit de la Terra i dins del qual s'han posat en òrbita les plataformes Sentinel-2 A i B. A més del sistema multiespectral Sentinel-2, s'han llançat o està previst llançar diversos satèl·lits com a part d'aquest programa encara en curs. Pel que fa a possibles aplicacions agrícoles dins del programa Copernicus, cal esmentar Sentinel-1 (<https://sentinel.esa.int/web/sentinel/missions/sentinel-1>) que porta a bord un radar d'obertura sintètica així com Sentinel-3 que proporciona dades dels infrarojos tèrmics a 1 km de resolució espacial (<https://sentinel.esa.int/web/sentinel/missions/sentinel-3>).

Comprendre les característiques dels satèl·lits és especialment rellevant per comprendre les millores que ofereixen les dades de Sentinel-2 en comparació amb altres satèl·lits en òrbita per les aplicacions agrícoles i ambientals. En aquest sentit, és fonamental entendre les escales d'observació i resolució dels sensors dels satèl·lits, ja que determinen les capacitats de monitoratge agrícola. Pel que fa a la resolució, quatre paràmetres són centrals per aproximar-se a una millor comprensió dels objectius i l'abast d'aquests instruments. En primer lloc, la resolució espectral, és a dir, el nombre, l'amplada i la franja de l'espectre electromagnètic de les bandes espectrals del sistema de sensors. En segon lloc, la resolució espacial, que és la mesura de l'objecte més petit que pot ser distingit pel sensor, és a dir la mida del píxel. En tercer lloc, cal tenir en compte la resolució temporal que és una mesura de la freqüència amb què un sensor torna a visitar la mateixa àrea de la superfície terrestre; i finalment, la resolució radiomètrica, que ve determinada pel nombre de bits en què es divideix la radiació registrada (Liang et al., 2012) i determina la precisió espectral de les mesures.

Sentinel-2 és una constel·lació de dues plataformes idèntiques: Sentinel-2A que es va posar en òrbita el 23 de juny de 2015 i Sentinel-2B en òrbita d'ençà el 7 de març de 2017, formant, a partir d'aleshores, la constel·lació Sentinel-2 A + B (<https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/sentinel-2>). Tots dos satèl·lits tenen una vida útil de 7,25 anys, inclosa una fase de posada en servei de 3 mesos amb un propulsor de 12 anys de durada. Tant els satèl·lits Sentinel-2A com Sentinel-2B tenen a bord exactament el mateix instrument multiespectral (MSI) amb bandes des del visible fins a l'infraroig d'ona curta (Taula 2): quatre (4)

bandes a 10 m, la clàssica banda ampla visible blava (490 nm), verda (560 nm), vermella (665 nm) i l'infraroig proper (842 nm); sis bandes a 20 m, quatre bandes estretes al domini espectral de la vora del vermell, rellevant per la vegetació (705, 740, 775 i 865 nm) i dues bandes d'infraroig d'ona curta (1610 i 2190 nm); i tres (3) bandes a 60 m dedicades a la correcció atmosfèrica (443 nm per als aerosols i 940 nm per al vapor d'aigua) i a la detecció de cirrus (núvols alts) (1380 nm) (Taula 2). Després del segon llançament (Sentinel-2 B), el satèl·lit proporciona una resolució temporal més alta amb una freqüència de 5 dies i fins a 2 dies a les parts superiors del nord i sud del Planeta.

**Taula 2.** Bandes multiespectrals del Sentinel-2

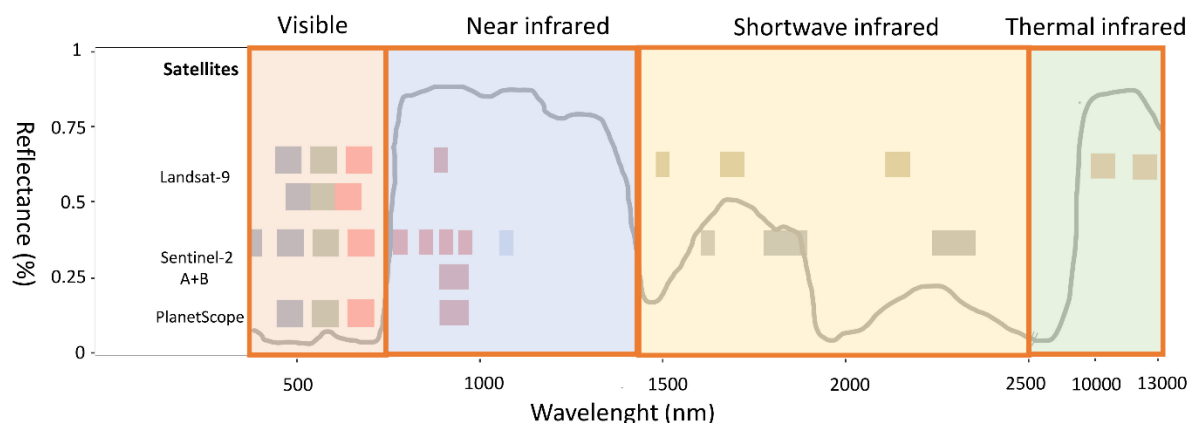
MSI Banda	Resolució espacial (m)	Longitud d'ona central (nm)	Amplada (nm)
B1: Coastal Aerosol	60	443	20
B2: Blau	10	490	65
B3: Verd	10	560	35
B4: Roig	10	665	30
B5: Red-Edge	20	705	15
B6: Red-Edge	20	740	15
B7: Red-Edge	20	783	20
B8: NIR	10	842	115
B8A: Vegetation RE	20	865	20
B9: Vapor d'aigua	60	945	20
B10: SWIR Cirrus	60	1375	30
B11: SWIR	20	1610	90
B12: SWIR	20	2190	180

### 3.3. La fisiologia amb la teledetecció: els models

Les plantes interaccionen amb la llum solar, l'espectre complet de la radiació electromagnètica emesa pel sol, de manera diferent. Això ho detectem depenent de la longitud d'ona observada. En aquest sentit, la radiació solar incident pot seguir principalment tres vies: pot ser transmesa, reflectida o absorbida. La radiació electromagnètica que reflecteixen les plantes conté informació sobre la seva composició biofísica i fisiologia i es pot mesurar mitjançant sensors com els del Sentinel-2. Centrant-nos en el domini visible, els pigments de clorofil·la presents a les fulles verdes absorbeixen fortament a la regió visible de l'espectre (400-660 nm), especialment a les

longituds blava i vermella, on es captura l'energia per a la fotosíntesi. Mentrestant, a l'infraroig proper, aproximadament dels 700 als 1300 nm, les fulles presenten alts valors de reflectància i transmissió, principalment relacionats amb les propietats estructurals de les fulles i la biomassa, i absorbeixen menys radiació en aquesta regió espectral (Avery and Berlin, 1992; Tucker, 1979). A més, l'estructura del dosser de la planta i l'àrea de les fulles també són trets fonamentals relacionats amb els patrons de reflectància del dosser (Disney et al., 2006; Rautiainen and Stenberg, 2005) i són paràmetres clau per controlar el creixement. Pel que fa a la regió infraroja d'ona curta, aproximadament dels 1300 als 2500 nm, l'absorció de la radiació està en gran part dominada per l'aigua, seguida d'altres components bioquímics presents a les fulles. En la Figura 5 es mostren diferents característiques dels satèl·lits i com la informació espectral es pot relacionar amb aplicacions concretes en agricultura. La combinació de les bandes espectrals permet el càlcul d'índexs de vegetació específics relacionats amb diferents característiques del cultiu com la biomassa, el contingut en clorofil·les o l'estat hídric.

Així doncs, pel que fa a l'agricultura, els cultius interactuen amb el seu entorn (variacions climàtiques estacionals, esdeveniments meteorològics extrems, plagues, propietats del sòl, etc.), creixen i passen per diferents estadis fenològics. Les interaccions entre les plantes i la reflectància de la llum es tradueixen en canvis en els patrons de senyal que es poden interpretar mitjançant dades obtingudes amb teledetecció. L'ús d'instruments de teledetecció per controlar els canvis de reflectància de la radiació electromagnètica als cultius està ben demostrat a la literatura científica (Blazquez and Edwards, 1986; Chen et al., 2010; Kokaly, 2001). Per exemple, el contingut de clorofil·la té un paper directe en la fotosíntesi i canvia com a conseqüència de l'estrès, per tant s'utilitza sovint com a indicador general de salut vegetal (Gitelson and Merzlyak, 1996) i s'interpreta juntament amb els senyals relacionats amb la biomassa i el contingut d'aigua. Aquests són exemples d'usos de les propietats de reflectància espectral de les plantes per al seguiment de la vegetació. Tanmateix, sovint són les propietats temporals i espacials dels satèl·lits les que determinen factors rellevants pel seu ús a l'agricultura. Això és conseqüència de les característiques dels conreus, que creixen amb més rapidesa i freqüència en parcel·les més petites i gestionades de manera activa que no pas en la vegetació natural menys intervinguda per l'home.

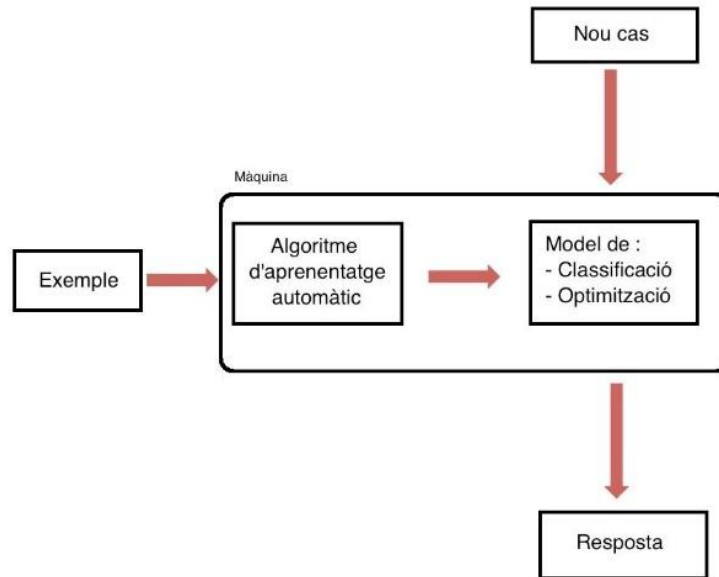


Target trait	Spectral Information	Applications
Field vegetation cover and greenness	Visible and near infrared	Stress detection Canopy cover Leaf area index Growth dynamics Senescence Crop greenness Agronomic and yield traits Plant emergence Phenology Leaf nitrogen content
Chlorophyll content	Visible, near infrared, and red edge	Crop greenness Stress detection Chlorophyll content Leaf nitrogen content Grain nitrogen content
Photosynthesis and yield	Visible and near infrared	Photosynthetic status Biomass and yield Senescence Chlorophyll content Stress detection
Water content	Shortwave infrared and thermal infrared	Water status Evapotranspiration Stress detection

**Figura 5.** Signatura espectral d'un cultiu amb les bandes disponibles per als sensors de satèl·lit (Landsat 9, Sentinel-2 i PlanetScope). S'indiquen les regions espectrals i es descriuen els trets agronòmics i fisiològics en relació amb la seva aplicació en agricultura i monitoratge de plantes.

Els enfocaments de modelització com el ML han mostrat un potencial considerable en aplicacions de teledetecció agrícoles (Chlingaryan et al., 2018). El ML consisteix en la generació de programes capaços de generalitzar comportaments a partir del reconeixement de patrons o classificacions per tal de generar predictors, Figura 6. Aquests algorismes informàtics de ML són especialment útils per estudiar sistemes biològics complexos, ja que poden capturar interaccions complexes entre variables i trobar patrons predictius generalitzables (Bzdok et al., 2018). L'ús de dades multiplataforma per desenvolupar models que permetin estimar el rendiment del blat abans de la collita és fonamental per guiar la gestió agrícola, tant per garantir el rendiment dels cultius

com per optimitzar l'ús dels recursos (Foley et al., 2011). Els enfocaments de ML analitzats presenten diferents models per analitzar les aproximacions més adequades pel monitoratge del rendiment i la qualitat del blat així com la classificació de cultius. L'ús de ML per estimar rendiments és considerat com una de les àrees de desenvolupament més importants associades a la teledetecció i l'agricultura (Weiss et al., 2020).



**Figura 6.** Exemple de funcionament del Machine Learning (ML), aprenentatge automàtic. Font: Viquipèdia



## Objectius

L'objectiu general de la tesi és emprar dades de teledetecció i sistemes d'informació geogràfica per estimar i comprendre el rendiment i qualitat de varietats de blat mediterrani tot considerant els efectes ambientals, del paisatge i el maneig agrícola.

Per dur a terme aquest objectiu, es van proposar els objectius específics següents:

1. Desenvolupar models de teledetecció multiescala per estimar els paràmetres de rendiment i qualitat.
2. Desenvolupar classificacions de tipus de cultius per tal de mapejar els camps i cultius agrícoles.
3. Comprendre els efectes ambientals, topogràfics i de maneig a nivell de camp i regional sobre el rendiment del blat i la qualitat.
4. Explorar els canvis d'ús del sòl i els efectes del paisatge en el cas del blat mediterrani.
5. Definir la idoneïtat del cultiu del blat pel que fa a les característiques regionals en diferents àrees climàtiques.
6. Entendre si cal una millora vegetal per un ambient específic o si els genotips tenen una plasticitat genètica que els permet adaptar-se a diferents agroclimes.
7. Avaluar les diferències entre el maneig convencional i ecològic del blat per definir zones d'idoneïtat del maneig.

## Informe dels directors



UNIVERSITAT DE  
BARCELONA

Departament de Biologia Evolutiva,  
Ecologia i Ciències Ambientals

Facultat de Biologia

Avda. Diagonal, 643  
08028 Barcelona

Fax +34 934 112 842  
Tel. +34 934 035 738

Els doctors Jose Luis Araus i Shawn C. Kefauver, com a directors de la tesi doctoral titulada **“El cultiu del blat a l’ambient mediterrani davant la crisi climàtica: Perspectives geogràfiques des de la teledetecció i l’ecofisiologia.”** presentada pel doctorand Joel Segarra informen sobre el factor d’impacte i la participació del doctorand en cadascun dels articles inclosos en la memòria d’aquesta tesi doctoral.

De les cinc publicacions que conformen el cos principal de resultats de la tesis doctoral d’en Joel Segarra, la primera està en fase d’avaluació i les altres quatre s’han publicat en revistes del primer quartil, i en la majoria de casos del primer decil, dins les especialitats d’agricultura, ecologia terrestre, teledetecció, etc. En totes elles el paper d’en Joel Segarra ha sigut preminent, tant a l’hora de concebre el treball com descriure els esborranys complets. També son de destacar les publicacions de l’annex que tenen a en Joel Segarra com a primer autor. Les publicacions de l’annex són de nou en revistes del primer quartil i sobre aspectes de teledetecció, maneig del territori i ecologia terrestre relacionats amb la tesis, però que no tenen com a fill conductor el conreu de blat o l’agricultura mediterrània. És per això que estan incloses com annex.

En referència als cinc articles que conformen el cos central de la tesis mencionar que tots son contribucions rellevants dins alguns dels àmbits mencionats més amunt.

**Capítol 1.** Article. Wheat agroecological production and its spatial implications in the Euro-Mediterranean region, enviat a la revista Agronomy for Sustainable Development (Factor d’impacte=5.83 Q1 Agronomy and Crop Science ).

És un treball de frontera entre diferents disciplines. S’hi desenvolupa un metaanàlisi del cultiu agroecològic de blat en la conca mediterrània europea per entendre espacialment les virtuts i

limitacions d'aquest sistema de producció. A tal efecte s'hi combinen aspectes d'agroecologia del blat amb ecologia del paisatge d'ambients mediterranis tot emmarcat en l'adaptació a nous escenaris climàtics. Així doncs a part del metaanàlisi, s'han fet servir bases de dades georeferenciades i de teledetecció per entendre la distribució del conreu de blat i les virtuts i mancances del marc agroecològic en la conca mediterrània europea per tal de discutir possibles millores i estendre el cultiu agroecològic del blat en aquesta regió. En aquest article el doctorand Joel Segarra ha fet tot el metaanàlisi, ha fet els mapes corresponents, ha escrit l'article i és l'autor de correspondència.

**Capítol 2.** Article. Suitability mapping and management monitoring in Castilian organic and conventional wheat fields with Sentinel-2 and spatial data, publicat a la revista Ecological Informatics (Factor d'impacte=5.1 Q1 Ecology ).

L'article que conforma aquest capítol, s'ha publicat a finals del 2023 en la línia de l'estudi anterior, en aquest cas s'han emprant dades geolocalitzades del paisatge, dades agronòmiques i fisiològiques, estadístiques de fonts oficials i informació de teledetecció per comparar el cultiu ecològic i convencional de blat a Castilla i Lleó i definir zones d'adequació pel cultiu del blat. És de destacar l'ús d'un sistema de priorització que ha demostrat ser efectiu per desxifrar les classes d'idoneïtat del blat. Per altra banda encara que la productivitat de l'agricultura orgànica de blat és marcadament inferior a la de l'agricultura convencional, les diferències de rendiment entre els dos tipus d'agricultura són molt menors en terrenys marginals i es podria prioritzar aquest tipus de cultiu en el 25% de terres de cultiu categoritzades com a marginals en la zona estudiada. En aquest article el doctorand Joel Segarra ha fet els models, ha fet els mapes corresponents, l'anàlisi de les dades, ha escrit l'article i és l'autor de correspondència.

**Capítol 3.** Article. Estimating wheat grain yield using Sentinel-2 imagery and exploring topographic features and rainfall effects on wheat performance in Navarre, publicat a la revista Remote Sensing (Factor d'impacte=5.35 Q1 Earth and Planetary Science )

En aquest estudi publicat l'any 2020, que fins a data d'avui ha acumulat 19 cites (segons Google Scholar), s'hi van desenvolupar models de predicció de rendiment a escala regional emprant imatges satel·litals de Sentinel 2 adquirides entre el guaijat i la floració, juntament amb dades topogràfiques i de precipitació. Com a regió d'estudi es va emprar la comunitat foral de Navarra la qual està caracteritzada per un ampli ventall de condicions agroecològiques i de rendiments. Per la classificació del tipus de cultiu es va emprar una classificació "random forest" i les estimacions es van comparar amb les estadístiques de rendiment i superfície cultivada posterior

a la collita publicades pel govern foral. Després el model es va validar amb dades agronòmiques específiques d'una xarxa de camps de cultiu. A continuació el model validat es va utilitzar per explorar les possibles relacions de l'estimació del rendiment de gra per parcel·la mapejada amb les característiques topogràfiques i la pluja mitjançant regressions ponderades geogràficament. Els resultats de la validació van ser robusts i posen de manifest la capacitat d'en Joel Segarra per desenvolupar i aplicar metodologies que permeten mapejar i estimar el rendiment del blat i comprendre el seu aspecte geogràfic. En aquest article el doctorand Joel Segarra ha fet els models, ha fet els mapes corresponents, l'anàlisi de les dades i ha escrit l'article.

**Capítol 4.** Article. Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content, publicat a la revista Information Processing in Agriculture (Factor d'impacte=7.42 Q1 Agronomy and Crop Science )

Aquest article va ser publicat a l'any 2023 i ha acumulat a data d'avui 5 cites (Google Scholar). El capítol representa una continuació del capítol 3, però amb l'objectiu d'expandir el model de predicció a la qualitat del blat; concretament al contingut de nitrogen del gra a la vegada que explorava dos nivells de mesura. Per una banda, la predicció del contingut de nitrogen en parcel·les experimentals i per l'altra la seva extrapolació amb dades del Sentinel-2 en parcel·les d'agricultors de Navarra. Els índexs multiespectrals van funcionar millor que el RGB i en termes de fenologia les mesures en la floració van produir models més robustos de predicció del contingut de nitrogen del gra que amb avaluacions fetes en estadis fenològics anteriors. Aquest estudi mostra la capacitat d'en Joel Segarra d'estudiar diferents característiques del blat i treballar amb mesures multiescala des del fenotipatge a la teledetecció. En aquest article el doctorand Joel Segarra ha fet els models, ha fet els mapes corresponents, l'anàlisi de les dades i ha escrit l'article.

**Capítol 5.** Article. Farming and Earth Observation: Sentinel-2 data to estimate within-field wheat grain yield, publicat a la revista International Journal of Applied Earth Observation and Geoinformation (Factor d'impacte=7.67 Q1 Earth-Surfaces Processes)

El darrer capítol inclou un treball publicat al 2022 i que ha acumulat a data d'avui 35 cites (Google Scholar 35). L'estimació de la variabilitat del rendiment del blat dins del camp és fonamental per a la gestió agrícola, especialment en el context de canvi global actual. Aquest treball contribueix en avançar en l'ús de satèl·lits d'accés lliure, com el Sentinel 2, en l'agricultura de precisió. Presenta una integració de dades georeferenciades provinents d'una recol·lectora a nivell de

camp i bandes multiespectrals de Sentinel-2. El treball demostra la capacitat d'en Joel Segarra d'emprar algoritmes d'aprenentatge automàtic per estimar amb alta precisió el rendiment del blat. En aquest article el doctorand Joel Segarra ha fet els models, ha fet els mapes corresponents, l'anàlisi de les dades i ha escrit l'article.

Resumint, els 5 capítols de la tesis d'en Joel Segarra mostren una varietat d'enfocaments al voltant del cultiu de blat a la conca mediterrània des de l'ecofisiologia vegetal, la teledetecció i la geografia. Que un estudiant de doctorat sigui capaç de treballar al voltant d'un problema de manera exitosa amb uns nivells d'organització tant diversos demostra la maduresa, curiositat i capacitat intel·lectual del doctorant.

I per a que així consti als efectes oportuns

Dr. Jose Luis Araus

JOSE LUIS  
ARAUS  
ORTEGA - DNI  
36924747H

Signat digitalment  
per JOSE LUIS  
ARAUS ORTEGA -  
DNI 36924747H  
Data: 2024.05.12  
19:22:40 +02'00'

Dr. Shawn C. Kefauver

Firmado por KEFAUVER ---  
SHAWN CARLISLE - \*\*\*\*8650\*  
el día 12/05/2024 con un  
certificado emitido por AC  
FNMT Usuarios

## **Capítol 1.** Evidències per balançar el cultiu agroecològic del blat amb la recuperació de terres de cultiu abandonades a l'euromediterrània.

### **Chapter 1.** Trade-offs between wheat agroecological cultivation and farming recovery on abandoned cropland in the Euro-Mediterranean region

Joel Segarra<sup>a,b,\*</sup>, Shawn Carlisle Kefauver<sup>a,b</sup>, Jose Luis Araus<sup>a,b</sup>

<sup>a</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain

<sup>b</sup> AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198, Lleida, Catalonia, Spain

\*Corresponding author: Joel Segarra (e-mail: [joel.segarra@ub.edu](mailto:joel.segarra@ub.edu))

Enviat a Agronomy For Sustainable Development



# Trade-offs between wheat agroecological cultivation and farming recovery on abandoned cropland in the Euro-Mediterranean region

Joel Segarra<sup>a,b,\*</sup>, Shawn Carlisle Kefauver<sup>a,b</sup>, Jose Luis Araus<sup>a,b</sup>

<sup>a</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain

<sup>b</sup> AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198, Lleida, Catalonia, Spain

\*Corresponding author: Joel Segarra (e-mail: joel.segarra@ub.edu)

## Abstract

Agroecology is recognized as low-input sustainable management and implies the application of ecological concepts to agriculture to cope with climate stressors and environmental degradation. Despite its potential, some studies claim that agroecology produces lower yields than high-input industrial agriculture, which poses several difficulties in terms of fulfilling food production requirements within the cropland available. We present this study with the purpose of identifying factors that limit agroecological wheat yields and to explore the potential of wheat agroecological management in the Euro-Mediterranean region. For this, we conducted a systematic review of the scientific literature and examined spatial data on wheat farming and changes in cropland use. After following a systematic review protocol, 69 studies were analyzed. The regions with the greatest number of wheat agroecological studies are the northeast of Spain and the southeast of France. Italy has also received significant attention in this regard. Central Spain, Portugal, and Greece have a limited number of studies on wheat agroecological production. We found that wheat agroecological approaches can achieve better competitiveness against weeds, improve soil structure, and better mitigate climate change (carbon sequestration) than conventional approaches. Moreover, an agroecological approach to wheat production would also help increase biodiversity, contributing to the control of pests and providing safe harbor to wild biodiversity. Furthermore, we found that because croplands have been progressively abandoned over the past few decades, expansion of wheat agroecological production into marginal agricultural land beyond the current Euro-Mediterranean croplands would not result in a loss of natural habitats. In that respect, yield gaps in marginal croplands were found to be reduced and therefore might be suitable for agroecological production. We believe that this article contributes to providing agroecological and spatial evidence in the current milieu of European Union agrarian policy debate so that better sustainable management legislation can be developed.

Keywords: *Triticum*, agroecology, land use changes, global change, sustainability

## 1.Introduction

In the Mediterranean region, the reduction in cultivated croplands has been observed since the mid-twentieth century, the abandonment of marginal croplands has been mostly attributed to enhanced productivity on the most suitable croplands (Levers et al. 2018), disappearance of livestock farming integrated in agroecosystems, and loss of roads to access marginal croplands (Segarra et al. 2023b). In this regard, less easily accessible farming areas such as mountainous croplands have been abandoned (Jiménez-Olivencia et al. 2021) with negative effects in landscape functionalities preservation, such as fire prevention. Intensive woody monocultures, for instance, have increased and threaten traditional Mediterranean agricultural landscapes (Zimmerer et al. 2024), as well as its importance in providing ecosystem services and sustaining biodiversity.

Food production has seen significant transformations since the mid-twentieth century, transitioning into an industrialized system characterized by enhanced agricultural productivity. These advancements have played a crucial role in alleviating hunger (Evenson and Gollin 2003). Nevertheless, the prevailing food system and practices of industrial farming, which are characterized by a globalized market and widespread utilization of chemical inputs in agricultural production, exert significant and enduring effects on environmental integrity (Rounsevell et al. 2003). These effects manifest in negative alterations to nutrient dynamics (Cameron et al. 2013), depletion of water resources (Rosegrant et al. 2009), and loss of biological diversity (Lark et al. 2020). In comparison to other regions around the globe, the Mediterranean region exhibits diminishing biological diversity after the loss of well-established agroecosystems due to intensified agricultural practices and the abandonment of cultivated areas (Donald et al. 2001; Queiroz et al. 2014). Moreover, the agricultural sector is accountable for more than 20% of anthropogenic greenhouse gas emissions worldwide, taking into account land-use changes (Tubiello et al. 2015).

The imperative of simultaneously achieving food production targets and safeguarding ecosystems within their respective carrying capacities constitutes a crucial focal point in the Mediterranean region. This region is regarded as a climate change hotspot due to the projected substantial decline in rainfall and increase in temperature over the coming decades, which distinguishes it from other global regions (Cos et al. 2022). These environmental changes are anticipated to impose severe strain on the already water-stressed Mediterranean crops and ecosystems (Tuel and Eltahir 2020). For example, wheat serves as the primary staple crop within the Mediterranean region, and it has been projected that there will be a decline of 15–30% in wheat grain yield as a consequence of a 2°C increase in temperature anticipated for the year 2040 (Bento et al. 2021). Therefore, to address the shrinking wheat yields in the Mediterranean region, it is imperative to prioritize sustainable wheat production in the context of biodiversity conservation and the

efficient utilization of cropland within the prevailing natural habitats, which might entail the sustainable recovery of abandoned croplands.

Hence, assuming no alterations in agronomic and breeding technologies, it may be inferred that agricultural yields in the Mediterranean region will decline as a result of environmental conditions, necessitating the expansion of agricultural land to meet food demands. A lively debate has taken place among the scientific community and policymakers in recent years on how to address this challenge (Giraldo and Rosset 2018; Canfield et al. 2021). From this perspective, some scholars contend that agroecology has potential as an alternative way to attain sustainable production while simultaneously safeguarding the environment (Altieri and Nicholls 2020; Espluga-Trenc et al. 2021). Agroecology involves a decrease in inputs and the utilization of naturally existing environmental and biodiversity services, such as the management of agricultural pests using natural enemies instead of chemical pesticides (Tilman et al. 2002). However, in addition to the inherent potential of agroecology to address the environmental degradation caused by industrial agriculture, other scholars have highlighted significant limitations associated with this method of food production when compared to industrial agriculture (Trewavas 2001). In this regard, it is crucial to consider the criticisms surrounding the lower yields of agroecological practices, which may pose challenges in sustaining current levels of food production within existing agricultural or convertible cropland areas (Smil 2004; Connor and Mínguez 2012).

The Mediterranean region has garnered considerable scholarly interest in relation to food production, consumption, sustainability, and biodiversity (Ponti et al. 2016; Blas et al. 2019; García-Martín et al. 2021). The integration of traditional farming practices and livestock (Aguilera et al. 2020), as well as the socioecological implications of agroecology in the Mediterranean region (Migliorini et al. 2018) have been reviewed. Agronomic aspects of agroecology in the region have centered less attention, besides some relevant studies on wheat agrobiodiversity importance to adapt to productive sustainable environments (Costanzo and Bàrberi 2014; Lopes et al. 2015). Yet, the trade-off between cropland abandonment and sustainable farming practices in the Mediterranean region has not centered much attention. We hypothesize that the agroecological approach could represent a valid strategy to recover abandoned cropland, thanks to the productive and qualitative improvement brought by wheat agroecological management to lost agricultural systems.

In this regard, this study is presented with the objective of investigating the potential of wheat agroecological approaches to achieve sustainable production and assess the abandoned cropland area within the Euro-Mediterranean region that could benefit from wheat agroecological production. The current improvements in remote sensing and historic land use change archives, as well as the accessibility

to geographically referenced data, allows the spatial integration of several factors to study the trade-offs between wheat agroecology and cropland abandonment recovery. We focused on the Euro-Mediterranean region because of the European Union's unified agriculture policies, which may differ from patterns observed in North African and Middle Eastern countries.

The study has been organized based on three research questions. First, what are the factors that limit the spread of agroecological practices in the areas under study? By doing a systematic review of the scholarly literature we found several agroecological production constraints mainly based on biological, management, and economic factors. Second, what is the contribution of wheat agroecological production to environmental sustainability and biodiversity? The systemic review also showed beneficial effects of wheat agroecological production to the environment. Wheat agroecological production is compared to industrial agriculture. Following these findings, we suggest ways to improve wheat agroecological sustainable production through research and management. Third, how can the agroecological approach help to convert lost and/or abandoned areas to agriculture? The present study examines the existing extent of cropland indicating which ecosystem services could favor the spread of agroecological practices in the reference area and which could be supported by European agricultural policy.

## **2. Materials and methods**

### **2.1. Wheat study area**

Wheat (*Triticum* spp.) is a cereal crop grown worldwide. It originated in the Fertile Crescent over 10,000 years ago. The Euro-Mediterranean region became its next cultivation area (Royo et al. 2017). Wheat species in the genus *Triticum* include wild and domesticated species. Hexaploid bread wheat (*Triticum aestivum* L.) and tetraploid durum wheat (*Triticum durum* L.) are the main species. Spelt (*Triticum spelta* L.) and diploid emmer (*Triticum monococcum* L.) are relics cultivated to a lesser extent.

This study focuses on the European Union member states situated in the Mediterranean region with significant wheat cultivation, namely Greece, Italy, France, Portugal, and Spain. These countries are collectively referred to as the Euro-Mediterranean region. It is worth noting that Malta, Slovenia, and Croatia were excluded from the study due to the limited Mediterranean wheat agricultural land available, while Cyprus was not considered due to its geographical proximity to the Middle East. The study primarily examines the Euro-Mediterranean ecoregions, as depicted in Figure 1 (Olson et al. 2001). It assesses the impact of farmland abandonment and wheat agroecological research conducted in this area.

Ecoregions are geographical demarcations that serve as representations of species distributions, ecological models pertaining to the implications of climate change, or conservation programs (Dinerstein et al. 2017). The Euro-Mediterranean region encompasses a total of fourteen distinct ecoregions, each of which is designated on the basis of the specific distribution patterns seen among a diverse array of plant species. Out of the total of 14 ecoregions considered, three were excluded from further analysis due to the low presence of wheat agriculture. These three ecoregions are specifically identified as Corsican montane broadleaf and mixed forests, as well as Illyrian deciduous forests, and Southeast Iberian shrubs and woods. Therefore, a total of 11 ecoregions within the Euro-Mediterranean region were examined.



**Fig. 1.** Euro-Mediterranean ecoregions based on (Olson et al. 2001). The ecoregions comprise a specific area with a distribution of a broad range of flora (from which they are named), and contain a distinct assemblage of natural communities and species within a boundary that approximates their original natural status before major land-use changes.

## 2.2. Data and processing

### 2.2.1. Cropland use change

We used the Historic Land Dynamics Assessment+ (HILDA+) created by (Winkler et al. 2021) to examine changes in land use patterns over the course of the last six decades (1960–2019), which includes

a substantial impact from agroecosystems. In order to comprehend cropland loss, primarily caused by abandonment, in 11 Euro-Mediterranean ecoregions, data relating cropland loss, gain, stability and multiple changes (gain and loss) were retrieved.

#### 2.2.2. Wheat cultivation in the Euro-Mediterranean region

Information can be lacking when national agronomic agencies provide spatial data on crop cover at coarser spatial levels. Furthermore, several states may have distinct crop-related resolutions, some of which are not made public. Alternatively, satellite-based products provide a better high-resolution homogenized picture of the spatial distribution of wheat production in the Euro-Mediterranean region. In this regard, we used a relatively recent parcel-level crop-type categorization based on the European satellite Sentinel-1 Synthetic Aperture Radar observations by (d'Andrimont et al. 2021) to pinpoint the spatial location of wheat (bread and durum) in the Euro-Mediterranean region. This tool provides a parcel-level regional understanding of European crops and has been validated in numerous European locations.

#### 2.2.4. Processing

Geographic information system data was processed with ArcGIS Pro 2.3.0 software. The maps and spatial representation of this study were also elaborated with this software. We calculated the percentage of cropland loss in relation to the stable cropland by dividing the cropland loss area by the stable cropland using the HILDA+ dataset. Moreover, at the Euro-Mediterranean ecoregion scale, we calculated the percentage of wheat (both bread and durum) area in comparison to the total agricultural area from the same temporal landscape by dividing the wheat area calculated by d'Andrimont et al. (2021) and the total agricultural area defined by the same authors.

#### 2.3. Systemic review of articles

In order to investigate the primary factors that restrict wheat agroecological production in the Euro-Mediterranean region, a comprehensive review of relevant literature was conducted in accordance with the Preferred Reporting Items for systemic Reviews and Meta-Analyses (PRISMA) statement as described by (Page et al. 2021). The process of systematically organizing scientific literature on a particular subject enables researchers to assess all pertinent information more easily. This involves identifying relevant articles, screening their abstracts to assess their suitability, determining eligibility based on full-text examination, and ultimately selecting the most appropriate articles.

A comprehensive search for scientific literature related to the subject of the review was conducted using the Web of Science and Scopus databases. Further, a search was carried out on Google Scholar to



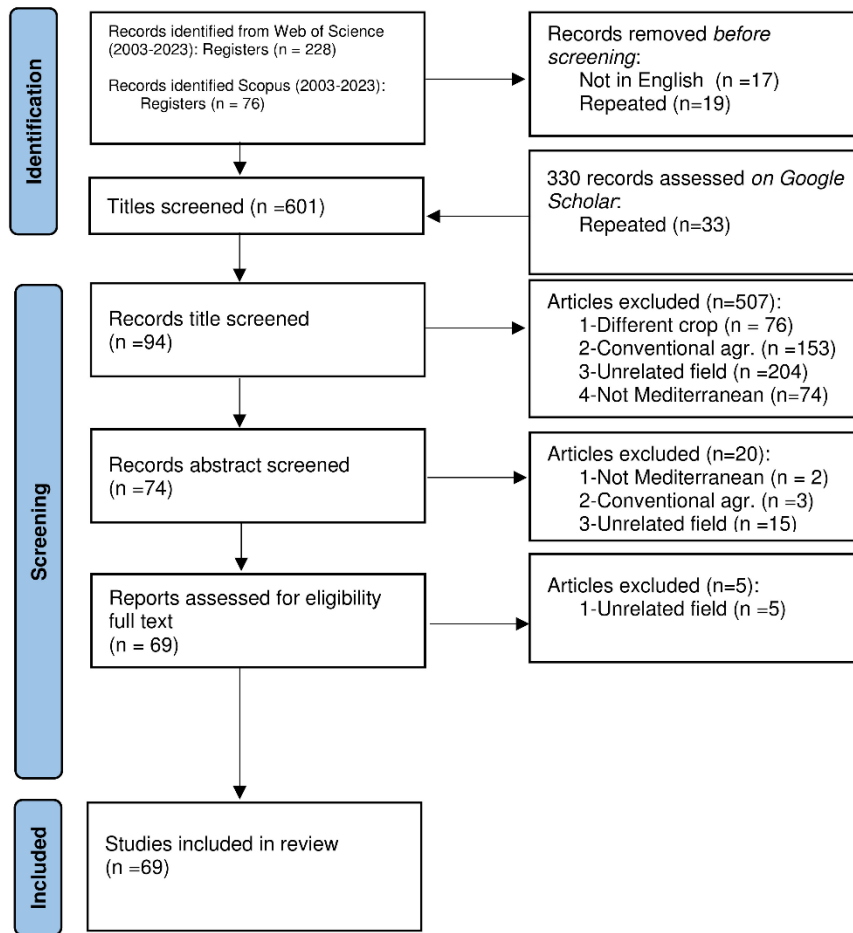
determine whether other articles could be found that might be missing from the previous databases. These platforms were chosen due to their widespread adoption within the scientific community and their reputation for providing reliable citations and analytical data (Li et al. 2018). The study encompassed two decades of literature from 2003 to 2023, and was structured and organized by our team. The search query employed syntax that incorporated specific keywords as follows: (Spain OR Spanish OR France OR French OR Italy OR Italian OR Portugal OR Portuguese OR Greek OR Greece) AND (Wheat OR Triticum) AND (ecological OR biological OR organic) AND (agrobiodiversity OR biodiversity conservation OR agroecology OR agroecosystem OR sustainability). The incorporation of biological or organic or ecological as a condition relates to the incorporation of research articles that either adhere to a certified biological approach to wheat management or incorporate certain management methods within the agroecological framework, such as low-input organic fertilization. The present study posits that the implementation of an agroecological approach to wheat management involves the utilization of organic inputs with minimal resource requirements, as well as the integration of environmental and biodiversity services to enhance wheat growth. Hence, it is not just limited to practices of certified organic agriculture. The papers that were searched were initially vetted by reading the abstracts and evaluating their suitability for the study's topic. After applying the filtering process, a total of 69 items were obtained. All of the articles considered in the study presented a case study that was grounded in evidence-based research (see Appendix S1). The publications were mapped and an indicator of the number of publications for each eco-region was calculated by dividing the number of publications by a specific unit of area (10,000 ha) of wheat production, representing a significant, yet proportional, area of the studied ecoregions. This was used to normalize the number of publications and better present the knowledge gaps spatially.

### **3. Results**

#### **3.1. Characteristics of the studies**

A total of 601 article titles retrieved from Web of Science, Scopus, and Google Scholar were obtained (Figure 2). Among the total number, 507 were eliminated from consideration during the title screening process for various reasons, such as not being specifically focused on wheat as a crop (n=76), being limited to conventional agriculture without any comparison or results related to agroecological management strategies (n=152), being otherwise unrelated to the field of study, particularly in the areas of chemistry and food science (n=204), or not being conducted in the Mediterranean region, such as studies conducted in northern Spain or France (n=74). Consequently, a total of 94 publications were evaluated based on their abstracts, out of which 20 were deemed unconnected with the aim of the study and eliminated

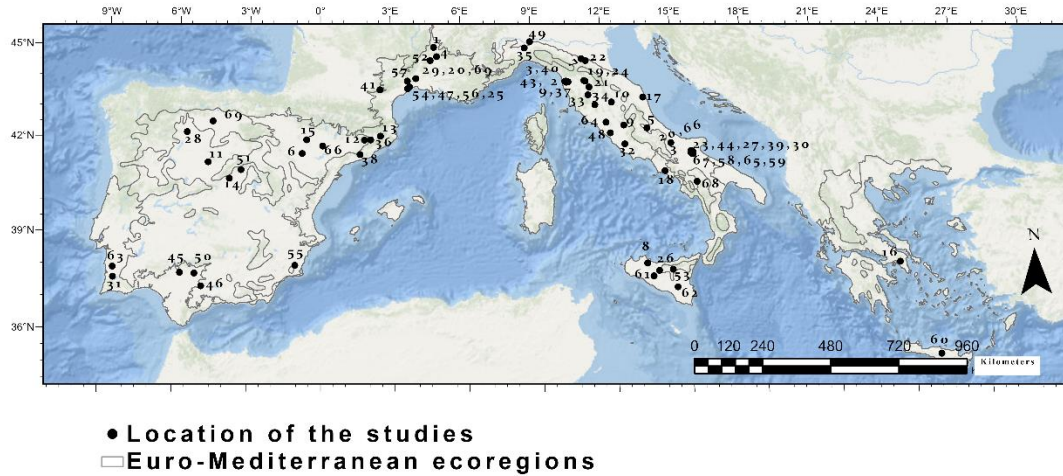
(Figure 2). Ultimately, following a comprehensive examination of the entire text, a total of 69 research studies were incorporated into the evaluation.



**Fig. 2.** Results obtained after following the diagram adapted from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram summarizing studies included and excluded.

The present study undertook a systematic assessment of scholarly literature, including a total of 69 studies derived from five different countries situated within the Euro-Mediterranean region. The countries with research articles on wheat agroecology were Italy (38), Spain (16), France (11), Portugal (2), and Greece (2) (see Figure 3 and 4). A general scarcity of research on wheat agroecology has been observed in Greece and Portugal. The research articles were disseminated through publications in scholarly journals that focus on specific areas of investigation, including agronomy, organic agriculture, landscape research,

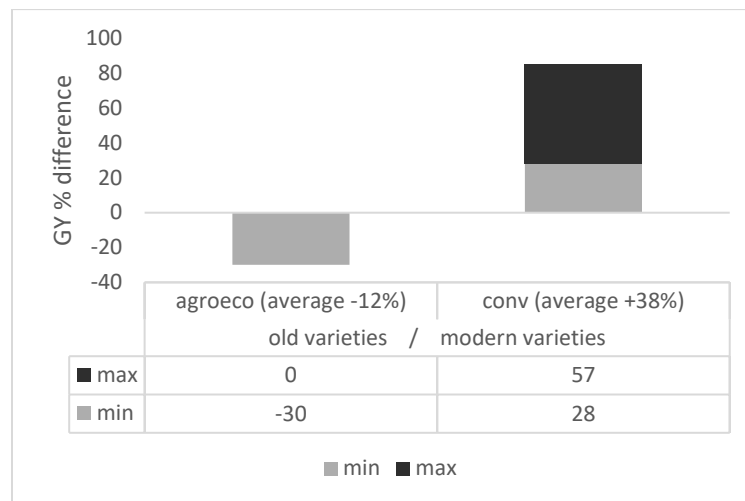
and ecology. In terms of the methods employed, the study encompassed a total of 37 cases conducted at research stations, alongside an additional 32 cases conducted in open agricultural areas. It is worth noting that a significant majority were published from the year 2012 onward, accounting for 77% of the total articles.



**Fig. 3.** Case study locations within the Euro-Mediterranean ecoregions. The numbers indicate the studies referenced in Table S1. Details of the number of studies by ecoregion are provided in Figure 4.

In terms of geographical representation, wheat case studies have been undertaken in 10 out of the 11 ecoregions under investigation (Figure 3 and Table 1). However, no relevant cases within our scope were found within the Pindus Mountains mixed forest ecoregion (Greece). In addition, it is noteworthy that Italian sclerophyllous and semi-deciduous forests ecoregions have been the subject of a considerable number of studies, totaling 31. The Northeast Spain and Southern France Mediterranean forest ecoregions have also attracted a relatively high amount of attention in the scientific community with 14 studies. The Tyrrhenian-Adriatic sclerophyllous and mixed forests ecoregion included four cases, and the Iberian sclerophyllous and semi-deciduous forests included eight. Five studies were conducted in the Southwest Iberian Mediterranean sclerophyllous and mixed forest ecoregions, while three were conducted in the South Apennine mixed montane forest ecoregion. The Mediterranean woods of Crete ecoregion was the subject of one study, and the Aegean ecoregion was also only represented by one study. Hence, the distribution of case studies is characterized by an uneven pattern, with the bulk concentrated in the Italic Peninsula, followed by south-eastern France and the northeast of the Iberian Peninsula. In relation to the wheat species examined in the studies, the majority (91%) of instances were based on bread wheat, with 32 cases, or durum wheat, with 31 cases. Additionally, six cases were dedicated to other varieties such as emmer or spelt, and one case did not provide a description of the wheat type utilized. Regarding the grain

yields obtained in the studies, the average grain yield increases of modern varieties in contrast with old varieties under conventional management versus agroecological settings were 38% increase and 12% decrease, respectively (Figure 4).

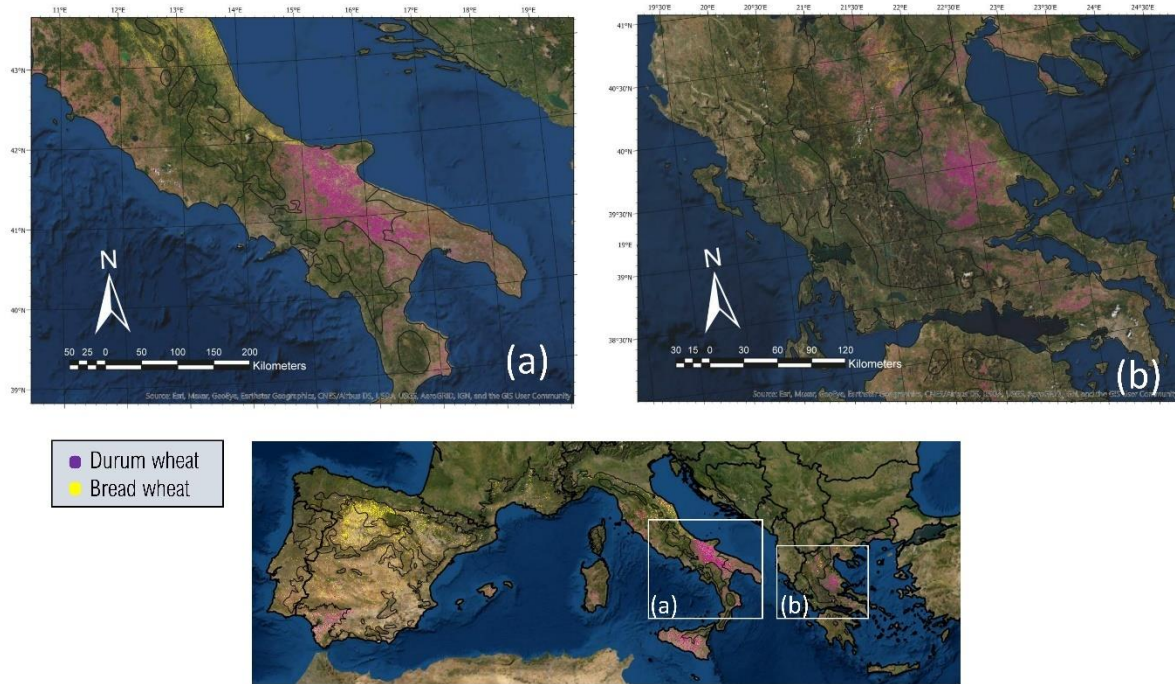


**Fig. 4.** Percentage of grain yield (GY) difference under agroecological management (agroeco) and conventional management (conv) with old varieties (landraces as well) and modern varieties (post-green revolution varieties adapted to high input management) based on the following studies results (De Vita et al. 2006; Pardo et al. 2011; Dinelli et al. 2013; Stagnari et al. 2013; Migliorini et al. 2014, 2016; Iannucci et al. 2018; Carranza-Gallego et al. 2018b). The grain yield average of old varieties in agroecological conditions is 2.6 t/ha (with a range between 2.1 and 3.6 t/ha) and in conventional conditions is 2.3 t/ha (with a range between 2 and 2.7 t/ha). Meanwhile the grain yield average of modern varieties in agroecological conditions is 2.9 t/ha (with a range between 2.2 and 3.6 t/ha) and in conventional conditions is 3.5 t/ha (with a range between 3.1 and 3.7 t/ha).

### 3.2. Wheat cultivation and cropland use

Figure 5 presents the current distribution patterns of the two primary wheat varieties, namely bread wheat and durum wheat, within the Euro-Mediterranean region. Additionally, it provides a detailed examination of specific ecoregions, namely the Italian sclerophyllous and semi-deciduous forests, the Tyrrhenian-Adriatic sclerophyllous and mixed forests, and the Aegean and Western Turkey sclerophyllous and mixed forests. These regions are of particular interest as they are known for extensive cultivation of durum wheat. Figure 5 illustrates the prominent growth of bread wheat in the northernmost regions, specifically the Iberian Peninsula and the Italic Peninsula. This cultivation is further supported by the data presented in Table 1. The Iberian sclerophyllous and semi-deciduous forests ecoregion, along with the Northwest

Iberian montane forests ecoregion, collectively account for 73% of the overall cropland dedicated to bread wheat cultivation within the Euro-Mediterranean region. The Iberian sclerophyllous and semi-deciduous forests ecoregion encompasses an area of 1,060,800 hectares, whereas the Northwest Iberian montane forests ecoregion spans 428,900 hectares dedicated to the cultivation of bread wheat. The Italian sclerophyllous and semi-deciduous forests ecoregion encompasses a notable expanse of land dedicated to the cultivation of bread wheat, spanning around 202,500 hectares. This agricultural activity is particularly prominent in the northern regions of the ecoregion, but it should be noted that it has a far larger area dedicated to the cultivation of durum wheat, with a total of 615,300 hectares. This makes the Italian sclerophyllous and semi-deciduous forests ecoregion the leading producer of durum wheat in terms of cultivated land area, as indicated in Table 1. In southern Italy, the Tyrrhenian-Adriatic sclerophyllous and mixed forests ecoregion is also home to a substantial area of durum wheat cultivation, spanning approximately 419,700 hectares. Furthermore, the ecoregion known as Aegean and Western Turkey sclerophyllous and mixed forests encompasses a substantial area dedicated to the cultivation of durum wheat, covering 234,100 hectares (Table 1). This information is visually depicted in Figure 5. In contrast to the prevalence of common bread wheat in the northern Iberian Peninsula, the spatial pattern distribution revealed a noteworthy presence of durum wheat within the Southwest Iberian Mediterranean sclerophyllous and mixed forests ecoregion. This area includes a substantial growing region for durum wheat, spanning around 189,000 hectares. Therefore, bread wheat is the predominant crop in this region, covering over 2 million hectares, while durum wheat occupies around 1.8 million hectares. As depicted in Figure 5 and Table 1, it is evident that the spatial distribution across the Euro-Mediterranean region exhibits contrasting characteristics. Specifically, the northern Mediterranean regions exhibit a greater prevalence of bread wheat cultivation, while the southern parts demonstrate a greater extent of durum wheat cultivation.



**Fig. 5.** Wheat growing areas within the Euro-Mediterranean regions based on (d’Andrimont et al. 2021). Panel A shows the Italian sclerophyllous and semi-deciduous forests and the Tyrrhenian-Adriatic sclerophyllous and mixed forests ecoregions, while Panel B shows the Aegean and Western Turkey sclerophyllous and mixed forests ecoregion. The yellow pixels correspond to bread wheat, while the purple pixels correspond to durum wheat.

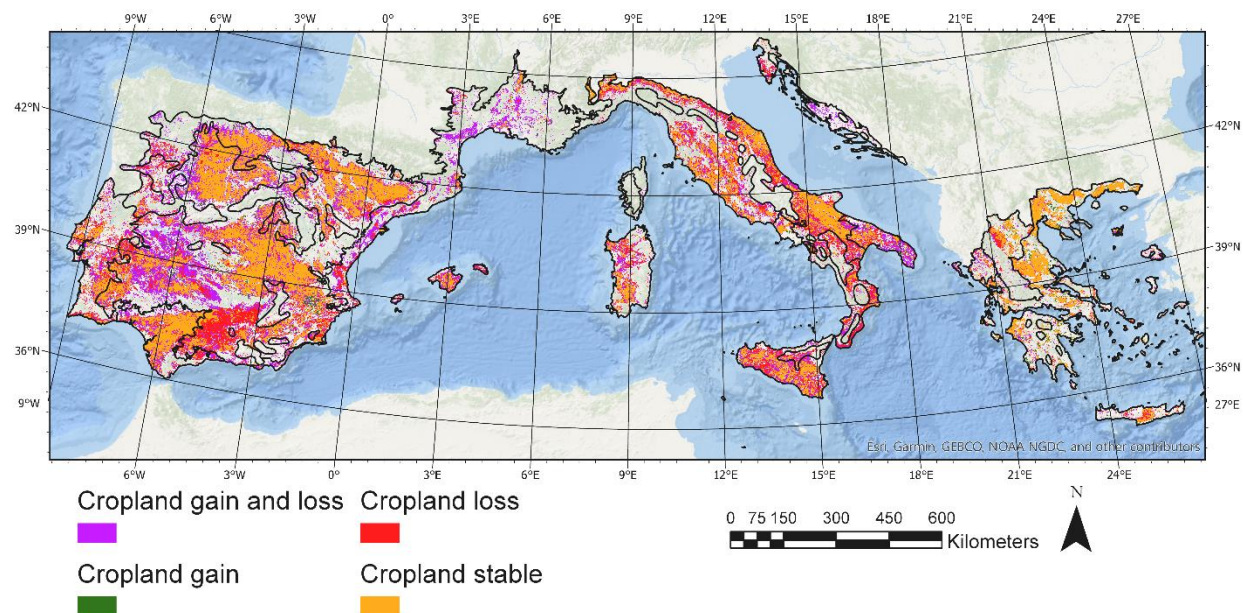
The number of studies examining wheat agroecological production was normalized with the area of wheat cultivation for each ecoregion (Table 1). The number of studies of 10,000 ha in size was greatest in the Northeast Spain and Southern France Mediterranean forests with 1.24 studies, and in the Crete Mediterranean forests with 3.85 studies. In the latter case, the limited area of wheat cultivation in the island produced a high value because there was only one study. It is remarkable that the area with the highest number of studies, namely the Italian sclerophyllous and semi-deciduous forests, when normalized with the land area under wheat cultivation had 0.4 studies per unit of land (10,000 ha). Meanwhile, the Iberian sclerophyllous and semi-deciduous forests, being the ecoregion with the largest production area of bread wheat, had only 0.06 studies on wheat agroecological production per unit of land (Table 1).

**Table 1.** The growing areas of durum wheat and bread wheat based on d’Andrimont et al. (2021) within the Euro-Mediterranean ecoregions (Olson et al. 2001), and the number of studies per 10,000 ha of wheat cultivated in each ecoregion is indicated.

Euro-Mediterranean ecoregion	Wheat growing area (ha)		Number of Studies · 10kha <sup>-1</sup> (total studies)
	Durum wheat	Bread wheat	
Aegean and Western Turkey sclerophyllous and mixed forests	234,100	46,700	0.04 (1)
Crete Mediterranean forests	1,900	700	3.85 (1)
Iberian conifer forests	1,300	100,000	0.1 (1)
Iberian sclerophyllous and semi-deciduous forests	184,200	1,060,800	0.06 (8)
Italian sclerophyllous and semi-deciduous forests	615,300	202,500	0.4 (31)
Northeast Spain and Southern France Mediterranean forests	38,600	82,000	1.20 (14)
Northwest Iberian montane forests	4,300	429,000	0.02 (1)
Pindus Mountains mixed forests	43,000	17,000	0 (0)
South Apennine mixed montane forests	52,200	2,600	0.55 (3)
Southwest Iberian Mediterranean sclerophyllous and mixed forests	189,000	92,200	0.18 (5)
Tyrrhenian-Adriatic sclerophyllous and mixed forests	419,700	16,700	0.1 (4)

The cropland use dynamics for the last 60 years are shown in Figure 6. Substantial agricultural land use changes have been observed, although the main spatial category was stable croplands in most of the studied ecoregions. Nevertheless, in the case of the Northeast Spain and Southern France Mediterranean forests, and the South Apennine mixed montane forests, the cropland area loss was larger than the stable cropland area. Indeed, the area that currently supports agricultural production is smaller than the agricultural land that has been abandoned in the last six decades. Furthermore, these two ecoregions include large areas of land with multiple spatial changes.





**Fig. 6.** Cropland dynamics 1960-2019 based on the HILDA+ dataset (Winkler et al. 2021) for the Euro-Mediterranean region. Cropland gain and loss (in purple) implies multiple change events within the studied period, cropland gain (in green) refers to spatial cropland expansion into non-agricultural areas, cropland loss (in red) refers to the loss of agricultural area, while cropland stable (in orange) refers to stable agricultural lands. A table with all the area data is presented in the appendix (Table S2).

**Table 2.** Cropland use dynamics for each of the Euro-Mediterranean ecoregions. The % of cropland used for wheat farming based on d’Andrimont et al. (2021) dataset and the % of cropland loss relative to stable cropland are indicated based on HILDA+ dataset.

Euro-Mediterranean ecoregion	% of wheat in cropland			% of cropland loss since 1960 relative to stable cropland
	Durum	Bread	Total	
Aegean and Western Turkey sclerophyllous and mixed forests	10.9	2.2	13.1	15
Crete Mediterranean forests	1.9	1	2.9	73
Iberian conifer forests	0.3	20.5	20.8	31
Iberian sclerophyllous and semi-deciduous forests	1.8	10.5	12.3	35
Italian sclerophyllous and semi-deciduous forests	19.3	6.3	25.6	60
Northeast Spain and Southern France Mediterranean forests	9.4	19.9	29.3	127
Northwest Iberian montane forests	0.4	40.1	40.5	30
Pindus Mountains mixed forests	15.1	6	21.1	55
South Apennine mixed montane forests	45.6	2.3	47.9	185
Southwest Iberian Mediterranean sclerophyllous and mixed forests	10.7	5.2	15.9	58

The stable gain of cropland area has been marginal in all the studied ecoregions (Table S2), especially in contrast with the areas where cropland loss and multiple changes in cropland have occurred. This indicates a trend of reductions in cropland area over recent decades. The two ecoregions with the largest cropland loss (Table 2) were Northeast Spain and Southern France Mediterranean forests (127%) and South Apennine mixed montane forest (185%). The following reductions in stable cropland area per ecoregion have been observed: 85% in the Tyrrhenian-Adriatic sclerophyllous and mixed forests ecoregion; 73% in the Crete Mediterranean forests ecoregion; 60% in the Italian sclerophyllous and semi-deciduous forests ecoregion; 58% in the Southwest Iberian Mediterranean sclerophyllous and mixed forests ecoregion; and 55% in the Pindus Mountains mixed forests ecoregion. Only four ecoregions (Aegean and Western Turkey sclerophyllous and mixed forests, Northwest Iberian montane forests, Iberian conifer forests, and Iberian sclerophyllous and semi-deciduous forests) had a cropland loss below 50% of the equivalent stable cropland.

#### 4. Discussion

Regarding the first research question, "what are the factors that limit the spread of agroecological practices in the areas under study?", we have observed several factors constraining wheat agroecological production after examining the scientific literature. These are few varieties adapted to such management approaches for adequate grain yield and quality performance, a lack of economic viability, low competitiveness against weeds, susceptibility to pests without effective treatments, and, in some regions, a deficit of studies providing evidence and management alternatives to reduce the yield gaps within conventional agriculture. Yet, some agroecological management practices are relatively already well-integrated in cereal systems, such as organic fertilization and reduced tillage.

Regarding the first factor, the general assumption is that the adaptability of preferred cultivars is facilitated by breeding for extensive, uniform environments that are controlled using standardized intensive inputs (such as chemical fertilizers, herbicides, and pesticides). However, it is important to note that this assumption is based on agricultural systems in which these varieties are cultivated at comparable levels of high-input conventional agricultural management. In contrast, in the context of marginal environments and agroecological managements characterized by reduced inputs, varieties specifically bred for narrower target environments appeared to exhibit greater stability under those specific conditions (Ceccarelli 1994). For instance, we observed an average 12% reduction in grain yield when using old varieties in contrast with modern varieties (Figure 4) under agroecological management (De Vita et al.

2006; Dinelli et al. 2013; Stagnari et al. 2013; Iannucci et al. 2018). Yet, under extremely marginal conditions, (Carranza-Gallego et al. 2018b) observed improved yields of old varieties under agroecological conditions, especially due to relatively wetter seasons and therefore higher weed infestation and improved viability of higher genotypes (i.e., older).

In general terms, however, modern varieties outperform older varieties, even in extremely marginal croplands such as those described by Wang et al. (2021). In that case from Greece, modern wheat varieties outperformed older varieties, and moreover, the agroecological approach outperformed the conventional, albeit in marginal yields of about 1,000 kg/ha. Overall, in conventional management, the average grain yield increase in modern varieties in contrast to old varieties is 38% (Figure 4), as averaged from the analyzed scientific literature (Stagnari et al. 2013; Migliorini et al. 2014, 2016; Carranza-Gallego et al. 2018b). This illustrates that under conventional management, modern varieties have a larger yield gap with old varieties than under agroecological management. Moreover, Pardo et al. (2011) described that under relatively marginal Mediterranean environmental conditions, the differences in production between agroecological and conventional are nonexistent with equivalent varieties. Dinelli et al. (2013) also observed no grain yield difference in Mediterranean semiarid regions for organic vs. conventional management with contrasting weed treatments or organic vs. inorganic fertilization. Water stress occurring during the onset of flowering is much more severe in chemically fertilized wheat fields due to the greater leaf area of the crop (Stagnari et al. 2013). Landraces, which are genetically heterogeneous populations, are traditionally cultivated in marginal land areas and can serve as a genetic resource to enhance wheat characteristics that are more suited to agroecological production (Ceccarelli 2011). These landraces with high chlorophyll concentrations in leaves have the potential to boost photosynthetic capacities and, when combined with high stomatal conductance, enhance leaf cooling, resulting in improved heat tolerance (Reynolds et al. 1994).

Agroecological approaches in poor nitrogen conditions make it challenging to achieve adequate grain protein content. (De Vita et al. 2007) described a reduction in grain protein content of 21.25% under agroecological management in contrast to conventional management. The availability of fertilization can be addressed by improving crop rotation and adding leguminous rotations. Moreover, crop rotations can increase wheat yields by 6.5% compared to continuous cropping (Montemurro and Maiorana 2015). Alongside management practices, a study of several organic bread wheat fields showed that the main reason for the differences in grain protein content was the cultivar's baking quality grade (Casagrande et al. 2009). This characteristic is highly associated with genotypic attributes. Even so, it is generally seen that higher grain protein production occurs in farming systems that use modern crop varieties and high-input fertilization management (Sanchez-Garcia et al. 2015). Therefore, in addition to utilizing a cultivar

with enhanced baking quality, the grain protein content in agroecological bread wheat may be augmented by enhancement of fertilization management. This can be achieved by opting for a legume fodder crop as a preceding crop or by refraining from late sowing dates (Casagrande et al. 2009). Therefore, improving management and having broad genetic variability are central to breeding in agroecological settings. Breeding a new generation of varieties (Casals et al. 2019) or a mixture of advanced varieties adapted to target environments, together with improvements in agricultural management, might reduce the yield gap between industrial and agroecological management. According to (Padel et al. 2021), there is a consensus among organic breeders in the European Union regarding the necessity of establishing dedicated breeding programs for organic durum and bread wheat.

Regarding the economic viability of wheat agroecological production in the Euro-Mediterranean, several authors have shown that it needs to overcome significant economic constraints. In contrast with conventional agriculture, (Bux et al. 2022) observed a 55% higher profit in conventional wheat production than in agroecological approaches in Italy, while (Pardo et al. 2009) observed that in semi-arid Mediterranean Spain, neither conventional nor organic wheat production are sustainable economically and need to be subsidized. However, a study conducted by (Varia et al. 2021) has demonstrated that organic systems cultivating landraces of durum wheat in Sicily exhibit a notable economic benefit when compared to the use of standard varieties because they could be sold at higher prices and were preferred by consumers. This was also observed by (Tudisca et al. 2014). Nonetheless, (Lago-Oliveira et al. 2023) observed significant economic improvements in wheat cultivation when it included rotations with chickpeas (a leguminous crop) due to a diversification of products and a reduction in inputs. Therefore, if varieties valued by consumers are cultivated using agroecological approaches, they might become economically viable; however, this is a very specific niche market that may limit the widespread adoption of wheat agroecological approaches. It is relevant to point out that all the studies dealing with economic aspects have focused on durum wheat and not on bread wheat, probably due to the singularity of this cultivar, which is virtually not present in other parts of Europe, in contrast with the widely extended bread wheat.

Regarding competition from weeds, this is a challenge in agroecological production due to the limitations in herbicide inputs. A way of finding alternatives has been addressed by examining both ancient and landrace genetic material when investigating characteristics that are not commonly prioritized in breeding for conventional agricultural settings, such as the capacity to compete against weeds (Lazzaro et al. 2019; Carranza-Gallego et al. 2019). In this regard, Lazzaro et al. (2019) discovered that older wheat cultivars exhibited the most effective weed suppression, with an average reduction of 67% in weed biomass at harvest when compared to newer cultivars. This outcome is likely attributed to the greater height, above-

ground biomass, and leaf area index observed in the older varieties. Furthermore, it was noted that mixes consisting of a greater number of components, specifically six and twelve genotypes, tended to enhance the overall crop performance in comparison to less diversified wheat stand types. The utilization of bread wheat variety mixes has been suggested as a potential approach to enhancing genetic variety and performance, particularly in terms of yield stability, as previously indicated. According to (Wolfe 2000), this tactic also has the potential to improve disease suppression. (Costanzo and Bàrberi 2016) observed that weeds in wheat production can also be reduced by 30–41% by establishing a living mulch system and avoiding the use of herbicides, although grain yield reductions up to 50% were observed. Therefore, some genetic improvements might contribute to improving competitiveness against weeds, but at the expense of a reduced final grain yield.

Regarding the effects of pests, aphid damage poses a great threat to wheat crops, resulting in production losses due to the depletion of plant nutrients and the transmission of diseases, which generate about 700,000 tonnes of wheat losses in Europe (Dedryver et al. 2010). Generalist predators are known to prey on these crop pests. (Chabert and Sarthou 2017) demonstrated that the population density of these predatory organisms is influenced by the level of complexity within the agricultural system rather than being only determined by management approaches, such as the choice between organic and industrial methods. Practices such as implementing a cover crop prior to seeding wheat or adopting long-term no-tillage techniques are significant determinants in the regulation of aphid populations. Despite their study's spatial limitations, (Moschini et al. 2012) observed that the implementation of agroecological management, particularly over an extended duration, resulted in an augmentation of the relative density of aphid predators within wheat fields. Additionally, the existence of ecological infrastructure surrounding the fields had a positive influence on their uniform distribution. Therefore, it is suggested that agroecological management contributes to the biological regulation of pests. However, it is important to acknowledge that other aspects related to the methodologies employed and the ecological framework encompassing the fields are equally crucial in fostering local biodiversity. (Pérez-Fuertes et al. 2015) demonstrated that irrigation has a positive impact on the diversity of several arthropod groups and also promotes plant production under arid conditions. Nevertheless, the applicability of this management strategy may be limited due to water constraints in the Mediterranean setting. According to (González-Estébanez et al. 2011), irrigation has the potential to contribute to the design of agroecosystems and promote biodiversity while also ensuring production stability. Despite this, detrimental impacts on soil erosion and biodiversity might be observed as a result of increased salinization, which can be attributed to factors such as inadequate irrigation practices or the utilization of substandard water sources. The limitation of inputs in agroecological management makes the control of pests challenging, but the benefits obtained from the surrounding diversity and management practices, such as no-tillage and crop rotations,

contribute to controlling pests. Yet the pest-derived yield losses are larger than in conventional management practices (De Ponti et al. 2012).

The areas with the largest knowledge gaps in wheat agroecological management practices are mainland Greece (0 and 0.04 studies per 10,000 ha of wheat cultivation) and central Spain (0.06 and 0.02 studies per 10,000 ha of wheat cultivation) (Table 1). This is relevant because central Spain has the largest area of bread wheat production in the Euro-Mediterranean, moreover, it has significant wheat cultivation in marginal areas that could benefit from agroecological approaches, with 25% of croplands in such marginal conditions in some areas (Segarra et al. 2023a). In particular, central Spain is constituted by a large cereal steppe land that would benefit from greater understanding of sustainability and improvement to management practices. Greece has significant durum wheat production and a rich agrobiodiversity of wheat types (Pecetti et al. 1992), which would favor agroecological management. Although traditional farming is widespread in Greece and might resemble some aspects of agroecology, climate change adaptation, nonetheless, necessitates the development of improved management approaches. Southern Italy has received more scholarly attention than southern Spain and Portugal. While central-north Italy and northeast Spain/southeast France are areas with the greatest number of studies on wheat agroecology. In this regard, wheat agroecological production varies greatly across the Mediterranean regions and the wheat varieties used (old or modern cultivars, for instance), as has been presented in Figure 4. This is also observed within relatively similar regions, such as southern France, where David et al. (2005) described grain yields ranging from 1.4 to 4.9 t/ha, depending on many factors. Hence, assessing the multiple factors discussed here at an ecoregional level could contribute to spatially improving wheat sustainable production and highlight areas lacking research on the topic so that more resources can be devoted to researching the mentioned limiting factors.

Regarding the second research question, “What is the contribution of wheat agroecological production to environmental sustainability and biodiversity?”, the review of the literature indicated that wheat agroecological production improves soil structure and avoids erosion. Moreover, it increases the wild biodiversity of flora and fauna and has a smaller carbon footprint than conventional approaches, these factors have been nonetheless little integrated in agricultural practices.

It is noteworthy that industrial agricultural methods, such as tillage and the use of chemical fertilizers, make a substantial contribution to the erosion of cropland, thus exerting an impact on a crucial component of agricultural systems. Soil erosion and inadequate soil management are particularly prevalent in Mediterranean regions. In the context of Mediterranean semi-arid rain-fed wheat production in Spain, the implementation of low-input management practices, such as no-tillage and the use of green manures, has been found to have positive effects on soil properties. Specifically, these practices have been observed to

enhance the accumulation of organic carbon and nutrients (nitrogen and phosphorus) in the soil. Additionally, the adoption of these low-input strategies has been associated with a reduction in both runoff (30% lower compared to conventional industrial practices) and erosion (65% lower compared to conventional industrial practices). Furthermore, these practices have been found to facilitate the mobilization of organic carbon and nutrients in sediments. These findings, as reported by (Martínez-Mena et al. 2020), suggest that the implementation of agroecological management practices has the potential to contribute to a long-term reduction in soil erosion. Such an approach has also been observed as a suitable way to reverse the erosion of wheat farming land in Italy (Siciliano 2009).

In that regard, agroecological systems, encompassing various crops such as wheat, have 13% and 9% greater efficacy in soil carbon (C) and soil nitrogen (N) storage, respectively, than conventional wheat farming (Migliorini et al. 2014). Moreover, no tillage practices increase the soil organic matter content (3.3%) and show a higher soil microbial biomass (71%) (Sapkota et al. 2012). The retention of crop residues in the field, as followed in conservation practices, is better at improving soil organic carbon sequestration than other practices and represents an improvement of soil structure in agroecological approaches (Di Bene et al. 2016). The spatial dimension and the presence of natural elements are key factors in determining the sustainability of biotic and abiotic interactions within soils.

The practice of industrial agriculture has led to landscape homogenization, an increase in cropland abandonment, and a decline in crucial interactions with various organisms, including weeds at field edges, insects, reptiles, and birds, among others. Furthermore, it is worth noting that there has been a general trend of increasing field size since the middle of the twentieth century. This has resulted in the fragmentation of croplands into huge, uniform areas that are often maintained using herbicides, pesticides, and other agricultural inputs. Despite the potential benefits of this agricultural paradigm in terms of increased yields and improved management, it has also had a substantial negative impact on the biodiversity of agricultural fields. It was found that in the northeastern part of the Iberian Peninsula, weed diversity was 58% reduced, was 47% less rich, and covered 69% less land area overall in industrial wheat fields compared to fields that were managed using agroecological practices (Chamorro et al. 2016). This reduction has significant implications for various organisms, including birds, pollinators, and other invertebrates, as well as segetal and rare species subsets, which have experienced decreases of 75% and 87%, respectively. These results are consistent with the findings observed in Mediterranean dryland cereal fields in other regions of the Iberian Peninsula. Research conducted by (Jose-Maria and Sans 2011) revealed that field edges and organic fields exhibited higher species richness and denser seedbanks compared to industrially managed fields. According to (Biaggini et al. 2007), a study conducted in Italy revealed that organic farming, when combined with semi-natural regions, exhibited the best levels of



plant species richness. This combination was proven to effectively protect and promote arthropods within agricultural landscapes. Moreover, the inclusion of legumes in wheat agroecosystem management can improve the conservation of beneficial arthropods and eventually contribute to ecosystem services such as natural pest control (Caballero-López et al. 2016). Agroecological wheat fields in Mediterranean drylands in Spain showed a three-fold higher abundance of weeds and a 1.4-fold higher abundance of arthropods than conventional farming (Ponce et al. 2011).

The management of industrial agriculture entails the extensive utilization of agrochemicals and the gradual depletion of various natural and semi-natural habitats. Consequently, this has resulted in a significant decline in biodiversity within agricultural systems. Wintering agricultural birds are subject to the influence of both farm management practices and the composition of the surrounding landscape. According to (Geiger et al. 2010), landscapes that are less fragmented and consist of arable crops and grasslands provide favorable conditions for a wide variety of farmland bird species throughout the winter season. Additionally, agroecological farming practices, such as rotation with fallows and no-tillage (Tarjuelo et al. 2019), have been found to have positive impacts on farmland birds.

Agroecological approaches contribute to increasing cultivated diversity. Intercropping is an agricultural technique employed to enhance crop performance and mitigate the potential hazards associated with crop losses by promoting production diversification. Tree-crop systems are recommended as a means to alleviate water and heat stressors, specifically in semi-arid regions of the Mediterranean area, as exemplified by olive groves. However, it has been observed that the diminished light caused by the shade of olive trees leads to a decrease in grain output ranging from 40% to 80% for durum wheat (Panozzo et al. 2020b). Light is indeed a constraining factor, but wheat-tree systems offer shade through their canopy and protection from wind due to the windbreak effect. Trees play a vital role in buffering microclimates and regulating water flow, thus providing shelter and contributing to the overall functioning of the ecosystem. Intercropping with wheat olive orchards can improve the profitability of low-productivity olive orchards (Panozzo et al. 2020a). Moreover, intercropping olive trees with wheat between rows results in a decrease in erosion rates (Vanwalleghem et al. 2011).

Another relevant aspect of wheat agroecological production is the reduction in carbon footprint and the capacity to increase carbon sequestration due to reduced tillage and higher residues remaining on the soil with increased microbiological activity. Bux et al. (2022) showed that wheat agroecological production in southern Italy can reduce the quantity of emissions associated with conventional farming by about 19%, yet a larger area for wheat cultivation might be needed along with changes in consumption patterns. Meanwhile, in southern Spain, a carbon footprint reduction of 5.8-fold has been observed in wheat

agroecological management relative to conventional management (G. Carranza-Gallego et al., 2018), although the cost was lower grain yields.

Regarding the third research question, “how can the agroecological approach help to convert lost and/or abandoned areas to agriculture?”, the abandoned cropland in all the studied eco-regions occupies significant areas and could host a re-cultivation of wheat with agroecological approaches while leaving the more productive lands to more productive approaches (Table 2). Therefore, we believe that Euro-Mediterranean cropland can sustain wheat agroecological production on at least 25% of the croplands in each eco-region, without occupying the more productive croplands. In that regard, the European Union aims to dedicate a 25% share of the overall European farmland to organic agriculture by the year 2030, as outlined in the European Commission's document titled "Action Plan for the Development of Organic Production in the EU" (2021). Hence, identifying marginal croplands can assist in delineating regions characterized by reduced yield gaps, following the data and evidence presented here. The primary factor contributing to the reduction in cropland is its conversion to forest as a result of abandonment. The abandonment of agricultural activities marginally enhances biodiversity conservation, increases the risk of fire, and overuses water resources in Mediterranean basins. Therefore, it is relevant to plan a recultivation of some of this cropland to reverse these challenges. The extension of wheat agriculture, for instance, in such areas should adhere to agroecological principles within a diverse landscape.

The imperative to preserve, augment, or rehabilitate cereal agricultural landscapes in the Euro-Mediterranean region is evident under the current climate change scenario, economic instability, and European Union agrarian policy debate. The contribution of wheat agroecology to recover abandoned croplands is especially significant regarding the ecosystem services it provides. It contributes to the four categories of ecosystem services. Namely, it has relevant provisioning services, such as the provision of wheat derived food products with high nutritious qualities (Dinelli et al. 2013). Moreover, it contributes to regulating services, such as the improved carbon sequestration of wheat under agroecological practices (Carranza-Gallego et al. 2018a), as well as the prevention of fire risk (Aguilera et al. 2020) by improving the landscape integrity of the Mediterranean eco-regions after the recovery of abandoned croplands. It also contributes to supporting services such as improved nutrient cycles following crop rotations and organic fertilization, as well as adequate integration with livestock in agroecosystems (Billen et al. 2021). Finally, wheat agroecological production in the recovery of abandoned cropland entails important cultural ecosystem services especially the heritage of Mediterranean diet, which entails the health of the population and the environment (Soler-Montiel et al. 2024), and the importance of well-integrated landscapes for the social and economic benefit (Giacomelli et al. 2024). These relevant ecosystem services of wheat agroecological production in the recovery of abandoned Mediterranean croplands

illustrate management approaches that the European Union could support for the benefit of Euro-Mediterranean agroecosystems and its population.

## **5. Conclusions**

Wheat agroecological production in the Euro-Mediterranean region makes proven contributions to environmental sustainability by reducing inputs (fertilization and energetic demands) while improving soil structure, reducing soil erosion, and increasing the diversity and quantity of weeds, arthropods, microorganisms, and birds. All these factors play a crucial role in the sustainability of Mediterranean cereal landscapes, which could be extrapolated to other global regions, and improve the resilience of agricultural production in a biodiversity, climate, and food crisis scenario. However, we have also observed several factors constraining wheat agroecological production in the Euro-Mediterranean region. The scarceness of adapted varieties to such management approaches for adequate grain yield and grain protein performance (up to 57% and 21% reduction in agroecological production compared to conventional, respectively) could be addressed by improving varieties for specific sub-regions and agroecological management (such as those with improved weed competitiveness) and focusing on production in marginal Mediterranean croplands (reduced yield gaps of 12% on average). In all of the studied regions, recently abandoned cropland (since the 1960s) would provide a sufficient margin for re-cultivation of wheat agroecological production in marginal croplands while leaving more suitable croplands to more productive approaches. Besides organic fertilization, crop rotation with leguminous crops is central to reducing the grain protein yield gap with conventional fertilization. The lack of economic viability observed in a broader economic market could be overcome by promoting high quality products at fair prices with short supply-chain distribution, following some initiatives being carried out in Italy (Piani et al. 2021). The susceptibility to pests in agroecological wheat production can be, to some extent, overcome by increasing biodiversity-ecosystem services, but increases in biological control research and application are also central. A limiting factor is also the lack of studies providing evidence and management alternatives to reduce the yield gaps with conventional agriculture; this deficiency is especially notorious in mainland Greece and central Spain. The novel study we present here evidences the benefits and limitations of wheat agroecological production in the Euro-Mediterranean and illustrates some of the spatial and management factors that can guide wheat agroecological research and policy within the framework of the current climate crisis.

## **Declarations**

**Acknowledgments:** We acknowledge the support of the project PID2019-106650RB-C21 from the Ministerio de Ciencia e Innovacion, Spain. J.S. is a recipient of a FPI doctoral fellowship from the same institution (grant: PRE2020- 091907). J.L.A. acknowledges support from the Institutio Catalana de Recerca i Estudis Avançats (ICREA), Generalitat de Catalunya, Spain. S.C.K. is supported by the Ramon y Cajal RYC-2019-027818-I research fellowship from the Ministerio de Ciencia e Innovacion, Spain.

**Author's contribution:** Joel Segarra conceptualized the study. Joel Segarra developed the methodology, conducted the literature search, and analyzed the spatial data. Jose Luis Araus and Shawn Carlisle Kefauver provided feedback during the search and analysis processes. The first draft of the manuscript was written by Joel Segarra, and Jose Luis Araus and Shawn Carlisle Kefauver reviewed and approved the final version of the manuscript before submission.

**Funding:** This article is part of Joel Segarra PhD studies funded by the Ministry of Science and Innovation (Spain) (grant: PRE2020- 091907).

**Data Availability:** All data used is publicly available. The articles analyzed and spatial data is presented in the supplementary material.

**Code Availability:** Not applicable.

**Conflict of interest:** The authors declare no competing interests.

**Ethics approval and consent to participate:** Not applicable.

**Consent for publication:** Not applicable.

## 6. References

- Aguilera E, Díaz-Gaona C, García-Laureano R, et al (2020) Agroecology for adaptation to climate change and resource depletion in the Mediterranean region. A review. *Agric Syst* 181:102809. <https://doi.org/10.1016/j.agsy.2020.102809>
- Altieri MA, Nicholls CI (2020) Agroecology: challenges and opportunities for farming in the Anthropocene. *International Journal of Agriculture and Natural Resources* 47:204–215. <https://doi.org/10.7764/ijanr.v47i3.2281>
- Bento VA, Ribeiro AFS, Russo A, et al (2021) The impact of climate change in wheat and barley yields in the Iberian Peninsula. *Sci Rep* 11:1–12. <https://doi.org/10.1038/s41598-021-95014-6>
- Biaggini M, Consorti R, Dapporto L, et al (2007) The taxonomic level order as a possible tool for rapid assessment of Arthropod diversity in agricultural landscapes. *Agric Ecosyst Environ* 122:183–191. <https://doi.org/10.1016/j.agee.2006.12.032>
- Billen G, Aguilera E, Einarsson R, et al (2021) Reshaping the European agro-food system and closing its nitrogen cycle: The potential of combining dietary change, agroecology, and circularity. *One Earth* 4:839–850. <https://doi.org/10.1016/j.oneear.2021.05.008>
- Blas A, Garrido A, Unver O, Willaarts B (2019) A comparison of the Mediterranean diet and current food consumption patterns in Spain from a nutritional and water perspective. *Science of The Total Environment* 664:1020–1029. <https://doi.org/10.1016/j.scitotenv.2019.02.111>
- Bux C, Lombardi M, Varese E, Amicarelli V (2022) Economic and Environmental Assessment of Conventional versus Organic Durum Wheat Production in Southern Italy. *Sustainability* 14:9143. <https://doi.org/10.3390/su14159143>
- Caballero-López B, Blanco-Moreno JM, Pujade-Villar J, et al (2016) Herbivores, saprovores and natural enemies respond differently to within-field plant characteristics of wheat fields. *J Insect Conserv* 20:467–476. <https://doi.org/10.1007/s10841-016-9879-5>
- Cameron KC, Di HJ, Moir JL (2013) Nitrogen losses from the soil/plant system: a review. *Annals of Applied Biology* 162:145–173. <https://doi.org/10.1111/aab.12014>
- Canfield M, Anderson MD, McMichael P (2021) UN Food Systems Summit 2021: Dismantling Democracy and Resetting Corporate Control of Food Systems. *Front Sustain Food Syst* 5:. <https://doi.org/10.3389/fsufs.2021.661552>
- Carranza-Gallego G, Guzmán GI, García-Ruíz R, et al (2018a) Contribution of old wheat varieties to climate change mitigation under contrasting managements and rainfed Mediterranean conditions. *J Clean Prod* 195:111–121. <https://doi.org/10.1016/j.jclepro.2018.05.188>
- Carranza-Gallego G, Guzmán GI, Garcia-Ruíz R, et al (2019) Addressing the Role of Landraces in the Sustainability of Mediterranean Agroecosystems. *Sustainability* 11:6029. <https://doi.org/10.3390/su11216029>

- Carranza-Gallego G, Guzmán GI, Soto D, et al (2018b) Modern Wheat Varieties as a Driver of the Degradation of Spanish Rainfed Mediterranean Agroecosystems throughout the 20th Century. *Sustainability* 10:3724. <https://doi.org/10.3390/su10103724>
- Casagrande M, David C, Valantin-Morison M, et al (2009) Factors limiting the grain protein content of organic winter wheat in south-eastern France: a mixed-model approach. *Agron Sustain Dev* 29:565–574. <https://doi.org/10.1051/agro/2009015>
- Casals J, Rull A, Segarra J, et al (2019) Participatory plant breeding and the evolution of landraces: A case study in the organic farms of the Collserola Natural Park. *Agronomy* 9:. <https://doi.org/10.3390/agronomy9090486>
- Ceccarelli S (2011) Landraces: Importance and use in breeding and environmentally friendly agronomic systems. In: *Agrobiodiversity Conservation: Securing the Diversity of Crop Wild Relatives and Landraces*
- Ceccarelli S (1994) Specific adaptation and breeding for marginal conditions. In: *Breeding Fodder Crops for Marginal Conditions*. pp 101–127
- Chabert A, Sarthou J-P (2017) Practices of conservation agriculture prevail over cropping systems and landscape heterogeneity in understanding the ecosystem service of aphid biocontrol. *Agric Ecosyst Environ* 249:70–79. <https://doi.org/10.1016/j.agee.2017.08.005>
- Chamorro L, Masalles RM, Sans FX (2016) Arable weed decline in Northeast Spain: Does organic farming recover functional biodiversity? *Agric Ecosyst Environ* 223:1–9. <https://doi.org/10.1016/j.agee.2015.11.027>
- Connor DJ, Mínguez MI (2012) Evolution not revolution of farming systems will best feed and green the world. *Glob Food Sec* 1:106–113. <https://doi.org/10.1016/j.gfs.2012.10.004>
- Cos J, Doblas-Reyes F, Jury M, et al (2022) The Mediterranean climate change hotspot in the CMIP5 and CMIP6 projections. *Earth System Dynamics* 13:321–340. <https://doi.org/10.5194/esd-13-321-2022>
- Costanzo A, Bàrberi P (2014) Functional agrobiodiversity and agroecosystem services in sustainable wheat production. A review. *Agron Sustain Dev* 34:327–348. <https://doi.org/10.1007/s13593-013-0178-1>
- Costanzo A, Bàrberi P (2016) Field scale functional agrobiodiversity in organic wheat: Effects on weed reduction, disease susceptibility and yield. *European Journal of Agronomy* 76:1–16. <https://doi.org/10.1016/j.eja.2016.01.012>
- d’Andrimont R, Verhegghen A, Lemoine G, et al (2021) From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. *Remote Sens Environ* 266:. <https://doi.org/10.1016/j.rse.2021.112708>
- David C, Jeuffroy M-H, Henning J, Meynard J-M (2005) Yield variation in organic winter wheat: a diagnostic study in the Southeast of France. *Agron Sustain Dev* 25:213–223. <https://doi.org/10.1051/agro:2005016>
- De Ponti T, Rijk B, Van Ittersum MK (2012) The crop yield gap between organic and conventional agriculture. *Agric Syst* 108:1–9. <https://doi.org/10.1016/J.AGSY.2011.12.004>

- De Vita P, Mastrangelo AM, Codianni P, Fornara M (2007) Bio-agronomic Evaluation of Old and Modern Wheat, Spelt and Emmer Genotypes for Low-input Farming in Mediterranean Environment. *Italian Journal of Agronomy* 2:291. <https://doi.org/10.4081/ija.2007.291>
- De Vita P, Riefolo C, Codianni P, et al (2006) Agronomic and qualitative traits of *T. turgidum* ssp. *dicoccum* genotypes cultivated in Italy. *Euphytica* 150:195–205. <https://doi.org/10.1007/s10681-006-9107-6>
- Dedryver C-A, Le Ralec A, Fabre F (2010) The conflicting relationships between aphids and men: A review of aphid damage and control strategies. *C R Biol* 333:539–553. <https://doi.org/10.1016/j.crvi.2010.03.009>
- Di Bene C, Marchetti A, Francaviglia R, Farina R (2016) Soil organic carbon dynamics in typical durum wheat-based crop rotations of Southern Italy. *Italian Journal of Agronomy* 11:209–216. <https://doi.org/10.4081/ija.2016.763>
- Dinelli G, Marotti I, Di Silvestro R, et al (2013) Agronomic, nutritional and nutraceutical aspects of durum wheat (*Triticum durum* Desf.) cultivars under low input agricultural management. *Italian Journal of Agronomy* 8:12. <https://doi.org/10.4081/ija.2013.e12>
- Dinerstein E, Olson D, Joshi A, et al (2017) An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *Bioscience* 67:534–545. <https://doi.org/10.1093/biosci/bix014>
- Donald PF, Green RE, Heath MF (2001) Agricultural intensification and the collapse of Europe's farmland bird populations. *Proc R Soc Lond B Biol Sci* 268:25–29. <https://doi.org/10.1098/rspb.2000.1325>
- Espluga-Trenc J, Calvet-Mir L, López-García D, et al (2021) Local agri-food systems as a cultural heritage strategy to recover the sustainability of local communities. Insights from the spanish case. *Sustainability (Switzerland)* 13:. <https://doi.org/10.3390/su13116068>
- Evenson RE, Gollin D (2003) Assessing the Impact of the Green Revolution, 1960 to 2000. *Science* (1979) 300:758–762. <https://doi.org/10.1126/science.1078710>
- García-Martín M, Torralba M, Quintas-Soriano C, et al (2021) Linking food systems and landscape sustainability in the Mediterranean region. *Landsc Ecol* 36:2259–2275. <https://doi.org/10.1007/s10980-020-01168-5>
- Geiger F, de Snoo GR, Berendse F, et al (2010) Landscape composition influences farm management effects on farmland birds in winter: A pan-European approach. *Agric Ecosyst Environ* 139:571–577. <https://doi.org/10.1016/j.agee.2010.09.018>
- Giacomelli M, Sargolini M, Felipe-Lucia MR (2024) Including the perspective of stakeholders in landscape planning through the Ecosystem Services co-production framework: an empirical exploration in Le Marche, Italy. *Reg Environ Change* 24:24. <https://doi.org/10.1007/s10113-024-02184-w>
- Giraldo OF, Rosset PM (2018) Agroecology as a territory in dispute: between institutionality and social movements. *J Peasant Stud* 45:545–564. <https://doi.org/10.1080/03066150.2017.1353496>

- González-Estébanez FJ, García-Tejero S, Mateo-Tomás P, Olea PP (2011) Effects of irrigation and landscape heterogeneity on butterfly diversity in Mediterranean farmlands. *Agric Ecosyst Environ* 144:262–270. <https://doi.org/10.1016/j.agee.2011.09.002>
- Iannucci A, Fares C, Codianni P (2018) Influence of nitrogen levels on bio-agronomic and quality traits of tetraploid wheats under organic farming. *Annals of Applied Biology* 173:1–15. <https://doi.org/10.1111/aab.12429>
- Jiménez-Olivencia Y, Ibáñez-Jiménez Á, Porcel-Rodríguez L, Zimmerer K (2021) Land use change dynamics in Euro-mediterranean mountain regions: Driving forces and consequences for the landscape. *Land use policy* 109:. <https://doi.org/10.1016/j.landusepol.2021.105721>
- Jose-Maria L, Sans FX (2011) Weed seedbanks in arable fields: effects of management practices and surrounding landscape. *Weed Res* 51:631–640. <https://doi.org/10.1111/j.1365-3180.2011.00872.x>
- Lago-Oliveira S, Rebolledo-Leiva R, Garofalo P, et al (2023) Environmental and economic benefits of wheat and chickpea crop rotation in the Mediterranean region of Apulia (Italy). *Science of The Total Environment* 896:165124. <https://doi.org/10.1016/j.scitotenv.2023.165124>
- Lark TJ, Spawn SA, Bougie M, Gibbs HK (2020) Cropland expansion in the United States produces marginal yields at high costs to wildlife. *Nat Commun* 11:4295. <https://doi.org/10.1038/s41467-020-18045-z>
- Lazzaro M, Bàrberi P, Dell’Acqua M, et al (2019) Unraveling diversity in wheat competitive ability traits can improve integrated weed management. *Agron Sustain Dev* 39:6. <https://doi.org/10.1007/s13593-018-0551-1>
- Levers C, Schneider M, Prishchepov A V., et al (2018) Spatial variation in determinants of agricultural land abandonment in Europe. *Science of The Total Environment* 644:95–111. <https://doi.org/10.1016/j.scitotenv.2018.06.326>
- Li K, Rollins J, Yan E (2018) Web of Science use in published research and review papers 1997–2017: a selective, dynamic, cross-domain, content-based analysis. *Scientometrics* 115:1–20. <https://doi.org/10.1007/s11192-017-2622-5>
- Lopes MS, El-Basyoni I, Baenziger PS, et al (2015) Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *J Exp Bot* 66:3477–3486. <https://doi.org/10.1093/jxb/erv122>
- Martínez-Mena M, Carrillo-López E, Boix-Fayos C, et al (2020) Long-term effectiveness of sustainable land management practices to control runoff, soil erosion, and nutrient loss and the role of rainfall intensity in Mediterranean rainfed agroecosystems. *Catena (Amst)* 187:104352. <https://doi.org/10.1016/j.catena.2019.104352>
- Migliorini P, Gkissakis V, Gonzalez V, et al (2018) Agroecology in Mediterranean Europe: Genesis, State and Perspectives. *Sustainability* 10:2724. <https://doi.org/10.3390/su10082724>
- Migliorini P, Moschini V, Tittarelli F, et al (2014) Agronomic performance, carbon storage and nitrogen utilisation of long-term organic and conventional stockless arable systems in Mediterranean area. *European Journal of Agronomy* 52:138–145. <https://doi.org/10.1016/j.eja.2013.09.017>



- Migliorini P, Spagnolo S, Torri L, et al (2016) Agronomic and quality characteristics of old, modern and mixture wheat varieties and landraces for organic bread chain in diverse environments of northern Italy. *European Journal of Agronomy* 79:131–141. <https://doi.org/10.1016/j.eja.2016.05.011>
- Montemurro F, Maiorana M (2015) Agronomic Practices at Low Environmental Impact for Durum Wheat in Mediterranean Conditions. *J Plant Nutr* 38:624–638. <https://doi.org/10.1080/01904167.2014.988356>
- Moschini V, Migliorini P, Sacchetti P, et al (2012) Presence of aphid predators in common wheat (*Triticum aestivum* L.) in organic and conventional agroecosystems of Tuscany. *New Medit* 11:57–60
- Olson DM, Dinerstein E, Wikramanayake ED, et al (2001) Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience* 51:933–938. [https://doi.org/10.1641/0006-3568\(2001\)051\[0933:TEOTWA\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2)
- Padel S, Orsini S, Solfanelli F, Zanoli R (2021) Can the market deliver 100% organic seed and varieties in Europe? *Sustainability (Switzerland)* 13:1–13. <https://doi.org/10.3390/su131810305>
- Page MJ, McKenzie JE, Bossuyt PM, et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* n71. <https://doi.org/10.1136/bmj.n71>
- Panozzo A, Bernazeau B, Desclaux D (2020a) Durum wheat in organic olive orchard: good deal for the farmers? *Agroforestry Systems* 94:707–717. <https://doi.org/10.1007/s10457-019-00441-0>
- Panozzo A, Huang H, Bernazeau B, et al (2020b) Morphology, Phenology, Yield, and Quality of Durum Wheat Cultivated within Organic Olive Orchards of the Mediterranean Area. *Agronomy* 10:1789. <https://doi.org/10.3390/agronomy10111789>
- Pardo G, Aibar J, Cervero J, Zaragoza C (2009) Economic evaluation of cereal cropping systems under semiarid conditions: minimum input, organic and conventional. *Sci Agric* 66:615–621. <https://doi.org/10.1590/S0103-90162009000500005>
- Pardo G, Aibar J, Ciria P, et al (2011) Organic Versus Conventional Methods of Fertilization and Weed Control in a Long Term Rotation of Cereals in Semiarid Spain. *Chil J Agric Res* 71:187–194. <https://doi.org/10.4067/S0718-58392011000200002>
- Pecetti L, Annicchiarico P, Damania AB (1992) Biodiversity in a germplasm collection of durum wheat. *Euphytica* 60:229–238. <https://doi.org/10.1007/BF00039403>
- Pérez-Fuertes O, García-Tejero S, Pérez Hidalgo N, et al (2015) Irrigation effects on arthropod communities in Mediterranean cereal agro-ecosystems. *Annals of Applied Biology* 167:236–249. <https://doi.org/10.1111/aab.12223>
- Piani L, Carzedda M, Carestiatto N (2021) Food solidarity economy: evaluating transition community initiatives in Friuli Venezia Giulia region. *Agricultural and Food Economics* 9:32. <https://doi.org/10.1186/s40100-021-00203-6>
- Ponce C, Bravo C, de León DG, et al (2011) Effects of organic farming on plant and arthropod communities: A case study in Mediterranean dryland cereal. *Agric Ecosyst Environ* 141:193–201. <https://doi.org/10.1016/j.agee.2011.02.030>

- Ponti L, Gutierrez AP, Altieri MA (2016) Preserving the Mediterranean Diet Through Holistic Strategies for the Conservation of Traditional Farming Systems. *Environmental History (Netherlands)* 5:453–469. [https://doi.org/10.1007/978-3-319-26315-1\\_24](https://doi.org/10.1007/978-3-319-26315-1_24)
- Queiroz C, Beilin R, Folke C, Lindborg R (2014) Farmland abandonment: threat or opportunity for biodiversity conservation? A global review. *Front Ecol Environ* 12:288–296. <https://doi.org/10.1890/120348>
- Reynolds M, Balota M, Delgado M, et al (1994) Physiological and Morphological Traits Associated With Spring Wheat Yield Under Hot, Irrigated Conditions. *Functional Plant Biology* 21:717. <https://doi.org/10.1071/PP9940717>
- Rosegrant MW, Ringler C, Zhu T (2009) Water for Agriculture: Maintaining Food Security under Growing Scarcity. *Annu Rev Environ Resour* 34:205–222. <https://doi.org/10.1146/annurev.enviro.030308.090351>
- Rounsevell MDA, Annetts JE, Audsley E, et al (2003) Modelling the spatial distribution of agricultural land use at the regional scale. *Agric Ecosyst Environ* 95:465–479. [https://doi.org/10.1016/S0167-8809\(02\)00217-7](https://doi.org/10.1016/S0167-8809(02)00217-7)
- Royo C, Soriano JM, Alvaro F (2017) Wheat: A Crop in the Bottom of the Mediterranean Diet Pyramid. In: *Mediterranean Identities - Environment, Society, Culture*. InTech
- Sanchez-Garcia M, Álvaro F, Peremarti A, et al (2015) Changes in bread-making quality attributes of bread wheat varieties cultivated in Spain during the 20th century. *European Journal of Agronomy* 63:79–88. <https://doi.org/10.1016/j.eja.2014.11.006>
- Sapkota TB, Mazzoncini M, Bàrberi P, et al (2012) Fifteen years of no till increase soil organic matter, microbial biomass and arthropod diversity in cover crop-based arable cropping systems. *Agron Sustain Dev* 32:853–863. <https://doi.org/10.1007/s13593-011-0079-0>
- Segarra J, Araus JL, Kefauver SC (2023a) Suitability mapping and management monitoring in Castilian organic and conventional wheat fields with Sentinel-2 and spatial data. *Ecol Inform* 78:102347. <https://doi.org/10.1016/j.ecoinf.2023.102347>
- Segarra J, Fernández-Martínez J, Araus JL (2023b) Managing abandoned Mediterranean mountain landscapes: The effects of donkey grazing on biomass control and floral diversity in pastures. *Catena (Amst)* 233:. <https://doi.org/10.1016/j.catena.2023.107503>
- Siciliano G (2009) Social multicriteria evaluation of farming practices in the presence of soil degradation. A case study in Southern Tuscany, Italy. *Environ Dev Sustain* 11:1107–1133. <https://doi.org/10.1007/s10668-008-9169-9>
- Smil V (2004) *Enriching the earth: Fritz Haber, Carl Bosch, and the transformation of world food production*. MIT Press
- Soler-Montiel M, Montero M, Copena D, Pérez-Neira D (2024) Climate impact reduction potential of agroecology: a case study of traditional Andalusian recipes (Spain). *Agroecology and Sustainable Food Systems* 1–21. <https://doi.org/10.1080/21683565.2023.2300656>

- Stagnari F, Onofri A, Codianni P, Pisante M (2013) Durum wheat varieties in N-deficient environments and organic farming: a comparison of yield, quality and stability performances. *Plant Breeding* 132:266–275. <https://doi.org/10.1111/pbr.12044>
- Tarjuelo R, Morales MB, Arribas L, Traba J (2019) Abundance of weeds and seeds but not of arthropods differs between arable habitats in an extensive Mediterranean farming system. *Ecol Res* 34:624–636. <https://doi.org/10.1111/1440-1703.12029>
- Tilman D, Cassman KG, Matson PA, et al (2002) Agricultural sustainability and intensive production practices. *Nature* 418:671–677. <https://doi.org/10.1038/nature01014>
- Trewavas A (2001) Urban myths of organic farming. *Nature* 410:409–410. <https://doi.org/10.1038/35068639>
- Tubiello FN, Salvatore M, Ferrara AF, et al (2015) The Contribution of Agriculture, Forestry and other Land Use activities to Global Warming, 1990-2012. *Glob Chang Biol* 21:2655–2660. <https://doi.org/10.1111/gcb.12865>
- Tudisca S, di Trapani AM, Sgroi F, Testa R (2014) Organic farming and economic sustainability: The case of Sicilian durum wheat. *Quality - Access to Success* 15:
- Tuel A, Eltahir EAB (2020) Why Is the Mediterranean a Climate Change Hot Spot? *J Clim* 33:5829–5843. <https://doi.org/10.1175/JCLI-D-19-0910.1>
- Vanwalleghe T, Amate JJ, de Molina MG, et al (2011) Quantifying the effect of historical soil management on soil erosion rates in Mediterranean olive orchards. *Agric Ecosyst Environ* 142:341–351. <https://doi.org/10.1016/j.agee.2011.06.003>
- Varia F, Macaluso D, Vaccaro A, et al (2021) The Adoption of Landraces of Durum Wheat in Sicilian Organic Cereal Farming Analysed Using a System Dynamics Approach. *Agronomy* 11:319. <https://doi.org/10.3390/agronomy11020319>
- Wang J, Baranski M, Korkut R, et al (2021) Performance of Modern and Traditional Spelt Wheat (*Triticum spelta*) Varieties in Rain-Fed and Irrigated, Organic and Conventional Production Systems in a Semi-Arid Environment; Results from Exploratory Field Experiments in Crete, Greece. *Agronomy* 11:890. <https://doi.org/10.3390/agronomy11050890>
- Winkler K, Fuchs R, Rounsevell M, Herold M (2021) Global land use changes are four times greater than previously estimated. *Nat Commun* 12:1–10. <https://doi.org/10.1038/s41467-021-22702-2>
- Wolfe MS (2000) Crop strength through diversity. *Nature* 406:681–682. <https://doi.org/10.1038/35021152>
- Zimmerer KS, Aumeeruddy-Thomas Y, Caillon S, et al (2024) Agrobiodiversity threats amid expanding woody monocultures and hopes nourished through farmer and food movements in the Mediterranean. *Elem Sci Anth* 12:. <https://doi.org/10.1525/elementa.2023.00093>
- (2021) Commission communication: action plan for the development of organic production in the EU

## Supplementary material

**Table S1.** References of case studies.

N°Article	Title	Author	Year	Journal	Wheat type	Location
1	Yield variation in organic winter wheat: a diagnostic study in the Southeast of France	David et al.	2005	Agronomy For Sustainable Development	Bread	Southeastern France
2	The taxonomic level order as a possible tool for rapid assessment of Arthropod diversity in agricultural landscapes	Biaggini et al.	2007	Agriculture Ecosystems & Environment	Durum	Tuscany, Italy
3	Bio-agronomic Evaluation of Old and Modern Wheat, Spelt and Emmer Genotypes for Low-input Farming in Mediterranean Environment	De Vita et al.	2007	Italian Journal Of Agronomy	Emmer, spelt, bread, and durum	Italy
4	Factors limiting the grain protein content of organic winter wheat in south-eastern France: a mixed-model approach	Casagrande et al.	2009	Agronomy For Sustainable Development	Bread	South Eastern France
5	Crop yield and grain quality of emmer populations grown in central Italy, as affected by nitrogen fertilization	Marino et al.	2009	European Journal Of Agronomy	Emmer	Molise, Italy
6	Economic evaluation of cereal cropping systems under semiarid	Pardo et al.	2009	Scientia Agricola	Durum	Aragon, Spain

	conditions: minimum input, organic and conventional					
7	Social multicriteria evaluation of farming practices in the presence of soil degradation. A case study in Southern Tuscany, Italy	Siciliano	2009	Environment Development And Sustainability	Durum	Tuscany, Italy
8	Nitrogen Use Efficiency and Nitrogen Fertilizer Recovery of Durum Wheat Genotypes as Affected by Interspecific Competition	Giambalvo et al.	2010	Agronomy Journal	Durum	Sicily, Italy
9	Response of common wheat varieties to organic and conventional production systems across Italian locations, and implications for selection	Annicchiarico et al.	2010	Field Crops Research	Bread	North and Center Italy
10	Durum wheat-faba bean temporary intercropping: Effects on nitrogen supply and wheat quality	Tosti and Guiducci	2010	European Journal Of Agronomy	Durum	Papiano, Italy
11	Landscape composition influences farm management effects on farmland birds in winter: A pan-European approach	Geiger et al .	2010	Agriculture Ecosystems & Environment	Bread	Spain

12	Weed seedbanks in arable fields: effects of management practices and surrounding landscape	Jose-Maria and Sans	2011	Weed Research	Bread	Catalonia, Spain
13	A novel index of land use intensity for organic and conventional farming of Mediterranean cereal fields	Armengot et al.	2011	Agronomy For Sustainable Development	Bread	Catalonia, Spain
14	Effects of organic farming on plant and arthropod communities: A case study in Mediterranean dryland cereal	Ponce et al.	2011	Agriculture Ecosystems & Environment	Bread	Castile, Spain
15	Organic versus conventional methods of fertilization and weed control in a long term rotation of cereals in semiarid Spain	Pardo . et al.	2011	Chilean Journal Of Agricultural Research	Durum	Aragon, Spain
16	Performance of Wheat Varieties (Triticum aestivum L.) under Conservation Tillage Practices in Organic Agriculture	Bilalis et al.	2011	Notulae Botanicae Horti Agrobotanici Cluj-Napoca	Bread	Athens, Greece
17	Diversity of different farmer and modern wheat varieties cultivated in contrasting organic farming conditions in western Europe and implications for European seed	Serpolay et al.	2011	Organic Agriculture	Bread	Southeastern France and Pescara, Italy

	and variety legislation					
18	Durum Wheat in Conventional and Organic Farming: Yield Amount and Pasta Quality in Southern Italy	Fagnano et al.	2012	Scientific World Journal	Durum	Naples, Italy
19	Presence of aphid predators in common wheat ( <i>Triticum aestivum</i> L.) in organic and conventional agroecosystems of Tuscany	Moschini et al.	2012	New Medit	Bread	Florence, Italy
20	Phenotypic diversity and evolution of farmer varieties of bread wheat on organic farms in Europe	Dawson et al.	2012	Genetic Resources And Crop Evolution	Bread	Southeastern France and Italy
21	Fifteen years of no till increase soil organic matter, microbial biomass and arthropod diversity in cover crop-based arable cropping systems	Sapkota et al.	2012	Agronomy For Sustainable Development	Durum	Tuscany, Italy
22	Agronomic, nutritional and nutraceutical aspects of durum wheat ( <i>Triticum durum</i> Desf.) cultivars under low input agricultural management	Dinelli et al.	2013	Italian Journal Of Agronomy	Durum	Bologna, Italy
23	Durum wheat varieties in N-deficient environments and organic farming: a comparison of yield, quality and	Stagnari et al.	2013	Plant Breeding	Durum	Foggia, Italy

	stability performances					
24	Agronomic performance, carbon storage and nitrogen utilisation of long-term organic and conventional stockless arable systems in Mediterranean area	Migliorini et al.	2014	European Journal Of Agronomy	Bread	Florence, Italy
25	Contribution of relay intercropping with legume cover crops on nitrogen dynamics in organic grain systems	Amossé et al.	2014	Nutrient Cycling In Agroecosystems	Bread	Southeastern France
26	Organic farming and economic sustainability: The case of Sicilian durum wheat	Tudisca et al.	2014	Quality-Access To Success	Durum	Sicily, Italy
27	Agronomic practices at low environmental impact for durum wheat in mediterranean conditions	Montemurro et al.	2015	Journal Of Plant Nutrition	Durum	Foggia, Italy
28	irrigation effects on arthropod communities in Mediterranean cereal agro-ecosystems	Perez-Fuertes et al.	2015	Annals Of Applied Biology	Bread	Castile, Spain
29	Impact of alley cropping agroforestry on stocks, forms and spatial distribution of soil organic carbon - A case study in a Mediterranean context	Cardinael et al.	2015	Geoderma	Durum	Occitania, France



30	Agronomic practices at low environmental impact for durum wheat in mediterranean conditions	Montemurro & Maiorana	2015	Journal Of Plant Nutrition	Durum	Apulia, Italy
31	Environmental evaluation and benchmarking of the traditional dryland Mediterranean crop farming system in the Alentejo region of Portugal	Rosado et al.	2015	International Journal Of Sustainable Society	Durum	Alentejo, Portugal
32	The long-term effects of conventional and organic cropping systems, tillage managements and weather conditions on yield and grain quality of durum wheat ( <i>Triticum durum</i> Desf.) in the Mediterranean environment of Central Italy	Campiglia et al.	2015	Field Crops Research	Durum	Latium, Italy
33	Field scale functional agrobiodiversity in organic wheat: Effects on weed reduction, disease susceptibility and yield	Costanzo and Barberi	2016	European Journal Of Agronomy	Bread	Pisa, Italy
34	Nutritional characteristics of ancient Tuscan varieties of <i>Triticum aestivum</i> L.	Ghiselli et al.	2016	Italian Journal Of Agronomy	Bread	Siena, Italy
35	Agronomic and quality characteristics of old, modern and	Migliorini et al.	2016	European Journal Of Agronomy	Bread	Piamonte, Italy

	mixture wheat varieties and landraces for organic bread chain in diverse environments of northern Italy					
36	Arable weed decline in Northeast Spain: Does organic farming recover functional biodiversity?	Chamorro et al.	2016	Agriculture Ecosystems & Environment	Bread	Catalonia, Spain
37	Long-term evaluation of productivity, stability and sustainability for cropping systems in Mediterranean rainfed conditions	Bonciarelli et al.	2016	European Journal Of Agronomy	Bread	Tuscany, Italy
38	Herbivores, saprovores and natural enemies respond differently to within-field plant characteristics of wheat fields	Caballero-Lopez, et al.	2016	Journal Of Insect Conservation	Bread	Catalonia, Spain
39	Soil organic carbon dynamics in typical durum wheat-based crop rotations of Southern Italy	Di Bene, et al.	2016	Italian Journal Of Agronomy	Durum	Apulia, Italy
40	Grain yield and competitive ability against weeds in modern and heritage common wheat cultivars are differently influenced by sowing density	Lazzaro et al.	2017	Italian Journal Of Agronomy	Bread	Pisa, Italy
41	Practices of conservation agriculture prevail over cropping	Chabert and Sarthou	2017	Agriculture Ecosystems & Environment	Bread	Occitania, France

	systems and landscape heterogeneity in understanding the ecosystem service of aphid biocontrol					
42	Influence of durum wheat-faba bean intercrop on specific quality traits of organic durum wheat	De Stefanis et al.	2017	Biological Agriculture & Horticulture	Durum	Italy
43	Single vs multiple agroecosystem services provided by common wheat cultivar mixtures: Weed suppression, grain yield and quality	Lazzaro et al.	2018	Field Crops Research	Bread	Pisa, Italy
44	Influence of nitrogen levels on bio-agronomic and quality traits of tetraploid wheats under organic farming	Iannucci et al.	2018	Annals Of Applied Biology	Emmer	Foggia, Italy
45	Modern Wheat Varieties as a Driver of the Degradation of Spanish Rainfed Mediterranean Agroecosystems throughout the 20th Century	Carranza-Gallego et al.	2018	Sustainability	Durum	Andalusia, Spain
46	Contribution of old wheat varieties to climate change mitigation under contrasting managements and rainfed Mediterranean conditions	Carranza-Gallego et al.	2018	Journal Of Cleaner Production	Durum	Andalusia, Spain
47	The effect of intercropping on the efficiency of	Kaci et al.	2018	Plant Environment Soil	Durum	Occitania, France

	fababean symbiosis and durum wheat					
48	Assessing naturalness of arable weed communities: a new index applied to a case study in central Italy	Fanfarillo et al.	2018	Biological Agriculture & Horticulture	Bread	Latium, Italy
49	Unraveling diversity in wheat competitive ability traits can improve integrated weed management	Lazzaro et al.	2019	Agronomy For Sustainable Development	Bread	Lombardia, Italy
50	Addressing the Role of Landraces in the Sustainability of Mediterranean Agroecosystems	Carranza-Gallego et al.	2019	Sustainability	Durum	Andalusia, Spain
51	Abundance of weeds and seeds but not of arthropods differs between arable habitats in an extensive Mediterranean farming system	Tarjuelo et al.	2019	Ecological Research	Bread	Castile, Spain
52	Wheat genotypic diversity and intercropping to control cereal aphids	Mansion-Vaquie et al.	2019	Agriculture Ecosystems & Environment	Bread	Southeastern France
53	Organic Cropping System Affects Grain Chemical Composition, Rheological and Agronomic Performance of Durum Wheat	Pandino et al.	2020	Agriculture-Basel	Durum	Sicily, Italy
54	Morphology, Phenology, Yield, and Quality of Durum Wheat Cultivated within	Panozzo et al.	2020	Agronomy-Basel	Durum	Occitania, France

	Organic Olive Orchards of the Mediterranean Area					
55	Long-term effectiveness of sustainable land management practices to control runoff, soil erosion, and nutrient loss and the role of rainfall intensity in Mediterranean rainfed agroecosystems	Martinez-Mena et al.	2020	Catena	Not specified	Murcia, Spain
56	Genetic Diversity and Stability of Performance of Wheat Population Varieties Developed by Participatory Breeding	Van Frank et al.	2020	Sustainability	Bread	Southeastern France
57	Durum wheat in organic olive orchard: good deal for the farmers?	Panozzo et al.	2020	Agroforestry Systems	Durum	Occitania, France
58	Yield, yield stability and farmers' preferences of evolutionary populations of bread wheat: A dynamic solution to climate change	Bocci et al.	2020	European Journal Of Agronomy	Bread	Italy
59	Crop and Soil Response to Organic Management Under Mediterranean Conditions	Leogrande, et al.	2020	International Journal Of Plant Production	Durum	Apulia, Italy
60	Performance of Modern and Traditional Spelt Wheat (Triticum	Wang et al.	2021	Agronomy-Basel	Spelt	Crete, Greece

	spelta) Varieties in Rain-Fed and Irrigated, Organic and Conventional Production Systems in a Semi-Arid Environment; Results from Exploratory Field Experiments in Crete, Greece					
61	The Adoption of Landraces of Durum Wheat in Sicilian Organic Cereal Farming Analysed Using a System Dynamics Approach	Varia et al.	2021	Agronomy-Basel	Durum	Sicily, Italy
62	Intra- and Inter-Population Genetic Diversity of Russello and Timilia Landraces from Sicily: A Proxy towards the Identification of Favorable Alleles in Durum Wheat	Taranto et al.	2022	Agronomy-Basel	Durum	Sicily, Italy
63	Grain Composition and Quality in Portuguese Triticum aestivum Germplasm Subjected to Heat Stress after Anthesis	Scotti-Campos et al.	2022	Plants	Bread	Alentejo, Portugal
64	Agronomic and ecophysiological evaluation of an early establishment of perennial wheat lines in Central Italy	Baronti et al.	2022	Genetic Resources And Crop Evolution	Perennial wheat	Latium, Italy
65	Economic and Environmental Assessment of	Bux et al.	2022	Sustainability	Durum	Apulia, Italy

	Conventional versus Organic Durum Wheat Production in Southern Italy					
66	A comprehensive assessment of diversified cropping systems on agro-environmental sustainability in three Mediterranean long-term field experiments	Vanino et al.	2022	European Journal Of Agronomy	Durum and Bread	Apulia, Italy, and Aragon, Spain
67	Environmental and economic benefits of wheat and chickpea crop rotation in the Mediterranean region of Apulia (Italy)	Lago-Oliveira et al.	2023	Science Of The Total Environment	Durum	Apulia, Italy
68	Effects of Different Types of Soil Management on Organic Carbon and Nitrogen Contents and the Stability Index of a Durum Wheat-Faba Bean Rotation under a Mediterranean Climate	Tedone, L.	2023	Agronomy	Durum	Basilicata, Italy
69	Suitability mapping and management monitoring in Castilian organic and conventional wheat fields with Sentinel-2 and spatial data	Segarra et al.	2023	Ecological Informatics	Bread	Castile, Spain

**Table S2.** Cropland area in the Euro-Mediterranean ecoregions

Euro-Mediterranean ecoregion	Cropland area (ha)			
	Stable	Loss	Gain	Gain and loss
Aegean and Western Turkey sclerophyllous and mixed forests	2,141,800	330,100	316,500	347,400
Crete Mediterranean forests	98,300	71,500	21,000	49,300
Iberian conifer forests	488,900	149,500	78,200	488,000
Iberian sclerophyllous and semi-deciduous forests	10,147,700	3,581,600	577,100	6,294,300
Illyrian deciduous forests	57,900	171,600	11,800	211,700
Italian sclerophyllous and semi-deciduous forests	3,195,500	1,920,900	60,400	1,421,200
Northeast Spain and Southern France Mediterranean forests	411,400	522,300	92,100	1,568,900
Northwest Iberian montane forests	1,070,000	325,800	78,100	969,400
Pindus Mountains mixed forests	284,900	155,300	31,300	89,200
South Apennine mixed montane forests	114,500	211,600	3,600	151,700
Southwest Iberian Mediterranean sclerophyllous and mixed forests	1,775,400	1,023,200	80,600	937,400
Tyrrhenian-Adriatic sclerophyllous and mixed forests	1,597,300	1,352,800	81,800	1,355,700
Total	21,383,600	9,816,200	1,432,500	13,884,200





## **Capítol 2.** Mapatge de l'idoneïtat i monitoratge de camps de blat ecològics i convencionals a Castella

### **Chapter 2.** Suitability mapping and management monitoring in Castilian organic and conventional wheat fields with Sentinel-2 and spatial data

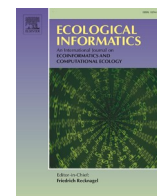
Joel Segarra<sup>a,b,\*</sup>, Jose Luis Araus<sup>a,b</sup>, Shawn Carlisle Kefauver<sup>a,b</sup>

<sup>a</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain

<sup>b</sup> AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198, Lleida, Catalonia, Spain

\*Corresponding author: Joel Segarra (e-mail: [joel.segarra@ub.edu](mailto:joel.segarra@ub.edu))

Publicat a Ecological Informatics



# Suitability mapping and management monitoring in Castilian organic and conventional wheat fields with Sentinel-2 and spatial data

Joel Segarra<sup>a,b,\*</sup>, Jose Luis Araus<sup>a,b</sup>, Shawn C. Kefauver<sup>a,b</sup>

<sup>a</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain

<sup>b</sup> AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198 Lleida, Catalonia, Spain

## ARTICLE INFO

### Keywords:

Organic agriculture  
Fuzzy interference system  
Envirotyping  
Satellite

## ABSTRACT

The Mediterranean region is a climate change hotspot due to projected precipitation decreases, soil erosion, and rising temperatures in the coming decades. In such a context, wheat productivity is expected to decrease; thus, all available croplands will need to be farmed to meet food demands. It is therefore crucial to have adequate management to help balance wheat production and environmental protection in specific target environments. With this aim, we present a study in Castile and Leon (Spain) in which we assess conventional and certified organic wheat grain yield and quality. Moreover, we spatially define wheat cultivation suitability categories and monitor crop rotation practices. For wheat suitability mapping, a fuzzy interference system was used to standardize topographic (height and slope), edaphic (pH, soil erosion, organic matter, and texture), and climatic (rainfall and temperature) variables and create marginal, suitable, and very suitable wheat cultivation suitability categories. We calculated the area of wheat suitability categories with confidence intervals and discussed factors affecting productivity. Additionally, we assessed the performance of certified organic and conventionally managed fields with organic and chemical fertilization. We also examined crop rotations using Sentinel-2 data over two seasons (2020–21). Certified organic management yielded 50% less than conventional management in similar environments while maintaining grain quality. Besides management, the main causes of the yield gap are the growing environment and variety. In conventional wheat management, organic fertilization achieves comparable yields to chemical fertilization. Crop rotation practices are uncommon, and over 50% of fields rotate wheat-to-barley or wheat-to-wheat, indicating poor soil management. The region has 25% of cropland classified as marginal wheat croplands ( $163,503 \pm 33,796$  ha) producing  $<2500$  kg/ha. Wheat organic agriculture has decreased yield gaps in marginal croplands compared to more productive areas. By farming organic wheat in marginal croplands, the best croplands might be used for more productive farming.

## 1. Introduction

Food production in the Mediterranean region has garnered considerable scientific and socioeconomic interest. This region is considered a climate change hotspot due to its projected decrease in precipitation and concurrent rise in world temperatures in the coming decades. Consequently, Mediterranean agriculture and ecosystems, which are currently experiencing water scarcity, would face greater challenges due to the decrease in precipitation (Tuel and Eltahir, 2020). Furthermore, it is worth noting that climate change has the potential to further extend the geographical distribution of Mediterranean climates in the Iberian Peninsula and other European regions (Beck et al., 2018). According to

Bento et al. (2021), it is anticipated that a rise in temperature of  $2^\circ\text{C}$  by the year 2040 will result in a decrease in wheat grain yield ranging from 15% to 30%. Therefore, in the context of climate change, it may be inferred that Mediterranean agricultural yields will decrease due to environmental factors unless there are advancements in technology, agronomy, and breeding methods. Consequently, there may be a need to farm all available cropland in order to compensate for these anticipated reductions.

Wheat is one of the most important crops worldwide (Peña, 2002) and is especially relevant in the Mediterranean region. It is therefore crucial to have adequate wheat cultivation management to ensure agricultural sustainability. Agriculture is a leading cause of negative

\* Corresponding author at: Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain.

E-mail address: [joel.segarra@ub.edu](mailto:joel.segarra@ub.edu) (J. Segarra).

<https://doi.org/10.1016/j.ecolinf.2023.102347>

Received 2 August 2023; Received in revised form 11 October 2023; Accepted 20 October 2023

Available online 30 October 2023

1574-9541/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

environmental transformations (Raven and Wagner, 2021) and global warming (Tubiello et al., 2015). Therefore, approaches that can help management balance production and environmental protection in specific target environments are pivotal. In the case of wheat, mapping cultivation suitability in contrasting environments and understanding performance under different managements can contribute to production sustainability. In this regard, we present a study in Castile and Leon (Spain) in which we compare conventional and certified organic wheat grain yield and quality, map wheat cultivation suitability, and monitor crop rotation practices. Mapping the suitability of wheat cultivation by combining environmental and land use factors can contribute to coping with yield reduction. In a wheat yield reduction scenario, all available croplands, either highly productive or marginal, will need to be cultivated to ameliorate the environmental and socioeconomic challenges posed by climate change. Hence, targeting the specific conditions of wheat-growing areas will increasingly require more attention to achieving food production targets and safeguarding ecosystems within their respective carrying capacities.

Different research approaches have been made to address wheat production requirements, from genotypic to phenotypic studies. In that respect, crop management and breeding efforts have led to a valuable burgeoning body of scientific literature and practical examples on selecting and obtaining specific traits and/or crop management practices adapted to drought, salinity, or pests, just to mention some (Araus et al., 2021; Barmerier and Schmidhalter, 2016; Rezzouk et al., 2020). However, many such studies have been generally limited in terms of experimental settings and sites (Resende et al., 2021) because of the complexities of having large datasets with actual field conditions across a wide environmental range. To overcome some of these limitations and to use most of the available data, Xu (2016) proposed envirotyping as a novel technical approach to help decipher environmental impacts on crops and overcome some of these limitations.

Envirotyping takes advantage of the most advanced geolocated sensors and available spatially interpolated data to increase data instances and modeling possibilities to help find optimal agricultural conditions for managing croplands. Within the envirotyping frame, we present a study in the central northern part of Spain to spatially define wheat suitability with regard to several typical environmental conditions. This area has a continental Mediterranean climate, with contrasting areas experiencing low and high rainfall in a relatively small region (García Fernández, 1986). The studied region comprises the current autonomous region of Castile and Leon and has been the traditional breadbasket of Spain. Current and future climate conditions urge the development of sub-regional-specific crop envirotyping approaches that can shed light on how to best define optimal breeding traits and agricultural management (Annicchiarico et al., 2005). Moreover, besides wheat suitability mapping, we have focused on monitoring soil sustainability practices by assessing crop rotations after wheat cultivation and have studied several wheat-growing fields with contrasting organic and conventional management. The European Union and its member states, such as Spain, have large geolocated datasets on relevant variables for crop production such as edaphology, topography, or climatology. In this study, we take advantage of publicly available datasets from the regional government and other international instances, as well as data derived from Sentinel-2, to envirotype wheat suitability and assess contrasting management approaches in the central northern part of Spain.

We have structured the study around three research questions:

1. What are the field-level grain yield and quality performances for contrasting management strategies in different environments? We have studied wheat grain yield and quality in organic and conventional managements with contrasting water regimes and fertilization conditions, which were further characterized using C and N isotopic compositions of grains. The C isotope composition ( $\delta^{13}\text{C}$ ), when analyzed in the mature grains of wheat, is an indicator of the water

status during crop growth (Araus et al., 2013, 2003). The natural abundance of the N isotope composition ( $\delta^{15}\text{N}$ ) in mature grains reveals the crop's source of nitrogen fertilization (Serret et al., 2008).

2. Are sustainable crop rotation practices being carried out in the region? We used Sentinel-2-derived crop type classification over two consecutive seasons (2020 and 2021) to determine crop rotation patterns. Crop rotations are directly linked to soil sustainability.
3. What is the spatial suitability of wheat cultivation in the region's croplands? To address this, we have used a fuzzy system to define suitability regarding topographic, edaphic, and climatic variables, which were subsequently divided into three classes (marginal, suitable, and very suitable) defined for the whole area.

## 2. Material and methods

### 2.1. Study area

The study area in Burgos and Palencia, in the central northern part of Spain (Castile and Leon), has a continental Mediterranean climate (García Fernández, 1986) with dry summers, intermittent precipitation, and wet winters. Annual rainfall ranges from 300 to 600 mm, and the average temperature is 11 to 15 °C. Mountains rising to 2000 m high surround the plateau-like region, which has an average elevation of 400 m above sea level. The region experiences droughts frequently, with exceptional droughts in recent years. Most wheat is farmed on rainfed croplands, which cover 6433 km<sup>2</sup>. We studied 36 wheat fields with contrasting management approaches (Fig. 1), namely certified organic and conventional. Conventional farming uses chemical fertilizers, pesticides, and herbicides, while certified organic farming employs manure and compost to fertilize the soil and uses environmental services and biological products to control pests. However, some conventionally managed fields are starting to implement sustainable practices such as organic fertilization while maintaining conventional pest management and using herbicides. As mentioned, most fields are rain-fed in the region; however, some of the studied fields had supplementary irrigation. Those fields were considered to have the maximum production range. In this regard, chemical fertilization was applied to 22 conventionally managed fields; 12 had supplementary irrigation, and 10 were rain-fed. Organic fertilization was applied to 11 conventionally managed fields from digestate, of which six had supplementary irrigation and five were rainfed. Three certified organic fields were rainfed. The bread wheat fields were located across the study area, but the spatial autocorrelation and distribution of study sites were not controlled. Organic management is limited in the region, making it difficult to obtain data from more fields for even field numbers. Average field sizes were 2–7 ha. Farmers reported the entire grain load from combine harvester-harvested wheat fields to the researchers, who adjusted for 10% moisture to compute grain yield for each site.

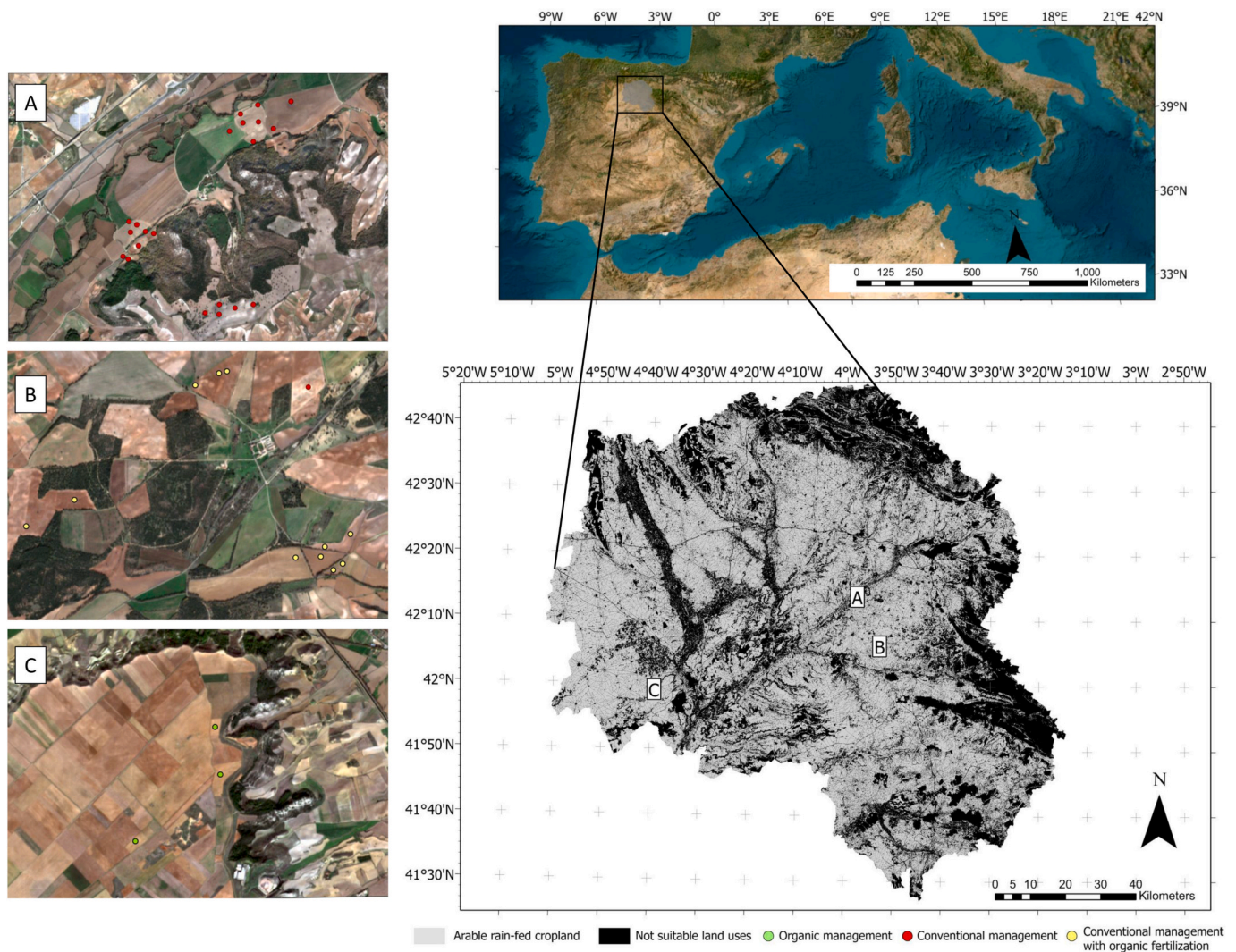
### 2.2. Carbon and nitrogen stable isotope composition

In the 36 fields studied, mature grains, collected at harvest, were dried at 60 °C for a minimum of 48 h finely ground using a grinder machine (Mixer Mill MM 400; Retsch GmbH, Haan; Germany), and then weighed in tin capsules (approximately 1 mg) for further analysis of the carbon and nitrogen stable isotope signatures and the total nitrogen and carbon contents. Stable isotope values were expressed in composition ( $\delta$ ) units, as the deviation of the isotopic composition of the material from the standard. Thus, the carbon and nitrogen isotopic compositions ( $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$ ) were expressed as:

$$\delta^{13}\text{C} \text{ or } \delta^{15}\text{N} (\text{‰}) = \left[ \left( \frac{R_{\text{sample}}}{R_{\text{standard}}} \right) - 1 \right] \times 1000 \quad (1)$$

Where the  $^{13}\text{C}/^{12}\text{C}$  and  $^{15}\text{N}/^{14}\text{N}$  ratios of the sample are notated as  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  and expressed in ‰, whereas  $R_{\text{standard}}$  is the molar abundance ratio of the secondary standard calibrated against the primary





**Fig. 1.** Location of the area of study within the Iberian Peninsula: the area is in the provinces of Palencia and Burgos (central north of Spain). Cropland uses and arable rainfed croplands are indicated on the map. The rest of the land uses, including urban, forest, and infrastructure, among others, are indicated as not suitable land uses. The data has been obtained from the regional government as freely accessible data (<https://www.itacyl.es/agro-y-geo-tecnologia/descarga-datos-geograficos/sigpac>). The locations of the 36 studied fields with contrasting management strategies (conventional and organic) in the study area are shown in the ABC labelled maps on the left. The background images in the ABC labelled maps corresponds to a Sentinel-2 satellite scene in true natural colour.

standard Pee Dee Belemnite ( $\delta^{13}\text{C}$ ) and  $\text{N}_2$  from air ( $\delta^{15}\text{N}$ ) (Farquhar et al., 1989). Different secondary standards were used for carbon (IAEA-CH7, IAEA-CH6 and IAEA-600, and USGS 40) and nitrogen (IAEA-600, N1, N2, NO3, UREA and Acetanilide) isotope analyses. Analytical precision of the  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  analyses were 0.1‰ and 0.3‰, respectively. Total carbon and nitrogen contents in grains were expressed as the percentage (%) of total carbon and nitrogen on dry matter basis. Isotopes and elemental analyses were performed employing an elemental analyzer operating in a continuous flow mode with a mass spectrometer (Delta C IRMS; ThermoFinnigan, Bremen; Germany), at the Scientific and Technical facilities of the University of Barcelona (Centres Científics i Tecnològics de la Universitat de Barcelona (CCiTUB)).

### 2.3. Statistical analysis

Statistical analyses were conducted using the open source software RStudio (R Core Team, 2021). Data for the set of grain yield, protein content and  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  were subjected to analyses of variance (ANOVAs) to test the effects of growing conditions (irrigation and fertilization management strategies). Differences were considered

significant at  $p$ -value  $\leq 0.05$ . Fisher's LSD (Least significant difference) test was used to determine post hoc differences for each contrasting condition.

### 2.4. Spatial data processing

In order to map wheat suitability, we used four criteria: topography, edaphology, climatology, and land use (Table 1). For the case of topography, we used two variables: slope and height. These variables are related to wheat root development, water runoff, or access to mechanization, which affect the development of the crop. We did not add aspect as the area is relatively flat with the surrounding mountains being mapped with the height data, hence a relatively even solar radiation was assumed for the studied croplands. These variables were derived from a digital elevation model elaborated by the regional government and were calculated in ArcGIS Pro 3.1.0. In regard to edaphology, we have used four variables (pH, organic matter, erosion, and texture). These edaphic variables have been collected by the Castile and Leon government based on extensive soil sampling throughout the region. These variables are pivotal for bionutrient accessibility and soil-water balance, among others. Regarding climatology, we used rainfall

**Table 1**

Spatial data used to map wheat suitability, 4 criterion, and nine variables obtained from the regional government soil maps website have been used (<https://www.atlas.itacyl.es/descarga>).

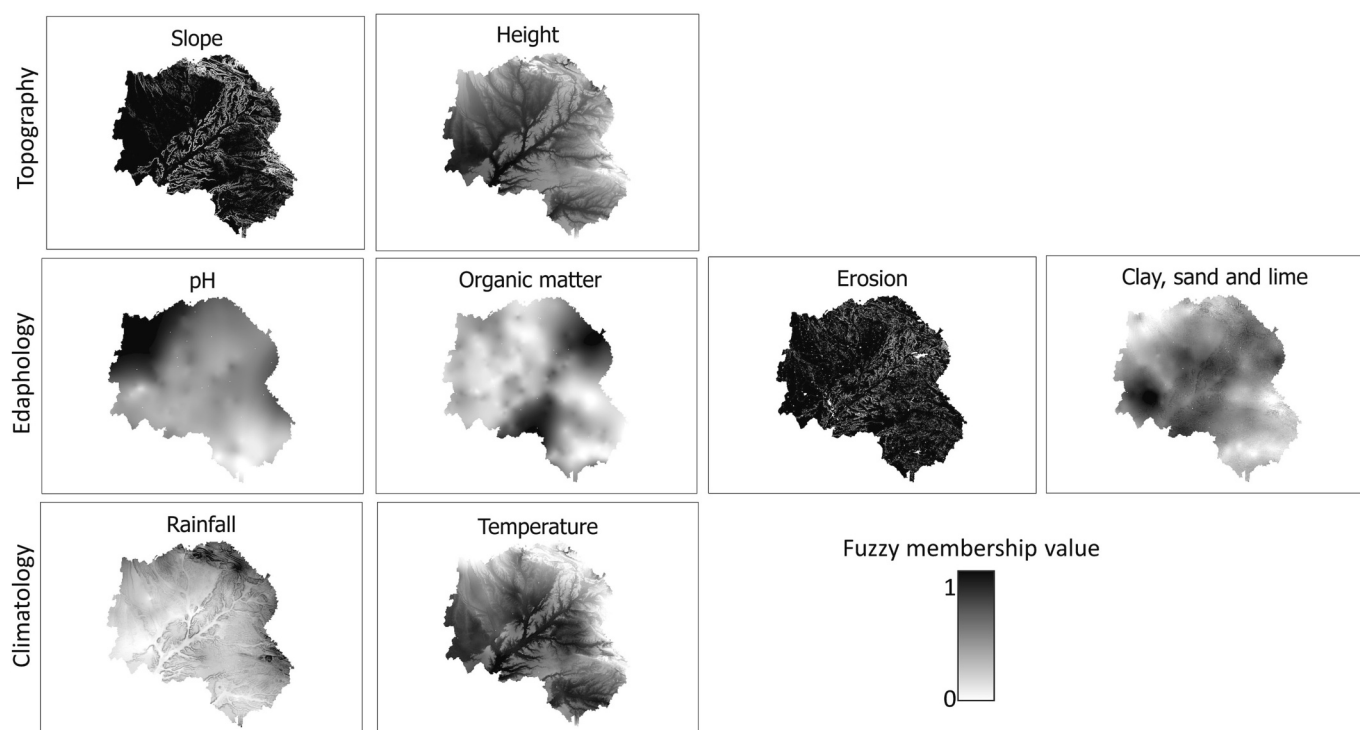
Criterion	Variables	Wheat physiology and agronomy	Reference
Topography	Slope (%)	Root development, water retention, runoff	(Kumhálová and Moudrý, 2014; Segarra et al., 2020b)
	Height (m)	Cold and freeze, access to mechanization	(Green et al., 2007)
Edaphology	pH	Nutrient bio-accessibility, microorganisms' activity	(Neina, 2019)
	Organic matter (%)	Soil quality, microorganisms' activity	(Masri and Ryan, 2006)
	Erosion (%)	Soil loss, root development, plant diseases	(Littleboy et al., 1992)
	Clay, sand, and lime (%)	Infiltration, structure development, soil-water balance	(He et al., 2014)
Climatology	Rainfall (mm)	Plant water status, growth, stress, grain yield	(Basso et al., 2012; Zhang and Oweis, 1999)
	Temperature (°C)	Plant water status, growth, stress, grain yield	(Porter and Gawith, 1999)
Land use	Rainfed croplands	Legal adequacy for rain-fed wheat cultivation	<a href="https://www.itacyl.es/agro-y-geo-tecnologia/descarga-datos-geograficos/sigpac">https://www.itacyl.es/agro-y-geo-tecnologia/descarga-datos-geograficos/sigpac</a>

and temperature historic averages for wheat-growing months, namely October to June. These parameters are central to plants water status, final yield, and quality performance. Land use datasets were used to circumscribe rain-fed croplands. All datasets were obtained from

publicly available governmental websites (<https://www.atlas.itacyl.es/descarga>).

For wheat suitability mapping, we used fuzzy logic, which has extensively been used for crop suitability mapping in a wide range of agricultural systems (Jamil et al., 2018; Mausio et al., 2020; Nisar Ahamed et al., 2000). Although it entails the application of expert knowledge to define the most important environmental variables, these variables can be backed and based on other empirical research, thus adding robustness, and making such crop suitability mapping accurate. All variables used for wheat suitability mapping were standardized using fuzzy membership functions, which were processed in ArcGIS Pro 3.1.0. Fuzzy membership allows standardizing heterogeneous values on a 0–1 scale. The value indicates the strength of a membership in a set, based on a specified fuzzification algorithm. Different algorithms can be used to standardize the variables depending on the relationship they have to the final output. A value of 1 indicates full membership in the fuzzy set, with membership decreasing to 0, indicating it is not a member of the fuzzy set. For instance, erosion is negatively related to adequate wheat suitability; hence, the smaller the erosion, the more suitable a cropland is, and therefore the value would be closer to 1. For rainfall, temperature, and organic matter, we used fuzzy large algorithms (following Eq. (1), where f1 is the spread and f2 is the midpoint), which related larger values of the variable to full membership in the fuzzy set. In contrast, for erosion, slope, and height, we used fuzzy small algorithms (following Eq. (2), where f1 is the spread and f2 is the midpoint). For clay, lime, and sand% and pH, we used a fuzzy Gaussian algorithm (following Eq. (3), where f1 is the spread and f2 is the midpoint) as the middle values are the most suitable for wheat cultivation. The fuzzy membership values for each variable are illustrated in Fig. 2.

$$\mu(x) = \frac{1}{1 + \left(\frac{x}{f_2}\right)^{-f_1}} \quad (1)$$



**Fig. 2.** Standardized layers with corresponding fuzzy membership values (between 0 and 1) for each variable used in wheat suitability mapping. The values in regard to the suitability of each variable are presented in Table 2 and are based on wheat suitability criteria (Naidu, 2006).

$$\mu(x) = \frac{1}{1 + \left(\frac{x}{f_2}\right)^{f_1}} \quad (2)$$

$$\mu(x) = e^{-f_1 \times (x - f_2)^2} \quad (3)$$

Subsequently, we used the fuzzy overlay function to combine fuzzy membership raster data together. In this sense, we generated three independent fuzzy overlays for the previously described topographic, edaphic, and climatic variables. We used sum as the overlay type to combine each variable to stress that the combined evidence is more important than the single evidence. The four categories correspond to the thresholds described in Table 2, adapted from wheat suitability criteria (Naidu, 2006). All maps were masked for the selected specific land use, namely rainfed croplands. The raster's pixel size was set at 10 m, which is a middle-to-high spatial resolution. Combining the four classes suitability for climatology, edaphology, and topography yielded the final wheat suitability map. To combine the three maps with the corresponding land use mask, we used the weighted overlay function on ArcGIS Pro 3.1.0 with the climatic, edaphic, and topographic rasters. We gave an importance percentage of 40% to both climatic and edaphic variables and 20% to topography Eq. (4) because the output raster from the weighted overlay is an integer the final value is rounded. We gave different weighted values based on the importance of these factors in a Mediterranean setting, with topographic variables having minor effects in contrast with climatic and edaphic factors (Masri and Ryan, 2006; Segarra et al., 2020b).

$$\text{Wheat suitability} = 0.4 \times \text{Climatology} \times 0.4 \times \text{Edaphology} \times 0.2 \times \text{Topography} \quad (4)$$

For the validation of the suitability classification, the resultant suitability map was compared to the FAO (Food and Agriculture Organization of the United Nations) dataset on the potential yield of rainfed wheat grain (Fischer, 2021). This dataset combines several agroclimatic variables, together with reported grain yields, to define agronomic potentialities at a sub-regional scale. The pixels have a resolution of 40 km<sup>2</sup>, and there is a specific dataset defined for wheat rain-fed yield. From this dataset, we extracted 100 random pixels across the studied subregional area, as illustrated in Fig. 3, for validating the obtained wheat suitability mapping results.

To assess the accuracy of the suitability map, we established a threshold corresponding to the FAO dataset on rain-fed wheat grain yield in the region. Marginal wheat cropland encompassed pixels with yields lower than 2500 kg/ha, suitable land between 2500 and 3500 kg/ha, and very suitable land over 3500 kg. A confusion matrix was computed by assessing the 100 random areas on ArcGIS Pro 3.1.0 (Fig. 3). The accuracy assessment obtained from the confusion matrix was used to compute the areas for each the four classes with corresponding confidence intervals. Error margins at 95% probability for each class were derived from the set of error pixels and matched pixels according to the following Eqs. (5) and (6):

$$S(\hat{A}_k) = A \times S(\hat{p}_{\bullet k}) \quad (5)$$

Where  $\hat{A}_k$  is the estimated area of the mapped change class k (in our

case it has four classes), A is the total map area,  $\hat{p}_{\bullet k}$  is the estimated area proportion of change class k derived from the error matrix (as described in Olofsson et al. (2014), page 52), and S is the standard error function of the estimators.  $S(\hat{p}_{\bullet k})$ , shown in Eq. (3), is the standard deviation of the area proportion of the corresponding class.

$$S(\hat{p}_{\bullet k}) = \sqrt{\sum_i W_i^2 \frac{\frac{n_{ik}}{n_i} \left(1 - \frac{n_{ik}}{n_i}\right)}{n_i - 1}} \quad (6)$$

## 2.5. Soil sustainable practices (rotations monitoring)

We used a Sentinel-2-derived crop type classification obtained from (<https://mcsncyl.itacyl.es/descarga>) for the 2020 and 2021 wheat crop seasons. For the studied rainfed cropland area, we extracted the crop type over two consecutive seasons to determine the two-year rotation patterns with regards to sustainable agriculture practices. In this sense, the SIGPAC-delineated field polygons were used, and the majority of pixel classes from within the field were determined as the crop type, assuming a general monoculture cropping approach. Moreover, we have assumed that in the conditions of the region for rainfed crops, only one harvest, namely, one crop, is considered annually. This data was processed in ArcGIS Pro 3.1.0.

## 3. Results

### 3.1. Field-level grain and quality yields

At the field level, we saw that irrigated fields had higher yields (Table 3), which means that they were much more productive than rainfed fields with the same management strategy. We have not observed any differences in yield between chemical and organic (digestate) sources of fertilization. This suggests that organic fertilizers maintain yields compared to conventional management. When it comes to certified organic management in rainfed conditions, wheat grain yield was found to be much lower than in conventionally managed fields, which had about twice as much wheat grain yield. In regard to grain protein content, no clear and significant differences have been observed between management conditions. The grain  $\delta^{13}\text{C}$  was significantly more negative under irrigation compared with rainfed conditions, in agreement with the better water status of irrigated fields. Organic management under rainfed conditions and conventional management with digestate fertilization and rainfed conditions had significantly different  $\delta^{15}\text{N}$  values in contrast with the other managements (Table 3).

### 3.2. Crop rotation practices

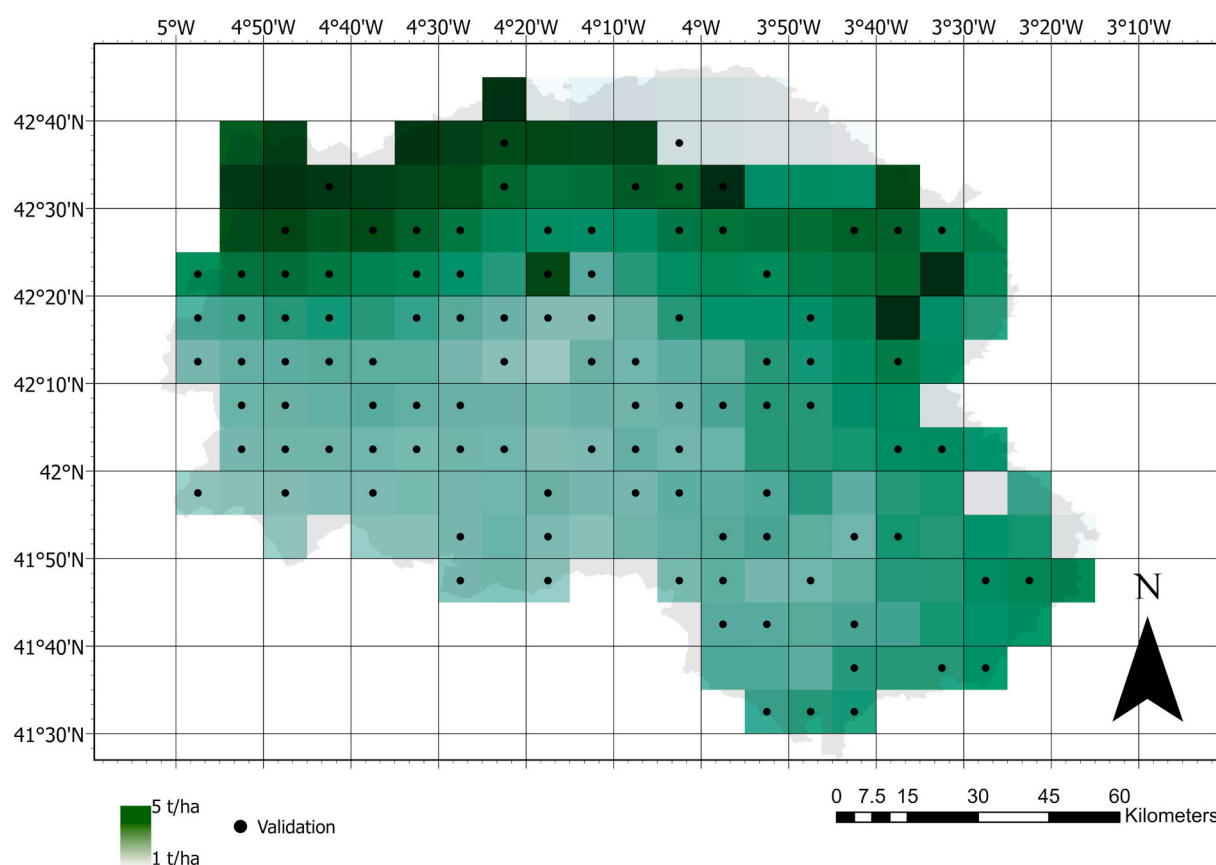
In Table 4, the subsequent crop cultivated after wheat in the following season is indicated. In this case, barley was the most common rotating crop, with 34% of fields growing wheat in the previous season (2020) that were subsequently cultivated with barley in 2021. Secondly, with comparable percentages, 20% of fields cultivated sunflower and 18.6% wheat, while 10% of wheat croplands had fallow fields in the following year, 3.5% were horticultural fields, and finally 2.8% were

**Table 2**

Value of variables in regard to suitability classification adapted from Naidu (2006) and the corresponding references presented in Table 1.

Suitability	Slope (%)	Height (m)	pH	Organic matter (%)	Erosion (Mg/ha)	Texture	Rainfall (mm)	Temperature, °C	Land use
Not suitable	>15	>1300	>10 or <5.5	<1	>150	Sandy	<200	2–8	Other uses
Marginal	10–15	850–1300	8.6–10	1–2	50–150	Clay (45–60%)	200–300	8–10	–
Suitable	5–10	600–850	7.5–8.5 and 5.5–6.4	2–3	10–50	Fine (sandy clay, silty clay, clay) and fine loamy	300–400	10–12	–
Very suitable	<5	110–600	6.5–7.5	>3	0–10	Coarse loam (loamy, silty loam)	>400	>13	Rain-fed croplands





**Fig. 3.** Rainfed wheat production for agricultural subunits published by FAO, the random agricultural areas used for the validation of the wheat suitability maps are indicated with dots.

**Table 3**

Results of the contrasting managements in regard with irrigation and fertilization. GY refers to grain yield, GPC to grain protein content and  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$ , refers to the isotopic composition of carbon and nitrogen of grains.

Fertilization	Irrigation	GY (kg•ha <sup>-1</sup> )	GPC (%)	$\delta^{13}\text{C}$ (‰)	$\delta^{15}\text{N}$ (‰)
Chemical	Irrigated	7347 <sup>a</sup> ± 302	16.2 <sup>a</sup> ± 1.1	-26.8 <sup>a</sup> ± 0.2	3.6 <sup>a</sup> ± 0.2
Chemical	Rainfed	3799 <sup>b</sup> ± 364	12.7 <sup>bc</sup> ± 0.5	-25.3 <sup>b</sup> ± 0.1	3.3 <sup>a</sup> ± 0.3
Digestate	Irrigated	7639 <sup>a</sup> ± 520	11.7 <sup>bc</sup> ± 0.3	-26.5 <sup>a</sup> ± 0.2	3.3 <sup>a</sup> ± 0.3
Digestate	Rainfed	4234 <sup>b</sup> ± 267	14.0 <sup>abc</sup> ± 0.3	-25.4 <sup>b</sup> ± 0.6	6.8 <sup>b</sup> ± 1.4
Organic	Rainfed	2042 <sup>c</sup> ± 133	13.0 <sup>abc</sup> ± 1.3	-24.9 <sup>b</sup> ± 0.1	1.7 <sup>c</sup> ± 0.2

**Table 4**

Percentage of fields doing rotation in the following season after cultivating wheat in the studied area. The crop type classification was extracted from Sentinel-2 images.

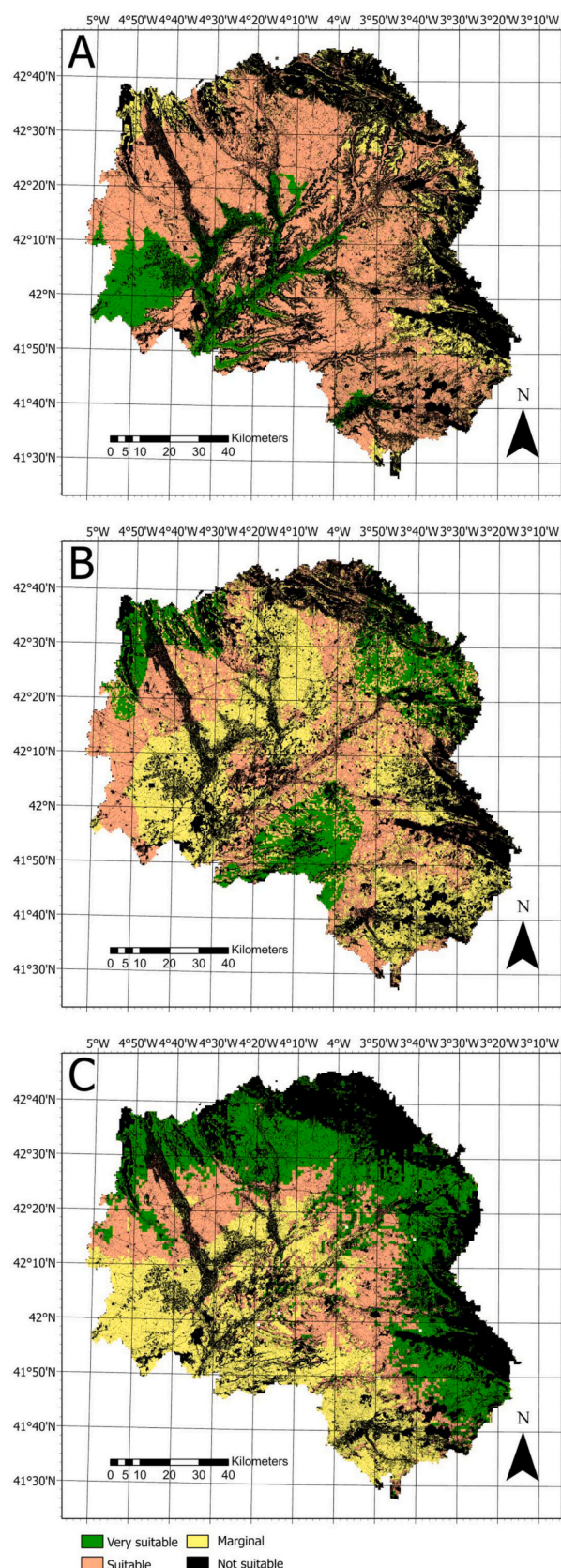
Crop type	% of fields with rotation after wheat
Barley	34.0
Sunflower	20.1
Wheat	18.6
Fallow	10.0
Horticultural crops	3.5
Forage crops	2.9
Other crops	10.8

forage crops in the 2021 season. The results presented in Table 4 indicate a relatively homogenous use of crop rotation in wheat croplands. About 10% of rotated crops included several crops, all of which account for <1% individually, such as pulses, which have been generally recommended as an adequate rotation for enriching soils and controlling pests.

### 3.3. Wheat suitability mapping

In Fig. 4, we present the three maps for each of the criteria reclassified from the fuzzy system into a class system. In this sense, Fig. 5 shows the final map of the three criteria plus the rain-fed cropland mask. As observed visually in Fig. 5 and described in Table 5, the marginal cropland for wheat cultivation occupied 163,503 ± 33,796 ha, which was around 25% of croplands, namely more than the highly productive very suitable croplands, which occupied 122,341 ± 29,014 ha. Meanwhile, the moderately suitable wheat cropland area occupied most of the area, with 357,446 ± 47,640 ha. Significant areas were devoted to wheat cultivation in the marginal croplands; this area represented about a fourth of the total marginal croplands (Table 5). Marginal wheat croplands are located in the southernmost parts of the studied region and are characterized by suboptimal precipitation and advanced soil erosion. In contrast, the majority of suitable croplands are located in regions with higher levels of precipitation and moderate temperatures with relatively adequate soil conditions. The northern sector of the studied area has wheat croplands that are very suited due to their relatively higher elevation, which enables consistent rainfall throughout the growing season, a mild temperature, and more suitable soil conditions than marginal and suitable categories. Non-suitable wheat croplands are characterized by inadequate topographic features as well as inadequate soil conditions, together with insufficient rainfall or too low





**Fig. 4.** Rainfed wheat suitability maps in regard with topographic (A), edaphic (B), and climatic (C) factors calculated following a fuzzy interference system.

temperatures. In regard to the accuracy of the wheat suitability map, as shown in Table 5, the overall accuracy in comparison with FAO classifications was 83%, a relatively high accuracy. This is especially relevant taking into account that the resolution of the obtained map (10 m) was improved, and some information might have been too complex to define at a finer scale.

#### 4. Discussion

Regarding the first research question (what are the field-level grain yield and quality performances for contrasting management strategies in different environments?), we observed that wheat with maximum water availability outperformed other management approaches in terms of grain yield, mostly as a consequence of a better water status. This agrees with the lower (i.e., more negative)  $\delta^{13}\text{C}$  of grains in irrigated compared to rainfed conditions, as has also been corroborated elsewhere (Rezzouk et al., 2020). This is supported by the use of  $\delta^{13}\text{C}$  as an indicator of the total water input (rainfed plus irrigation) received by the crop (Araus et al., 2003). Nonetheless, irrigated croplands devoted to cereal cultivation were limited in the region. Moreover, a general lack of sufficient irrigation happens throughout Spain (Berbel et al., 2019) due to insufficient modernization, a lack of water reservoirs, and extreme seasonal droughts in some areas. Moreover, the water available for irrigation is usually deployed to other, economically more profitable crops such as alfalfa, silage maize, or even sugar beet. Nevertheless, climate change may force us to reverse this scenario concerning wheat cultivation. Thus, increases in temperatures and decreases in precipitation in the region as a consequence of climate change (Calvin et al., 2023) will worsen conditions for rainfed wheat, making irrigation more needed.

A relevant point is that fertilization with organic sources, such as digestate, achieved significantly equivalent yields compared to chemically fertilized wheat crop areas. This is consistent with other reports in wheat (Yan and Gong, 2010; Yang et al., 2020), which did not find significant differences between organic fertilization sources in contrast with chemical fertilizer sources. This indicates a path towards more sustainable practices in the region without compromising yields, as chemical fertilization has been linked to high nutrient losses (nitrogen leaching, nitrous oxide emissions, and ammonia emissions) (Tuomisto et al., 2012) and also results in negative effects on soils, affecting the future of sustainable yields and the environment.

Under rainfed conditions, organic management has yielded about half the yield obtained with rain-fed conventional management approaches. The yield gap observed in this study is larger than that reported by De Ponti et al. (2012), who, after addressing a meta-analysis, observed <20% of the organic/conventional yield gap for wheat. For the studied region, the official yield evaluation for most optimal organic agriculture wheat performance is 3354 kg/ha, while in conventional agriculture it is 5504 kg/ha (<https://genvce.org/productos-genvce/informes/>), indicating an approximate difference of 40%. Other authors have suggested that such yield gaps are dramatically reduced or eventually offset if wheat is grown with low inputs in marginal environments (Dinelli et al., 2013). Another issue is the cultivar chosen. It is common for organic farmers to use landraces or traditional varieties (Ortman et al., 2023), which we also observed in the studied region. These genotypes have lower grain yields, even if the total biomass may be similar or even larger than that of modern varieties. A general lack of crop varieties bred to be specifically adapted to organic conditions and even specific agro-environments (not only management approaches) has been discussed for the European Union (Padel et al., 2021), suggesting that specific breeding for such conditions and approaches is considered pivotal for reducing yield gaps. Several factors affect the reduced production of organic wheat; for instance, the location of the organically managed fields was in the marginal wheat croplands we classified. Meanwhile, the location of conventionally managed fields under chemical and organic (digestate) fertilization, were respectively located in very suitable and suitable wheat cultivation croplands (Fig. 5). Thus,

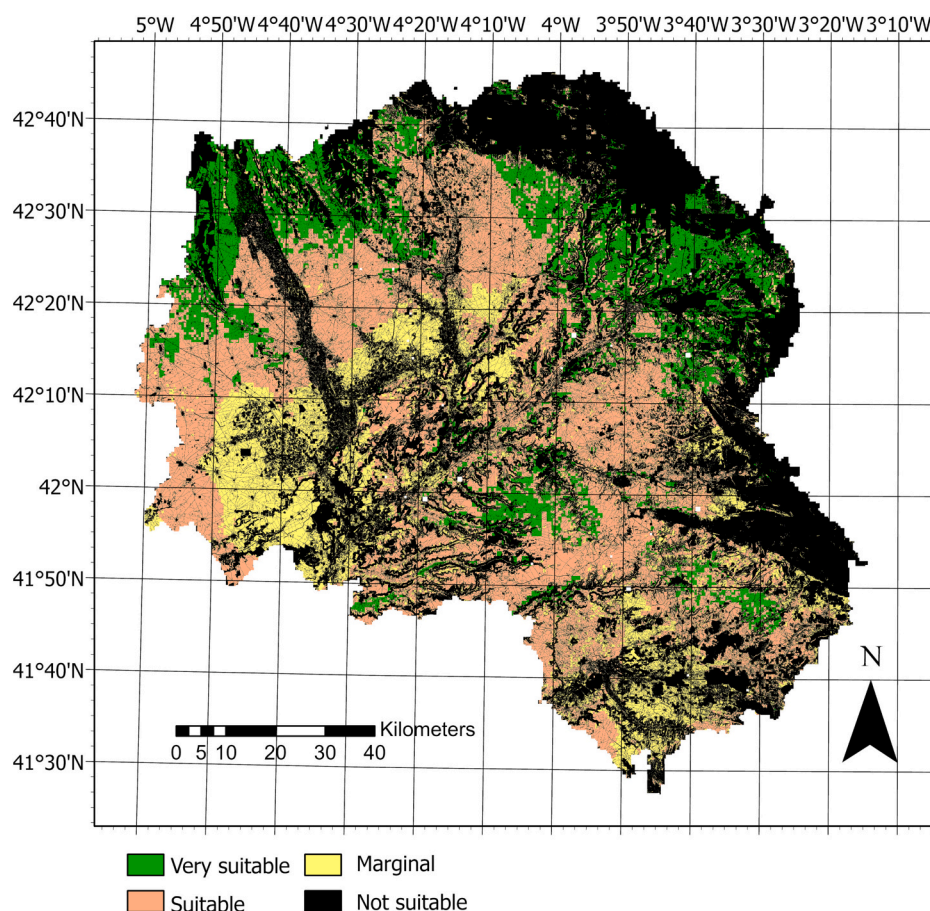


Fig. 5. Rainfed wheat suitability map for the whole region after weighting the suitability factors presented in Fig. 4.

Table 5

Total surface of rain-fed croplands for each suitability class and the current wheat surface within each suitability class with confidence intervals. The estimated wheat suitability area and confidence intervals calculated following Olofsson et al., (2014) according to the map shown in Fig. 5 and the accuracy assessment obtained with the confusion matrix shown in Table 6 are also presented.

Suitability	Wheat surface (ha)	Total surface (ha)
Marginal	44,815	163,503 ± 33,796
Suitable	114,218	357,446 ± 47,640
Very suitable	48,811	122,341 ± 29,014

Table 6

Confusion matrix for the wheat suitability shown in Fig. 5 validated with FAO rainfed wheat production raster for the study region. User's accuracy (UA), producer's accuracy (PA), and overall accuracy (OA) are shown with the 100 delimited agricultural areas used for the accuracy assessment.

	Marginal	Suitable	Very suitable	UA [%]
Marginal	22	2	3	81.5
Suitable	4	42	3	85.7
Very suitable	1	4	19	79.2
PA [%]	81.5	87.5	76.0	OA = 83.0

besides management, the genotype and the environment in which it is grown significantly affect the agricultural output. This might have played a role in increasing the yield gap in our study in comparison with other studies and the official statistics. In regard to quality traits, no clear differences were observed for the case of grain protein content, which suggests management (water regime and fertilization) together

with cultivar choice did not affect this trait, except for the irrigated and chemically fertilized trials. In this last case, the higher protein values may be the consequence of intensive nitrogen fertilization associated with the highly productive conditions of an irrigated crop (Ibrahim et al., 2018). Meanwhile,  $\delta^{15}\text{N}$  isotopes indicated the probable precedence of pulses (forage legumes) in the rotation (Peoples et al., 2015) for the organically managed rainfed fields, while the rainfed fields fertilized with digestate exhibited higher inputs from organic-derived nitrogen.

In regard to the second research question (are sustainable crop rotation practices being carried out in the region?) we observed a poor overall crop rotation in the region. Most fields growing wheat cultivate subsequently either wheat again (with 18.5% of cases) or another small grain cereal such as barley, with 34% of cases. This indicates that over 50% of fields cultivate the same or a relatively equivalent crop throughout consecutive seasons. This suggests a high dependence on pesticides and herbicides, as in continuous cropping scenarios, the emergence of pests and other limiting factors for crops is common, especially in high input systems (Jalli et al., 2021). But also, crop rotations have been shown to improve wheat performance, soils, and the environment in Mediterranean areas (López-Bellido et al., 2020). In this sense, multiple ecosystem services can benefit from crop rotations to maintain soil structure and microorganisms, improve water and nutrient use efficiency, and regulate pests in organic management (Barbieri et al., 2017).

The European-launched satellite constellation Sentinel-2 is a good tool for monitoring agricultural activity (Segarra et al., 2020a) and mapping crop types for assessing adequate agricultural policies (Defourny et al., 2019). These technological advances can guide policies to promote sustainable agriculture practices such as rotations and monitor their effectiveness from the field to the landscape level. Besides



rotations, conservation agriculture (i.e., minimum or no tillage) is being widely adopted in the region, which may contribute to the sustainability of marginal lands. Certain agricultural management practices have long-term impacts and considerable consequences for environmental integrity (Rounsevell et al., 2003). These consequences include changes in nutrient dynamics affecting soils (Cameron et al., 2013), the use of water resources (Rosegrant et al., 2009) (Rosegrant et al., 2009), and biological diversity (Lark et al., 2020). The environmental repercussions endanger crop performances while negatively impacting ecosystems and human well-being (Willett et al., 2019). Therefore, changing management has a wider impact on the ecosystem and its socioeconomic effects as territories are shaped in a certain manner. The current satellite data infrastructure allows for monitoring these aspects and potentially improving sustainable practices.

Regarding the third research question (what is the spatial suitability of wheat cultivation in the region's croplands?) we have determined through a fuzzy interference system approach that in the studied region in central-north Spain,  $163,503 \pm 33,796$  ha correspond to marginal croplands,  $357,446 \pm 47,640$  ha to suitable, and  $122,341 \pm 29,014$  ha to very suitable wheat cultivation croplands. The modeling approach has been confirmed as an adequate crop suitability method, as the validation we addressed shows (Table 5). This fuzzy interference modeling has also been successfully used for wheat suitability mapping in Mediterranean and semi-arid conditions (Dedeoğlu and Dengiz, 2019; Mendas and Delali, 2012; Pilevar et al., 2020). We confirm the appropriateness of this approach for wheat cultivation suitability mapping in the studied region. To our knowledge, no other study has addressed the suitability of wheat cultivation in the studied region. Therefore, the study presents a novelty by spatially defining wheat suitability using publicly available datasets, which can be relevant in circumscribing adequate management and breeding approaches for wheat cultivation in target environments in Castile and Leon (Spain).

The use of geolocate datasets affecting crop growth and development allows defining spatially envirotypes that can guide specific local breeding or management approaches. This allows us to overcome focusing only on a few experimental sites and conditions. In that respect, other studies have addressed, to some extent, this issue and have used satellite imagery to extrapolate experimental field data into a regional and landscape level for the cases of grain yield (Lambert et al., 2018; Segarra et al., 2020b) and quality traits (Segarra et al., 2022; Zhao et al., 2019), combining several remote sensing instruments and physiological measurements. The current data flow allows defining envirotypes and addressing adequate breeding and management approaches in marginal croplands that occupy significant areas, such as in the case of central-north Spain described here.

In that respect, the European Union plans to achieve 25% of total European cropland in organic production by the next decade (2030), as stated in the European Commission statement titled "Action Plan for the Development of Organic Production in the EU" (2021). Therefore, finding marginal croplands can help in defining areas with reduced yield gaps when comparing organic with conventional management, as discussed in the first research question, while leaving the most suitable lands to more productive approaches. The most optimal agricultural areas are generally found in intensively managed croplands with a localized concentration of the most advanced technological improvements, while marginal croplands are often found in environments with agroecological conditions that limit their productivity (Ceccarelli, 1994). In marginal areas, soil erosion, rainfall, and temperature do not meet the optimal conditions for crop development, resulting in crop failure or lower yields (Zampieri et al., 2020). Since the implementation of industrial agriculture, such croplands have been progressively abandoned in several Mediterranean regions (Jiménez-Olivencia et al., 2021; Segarra et al., 2023), as observed in the cropland use changes dataset elaborated by Winkler et al. (2021). Hence, mapping wheat suitability, monitoring soil sustainability practices, and evaluating field-level wheat performance with contrasting management approaches can contribute

to managing croplands and applying adequate farming approaches to specific target environments. Furthermore, this approach can be applied to other Mediterranean regions and parts of the world, adapting the methodology to the available spatially relevant environmental datasets. This approach could play a role in avoiding mismatches between production supply and demand. Yet, some limitations need to be addressed in regard to the integration of several datasets and sensors for an accurate assessment of wheat agroecological status in the region and elsewhere.

## 5. Conclusion

We conclude that there is a significant yield gap, about 50%, between organic and conventional wheat production in Castile and Leon (Spain). However, the genotypes generally used in the region for organic farming and the marginal environment where organically managed fields were located also played a significant role in the final agricultural output. In that respect, marginal croplands occupy  $163,503 \pm 33,796$  ha in central-north Spain, representing 25% of the total rainfed cropland. We suggest that this area could be organically farmed, with the use of wheat cultivars particularly adapted to the drought and low fertile conditions of these marginal lands. This could contribute to achieving up to a 25% of the total cropland in organic production by the next decade (2030) in the studied region, while leaving the most suitable croplands (suitable:  $357,446 \pm 47,640$  ha, and very suitable:  $122,341 \pm 29,014$  ha) to more productive management approaches. We also found that current sustainable production practices, such as crop rotation, are scarce and do not meet the optimal conditions to ensure long-term soil sustainability, with over 50% of fields cultivating either wheat or barley after wheat. Hence, geospatial datasets from fully accessible sources collected by public institutions or open-access satellites constitute a pivotal framework for envirotyping croplands and establishing adequate wheat cultivation approaches, such as organic or other more productive sustainable management. A fuzzy interference system has proven to be an adequate modeling approach to decipher wheat suitability classes and standardize different environmental data. The use of suitability modeling with fine agricultural and environmental data allows envirotyping and potentially guides sustainable management and research initiatives in the studied region. The methodology applied could be adapted to other wheat-growing regions.

## Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Joel Segarra reports was provided by University of Barcelona.

## Data availability

All data used is publicly available and the websites have been added in the manuscript, the farmers field-level data can be provided on request

## Acknowledgements

We acknowledge the support of the project PID2022-138307OB-C21 from the Ministerio de Ciencia e Innovación, Spain. J.S. is a recipient of a FPI (formación de personal investigador) doctoral fellowship from the Ministerio de Ciencia e Innovación, Spain (grant: PRE2020-091907). J. L.A. acknowledges support from the Institut Català de Recerca i Estudis Avançats (ICREA), Generalitat de Catalunya, Spain. S.C.K. is supported by the Ramon y Cajal RYC-2019-027818-I research fellowship from the Ministerio de Ciencia e Innovación, Spain. We acknowledge all farmers that shared the information of their wheat production with us.

## References

- Annicchiarico, P., Bellah, F., Chiaric, T., 2005. Defining subregions and estimating benefits for a specific-adaptation strategy by breeding programs. *Crop* 45, 1741–1749.
- Araus, J.L., Villegas, D., Aparicio, N., del Moral, L.F.G., El Hani, S., Rharrabti, Y., Ferrio, J.P., Royo, C., 2003. Environmental factors determining carbon isotope Discrimination and yield in durum wheat under Mediterranean conditions. *Crop Sci.* 43, 170. <https://doi.org/10.2135/cropsci2003.1700>.
- Araus, J.L., Cabrera-Bosquet, L., Serret, M.D., Bort, J., Nieto-Taladriz, M.T., 2013. Comparative performance of  $\delta^{13}C$ ,  $\delta^{18}O$  and  $\delta^{15}N$  for phenotyping durum wheat adaptation to a dryland environment. *Funct. Plant Biol.* 40, 595. <https://doi.org/10.1071/FP12254>.
- Araus, J.L., Kefauver, S.C., Díaz, O.V., Gracia-Romero, A., Rezzouk, F.Z., Segarra, J., Buchailot, M.L., Chang-Espino, M., Vatter, T., Sanchez-Bragado, R., Gallego, J.A.F., Serret, M.D., Bort, J., 2021. Crop phenotyping in a context of global change: what to measure and how to do it. *J. Integr. Plant Biol.* <https://doi.org/10.1111/jipb.13191>.
- Barbieri, P., Pellerin, S., Nesme, T., 2017. Comparing crop rotations between organic and conventional farming. *Sci. Rep.* 7, 13761. <https://doi.org/10.1038/s41598-017-14271-6>.
- Barmer, G., Schmidhalter, U., 2016. High-throughput phenotyping of wheat and barley plants grown in single or few rows in small plots using active and passive spectral proximal sensing. *Sensors (Switzerland)* 16, 1–14. <https://doi.org/10.3390/s16111860>.
- Basso, B., Fiorentino, C., Cammarano, D., Cafiero, G., Dardanelli, J., 2012. Analysis of rainfall distribution on spatial and temporal patterns of wheat yield in Mediterranean environment. *Eur. J. Agron.* 41, 52–65. <https://doi.org/10.1016/j.eja.2012.03.007>.
- Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data* 5, 180214. <https://doi.org/10.1038/sdata.2018.214>.
- Bento, V.A., Ribeiro, A.F.S., Russo, A., Gouveia, C.M., Cardoso, R.M., Soares, P.M.M., 2021. The impact of climate change in wheat and barley yields in the Iberian Peninsula. *Sci. Rep.* 11, 1–12. <https://doi.org/10.1038/s41598-021-95014-6>.
- Berbel, J., Expósito, A., Gutiérrez-Martín, C., Mateos, L., 2019. Effects of the irrigation modernization in Spain 2002–2015. *Water Resour. Manag.* 33, 1835–1849. <https://doi.org/10.1007/s11269-019-02215-w>.
- Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P.W., Trisos, C., Romero, J., Aldunce, P., Barrett, K., Blanco, G., Cheung, W.W.L., Connors, S., Denton, F., Dongue-Niang, A., Dodman, D., Garschagen, M., Geden, O., Hayward, B., Jones, C., Jotzo, F., Krug, T., Lasco, R., Lee, Y.-Y., Masson-Delmotte, V., Meinshausen, M., Mintenbeck, K., Mokssit, A., Otto, F.E.L., Pathak, M., Pirani, A., Poloczanska, E., Portner, H.-O., Revi, A., Roberts, D.C., Roy, J., Ruane, A.C., Skea, J., Shukla, P.R., Slade, R., Slangen, A., Sokona, Y., Sörensen, A.A., Tignor, M., van Vuuren, D., Wei, Y.-M., Winkler, H., Zhai, P., Zommers, Z., Hourcade, J.-C., Johnson, F.X., Pachauri, S., Simpson, N.P., Singh, C., Thomas, A., Totin, E., Alegría, A., Armour, K., Bednar-Friedl, B., Blok, K., Cissé, G., Dentener, F., Eriksen, S., Fischer, E., Garner, G., Guivarch, C., Haasnoot, M., Hansen, G., Hauser, M., Hawkins, E., Hermans, T., Kopp, R., Leprince-Ringuet, N., Lewis, J., Ley, D., Ludden, C., Namiar, L., Nicholls, Z., Some, S., Szopa, S., Trewin, B., van der Wijst, K.-I., Winter, G., Witting, M., Birt, A., Ha, M., 2023. In: Core Writing Team, Lee, H., Romero, J. (Eds.), IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland. <https://doi.org/10.59327/IPCC/AR6-9789291691647>.
- Cameron, K.C., Di, H.J., Moir, J.L., 2013. Nitrogen losses from the soil/plant system: a review. *Ann. Appl. Biol.* 162, 145–173. <https://doi.org/10.1111/aab.12014>.
- Ceccarelli, S., 1994. Specific adaptation and breeding for marginal conditions. *Breed. Fodder Crops Margin. Cond.* 101–127. [https://doi.org/10.1007/978-94-011-0966-6\\_15](https://doi.org/10.1007/978-94-011-0966-6_15).
- De Ponti, T., Rijk, B., Van Ittersum, M.K., 2012. The crop yield gap between organic and conventional agriculture. *Agric. Syst.* 108, 1–9. <https://doi.org/10.1016/j.agsy.2011.12.004>.
- Dedeoğlu, M., Dengiz, O., 2019. Generating of land suitability index for wheat with hybrid system approach using AHP and GIS. *Comput. Electron. Agric.* 167, 105062. <https://doi.org/10.1016/j.compag.2019.105062>.
- Defourny, P., Bontemps, S., Bellemans, N., Cara, C., Dedieu, G., Guzzonato, E., Hagolle, O., Inglada, J., Nicola, L., Rabatte, T., Savinaud, M., Udrou, C., Valero, S., Bégué, A., Dejoux, J.F., El Harti, A., Ezzahar, J., Kussul, N., Labbassi, K., Lebourgeois, V., Miao, Z., Newby, T., Nyamugama, A., Salh, N., Shelestov, A., Simonneau, V., Traore, P.S., Traore, S.S., Koetz, B., 2019. Near real-time agriculture monitoring at national scale at parcel resolution: performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sens. Environ.* 221, 551–568. <https://doi.org/10.1016/j.rse.2018.11.007>.
- Dinelli, G., Marotti, I., Di Silvestro, R., Bosi, S., Bregola, V., Accorsi, M., Di Loreto, A., Benedettelli, S., Ghiselli, L., Catizone, P., 2013. Agronomic, nutritional and nutraceutical aspects of durum wheat (*Triticum durum* Desf.) cultivars under low input agricultural management. *Ital. J. Agron.* 8, 12. <https://doi.org/10.4081/ija.2013.e12>.
- Farquhar, G.D., Ehleringer, J.R., Hubick, K.T., 1989. CARBON ISOTOPE DISCRIMINATION AND PHOTOSYNTHESIS. *Annu. Rev. Plant Biol.* 40, 503–537.
- Fischer, G., 2021. *Glob. Agro Ecol. Zones Vers. 4 (GAEZ v4)*.
- García Fernández, J., 1986. El clima en Castilla y León. *Ambito*.
- Green, T.R., Salas, J.D., Martinez, A., Erskine, R.H., 2007. Relating crop yield to topographic attributes using spatial analysis neural networks and regression. *Geoderma* 139, 23–37. <https://doi.org/10.1016/j.geoderma.2006.12.004>.
- He, Y., Hou, L., Wang, H., Hu, K., McConkey, B., 2014. A modelling approach to evaluate the long-term effect of soil texture on spring wheat productivity under a rain-fed condition. *Sci. Rep.* 4, 5736. <https://doi.org/10.1038/srep05736>.
- Ibrahim, A., Csűr-Varga, A., Jolánkai, M., Mansour, H., Hamed, A., 2018. Monitoring some quality attributes of different wheat varieties by infrared technology. *Agric. Eng. Int. CIGR J.* 20, 201–210.
- Jalli, M., Huusela, E., Jalli, H., Kauppi, K., Niemi, M., Himanen, S., Jauhainen, L., 2021. Effects of crop rotation on spring wheat yield and Pest occurrence in different tillage systems: a multi-year experiment in Finnish growing conditions. *Front. Sustain. Food Syst.* 5. <https://doi.org/10.3389/fsufs.2021.647335>.
- Jamil, M., Sahana, M., Sajjad, H., 2018. Crop suitability analysis in the Bijnor District, UP, using geospatial tools and fuzzy analytical hierarchy process. *Agric. Res.* 7, 506–522. <https://doi.org/10.1007/s40003-018-0335-5>.
- Jiménez-Olivencia, Y., Ibáñez-Jiménez, Á., Porcel-Rodríguez, L., Zimmerer, K., 2021. Land use change dynamics in euro-mediterranean mountain regions: driving forces and consequences for the landscape. *Land Use Policy* 109. <https://doi.org/10.1016/j.landusepol.2021.105721>.
- Kumhálová, J., Moudrý, V., 2014. Topographical characteristics for precision agriculture in conditions of the Czech Republic. *Appl. Geogr.* 50, 90–98. <https://doi.org/10.1016/j.apgeog.2014.02.012>.
- Lambert, M.J., Traoré, P.C.S., Blaes, X., Baret, P., Defourny, P., 2018. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens. Environ.* 216, 647–657. <https://doi.org/10.1016/j.rse.2018.06.036>.
- Lark, T.J., Spawn, S.A., Bougie, M., Gibbs, H.K., 2020. Cropland expansion in the United States produces marginal yields at high costs to wildlife. *Nat. Commun.* 11, 4295. <https://doi.org/10.1038/s41467-020-18045-z>.
- Littleboy, M., Freebairn, D., Hammer, G., Silburn, D., 1992. Impact of soil erosion on production in cropping systems. II. Simulation of production and erosion risks for a wheat cropping system. *Soil Res.* 30, 775. <https://doi.org/10.1071/SR9920775>.
- López-Bellido, L., López-Bellido, R., Fernández-García, P., Muñoz-Romero, V., López-Bellido, F.J., 2020. Carbon storage in a rainfed Mediterranean vertisol: effects of tillage and crop rotation in a long-term experiment. *Eur. J. Soil Sci.* 71, 472–483. <https://doi.org/10.1111/ejss.12883>.
- Masri, Z., Ryan, J., 2006. Soil organic matter and related physical properties in a Mediterranean wheat-based rotation trial. *Soil Tillage Res.* 87, 146–154. <https://doi.org/10.1016/j.still.2005.03.003>.
- Mausio, K., Miura, T., Lincoln, N.K., 2020. Cultivation potential projections of breadfruit (*Artocarpus altilis*) under climate change scenarios using an empirically validated suitability model calibrated in Hawai'i. *PLOS One* 15, e0228552. <https://doi.org/10.1371/journal.pone.0228552>.
- Mendas, A., Delai, A., 2012. Integration of MultiCriteria decision analysis in GIS to develop land suitability for agriculture: application to durum wheat cultivation in the region of Mleta in Algeria. *Comput. Electron. Agric.* 83, 117–126. <https://doi.org/10.1016/j.compag.2012.02.003>.
- Naidu, L.G.K., 2006. *Soil-site Suitability Criteria for Major Crops*.
- Neina, D., 2019. The role of soil pH in plant nutrition and soil remediation. *Appl. Environ. Soil Sci.* 2019, 1–9. <https://doi.org/10.1155/2019/5794869>.
- Nisar Ahmed, T.R., Gopal Rao, K., Murthy, J.S.R., 2000. GIS-based fuzzy membership model for crop-land suitability analysis. *Agric. Syst.* 63, 75–95. [https://doi.org/10.1016/S0308-521X\(99\)00036-0](https://doi.org/10.1016/S0308-521X(99)00036-0).
- Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* 148, 42–57. <https://doi.org/10.1016/j.rse.2014.02.015>.
- Ortman, T., Sandström, E., Bengtsson, J., Watson, C.A., Bergkvist, G., 2023. Farmers' motivations for landrace cereal cultivation in Sweden. *Biol. Agric. Hortic.* 1–22. <https://doi.org/10.1080/01448765.2023.2207081>.
- Padel, S., Orsini, S., Solfanelli, F., Zanoli, R., 2021. Can the market deliver 100% organic seed and varieties in Europe? *Sustain.* 13, 1–13. <https://doi.org/10.3390/su131810305>.
- Peña, R.J., 2002. *Wheat for Bread and Other Foods. Bread wheat improvement and production*, Rome.
- Peoples, M.B., Chalk, P.M., Unkovich, M.J., Boddey, R.M., 2015. Can differences in  $^{15}N$  natural abundance be used to quantify the transfer of nitrogen from legumes to neighbouring non-legume plant species? *Soil Biol. Biochem.* 87, 97–109. <https://doi.org/10.1016/j.soilbio.2015.04.010>.
- Pilevar, A.R., Matinfar, H.R., Sohrabi, A., Sarmadian, F., 2020. Integrated fuzzy, AHP and GIS techniques for land suitability assessment in semi-arid regions for wheat and maize farming. *Ecol. Indic.* 110, 105887. <https://doi.org/10.1016/j.ecolind.2019.105887>.
- Porter, J.R., Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. *Eur. J. Agron.* 10, 23–36. [https://doi.org/10.1016/S1161-0301\(98\)00047-1](https://doi.org/10.1016/S1161-0301(98)00047-1).
- R Core Team, 2021. *A language and environment for statistical computing*.
- Raven, P.H., Wagner, D.L., 2021. Agricultural intensification and climate change are rapidly decreasing insect biodiversity. *Proc. Natl. Acad. Sci.* 118. <https://doi.org/10.1073/pnas.2002548117>.
- Resende, R.T., Piepho, H.-P., Rosa, G.J.M., Silva-Junior, O.B., Silva, F.F., de Resende, M. D.V., Grattapaglia, D., 2021. Enviromics in breeding: applications and perspectives on envirotypic-assisted selection. *Theor. Appl. Genet.* 134, 95–112. <https://doi.org/10.1007/s00122-020-03684-z>.
- Rezzouk, F.Z., Gracia-Romero, A., Kefauver, S.C., Gutiérrez, N.A., Aranjuelo, I., Serret, M.D., Araus, J.L., 2020. Remote sensing techniques and stable isotopes as phenotyping tools to assess wheat yield performance: effects of growing temperature and vernalization. *Plant Sci.* 295, 110281. <https://doi.org/10.1016/j.plantsci.2019.110281>.

- Rosegrant, M.W., Ringler, C., Zhu, T., 2009. Water for agriculture: maintaining food security under growing scarcity. *Annu. Rev. Environ. Resour.* 34, 205–222. <https://doi.org/10.1146/annurev.enviro.030308.090351>.
- Rounsevell, M.D.A., Annetts, J.E., Audsley, E., Mayr, T., Reginster, I., 2003. Modelling the spatial distribution of agricultural land use at the regional scale. *Agric. Ecosyst. Environ.* 95, 465–479. [https://doi.org/10.1016/S0167-8809\(02\)00217-7](https://doi.org/10.1016/S0167-8809(02)00217-7).
- Segarra, J., Buchailot, M.L., Araus, J.L., Kefauver, S.C., 2020a. Remote sensing for precision agriculture : Sentinel-2 improved features and applications. *Agronomy* 1–18.
- Segarra, J., González-Torralba, J., Aranjuelo, Í., Araus, J.L., Kefauver, S.C., 2020b. Estimating wheat grain yield using Sentinel-2 imagery and exploring topographic features and rainfall effects on wheat performance in Navarre, Spain. *Remote Sens.* 12, 1–24. <https://doi.org/10.3390/rs12142278>.
- Segarra, J., Rezzouk, F.Z., Aparicio, N., González-Torralba, J., Aranjuelo, I., Gracia-Romero, A., Araus, J.L., Kefauver, S.C., 2022. Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content. *Inf. Process. Agric.* <https://doi.org/10.1016/j.inpa.2022.05.004>.
- Segarra, J., Fernández-Martínez, J., Araus, J.L., 2023. Managing abandoned Mediterranean mountain landscapes: the effects of donkey grazing on biomass control and floral diversity in pastures. *Catena* 233. <https://doi.org/10.1016/j.catena.2023.107503>.
- Serret, M., Ortiz-Monasterio, I., Pardo, A., Araus, J., 2008. The effects of urea fertilisation and genotype on yield, nitrogen use efficiency,  $\delta^{15}\text{N}$  and  $\delta^{13}\text{C}$  in wheat. *Ann. Appl. Biol.* <https://doi.org/10.1111/j.1744-7348.2008.00259.x>, 080617165316730-???
- Tubiello, F.N., Salvatore, M., Ferrara, A.F., House, J., Federici, S., Rossi, S., Biancalani, R., Condor Golec, R.D., Jacobs, H., Flammini, A., Prosperi, P., Cardenas-Galindo, P., Schmidhuber, J., Sanz Sanchez, M.J., Srivastava, N., Smith, P., 2015. The contribution of agriculture, forestry and other land use activities to global warming, 1990–2012. *Glob. Chang. Biol.* 21, 2655–2660. <https://doi.org/10.1111/gcb.12865>.
- Tuel, A., Eltahir, E.A.B., 2020. Why is the Mediterranean a climate change hot spot? *J. Clim.* 33, 5829–5843. <https://doi.org/10.1175/JCLI-D-19-0910.1>.
- Tuomisto, H.L., Hodge, I.D., Riordan, P., Macdonald, D.W., 2012. Does organic farming reduce environmental impacts? - a meta-analysis of European research. *J. Environ. Manag.* 112, 309–320. <https://doi.org/10.1016/j.jenvman.2012.08.018>.
- Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., Jonell, M., Clark, M., Gordon, L.J., Fanzo, J., Hawkes, C., Zurayk, R., Rivera, J.A., De Vries, W., Majele Sibanda, L., Afshin, A., Chaudhary, A., Herrero, M., Agustina, R., Branca, F., Lartey, A., Fan, S., Crona, B., Fox, E., Bignet, V., Troell, M., Lindahl, T., Singh, S., Cornell, S.E., Srinath Reddy, K., Narain, S., Nishtar, S., Murray, C.J.L., 2019. Food in the Anthropocene: the EAT–lancet commission on healthy diets from sustainable food systems. *Lancet* 393, 447–492. [https://doi.org/10.1016/S0140-6736\(18\)31788-4](https://doi.org/10.1016/S0140-6736(18)31788-4).
- Winkler, K., Fuchs, R., Rounsevell, M., Herold, M., 2021. Global land use changes are four times greater than previously estimated. *Nat. Commun.* 12, 1–10. <https://doi.org/10.1038/s41467-021-22702-2>.
- Xu, Y., 2016. Envirotyping for deciphering environmental impacts on crop plants. *Theor. Appl. Genet.* 129, 653–673. <https://doi.org/10.1007/s00122-016-2691-5>.
- Yan, X., Gong, W., 2010. The role of chemical and organic fertilizers on yield, yield variability and carbon sequestration— results of a 19-year experiment. *Plant Soil* 331, 471–480. <https://doi.org/10.1007/s11104-009-0268-7>.
- Yang, Y.J., Lei, T., Du, W., Liang, C.L., Li, H.D., Lv, J.L., 2020. Substituting chemical fertilizer nitrogen with organic manure and comparing their nitrogen use efficiency and winter wheat yield. *J. Agric. Sci.* 158, 262–268. <https://doi.org/10.1017/S0021859620000544>.
- Zampieri, M., Toreti, A., Ceglar, A., Naumann, G., Turco, M., Tebaldi, C., 2020. Climate resilience of the top ten wheat producers in the Mediterranean and the Middle East. *Reg. Environ. Chang.* 20 <https://doi.org/10.1007/s10113-020-01622-9>.
- Zhang, H., Oweis, T., 1999. Water–yield relations and optimal irrigation scheduling of wheat in the Mediterranean region. *Agric. Water Manag.* 38, 195–211. [https://doi.org/10.1016/S0378-3774\(98\)00069-9](https://doi.org/10.1016/S0378-3774(98)00069-9).
- Zhao, H., Song, X., Yang, G., Li, Z., Zhang, D., Feng, H., 2019. Monitoring of nitrogen and grain protein content in winter wheat based on sentinel-2A data. *Remote Sens.* 11, 1–25. <https://doi.org/10.3390/rs11141724>.



### **Capítol 3.** Estimació de la producció de blat amb imatges de Sentinel-2 i avaluació dels efectes topogràfics i de distribució de pluja en el rendiment a Navarra

#### **Chapter 3.** Estimating Wheat Grain Yield Using Sentinel-2 Imagery and Exploring Topographic Features and Rainfall Effects on Wheat Performance in Navarre

Joel Segarra <sup>1,2</sup>, Jon González-Torralba <sup>3</sup>, Íker Aranjuelo <sup>4</sup>, Jose Luis Araus <sup>1,2</sup> and Shawn C. Kefauver <sup>1,2,\*</sup>

<sup>1</sup>Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain; joel.segarra@ub.edu (J.S.); jaraus@ub.edu (J.L.A.)

<sup>2</sup>AGROTECNIO (Center for Research in Agrotechnology), 25198 Lleida, Spain

<sup>3</sup>Grupo AN, Campo Tajonar, 31192 Tajonar, Spain; jon.gonzalez@grupoan.com




<sup>4</sup>Instituto de Agrobiotecnología (IdAB), CSIC-Gobierno de Navarra, 31192 Pamplona, Spain; iker.aranjuelo@csic.es

\*Correspondence: sckefauver@ub.edu

Publicat a la revista Remote Sensing

## Article

# Estimating Wheat Grain Yield Using Sentinel-2 Imagery and Exploring Topographic Features and Rainfall Effects on Wheat Performance in Navarre, Spain

Joel Segarra <sup>1,2</sup>, Jon González-Torralba <sup>3</sup>, Íker Aranjuelo <sup>4</sup> , Jose Luis Araus <sup>1,2</sup>  and Shawn C. Kefauver <sup>1,2,\*</sup> 

<sup>1</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain; joel.segarra@ub.edu (J.S.); jaraus@ub.edu (J.L.A.)

<sup>2</sup> AGROTECNIO (Center for Research in Agrotechnology), 25198 Lleida, Spain

<sup>3</sup> Grupo AN, Campo Tajonar, 31192 Tajonar, Spain; jon.gonzalez@grupoan.com

<sup>4</sup> Instituto de Agrobiotecnología (IdAB), CSIC-Gobierno de Navarra, 31192 Pamplona, Spain; iker.aranjuelo@csic.es

\* Correspondence: sckefauver@ub.edu

Received: 4 June 2020; Accepted: 13 July 2020; Published: 15 July 2020



**Abstract:** Reliable methods for estimating wheat grain yield before harvest could help improve farm management and, if applied on a regional level, also help identify spatial factors that influence yield. Regional grain yield can be estimated using conventional methods, but the typical process is complex and labor-intensive. Here we describe the development of a streamlined approach using publicly accessible agricultural data, field-level yield, and remote sensing data from Sentinel-2 satellite to estimate regional wheat grain yield. We validated our method on wheat croplands in Navarre in northern Spain, which features heterogeneous topography and rainfall. First, this study developed stepwise multilinear equations to estimate grain yield based on various vegetation indices, which were measured at various phenological stages in order to determine the optimal timings. Second, the most suitable model was used to estimate grain yield in wheat parcels mapped from Sentinel-2 satellite images. We used a supervised pixel-based random forest classification and the estimates were compared to government-published post-harvest yield statistics. When tested, the model achieved an  $R^2$  of 0.83 in predicting grain yield at field level. The wheat parcels were mapped with an accuracy close to 86% for both overall accuracy and compared to official statistics. Third, the validated model was used to explore potential relationships of the mapped per-parcel grain yield estimation with topographic features and rainfall by using geographically weighted regressions. Topographic features and rainfall together accounted for an average for 11 to 20% of the observed spatial variation in grain yield in Navarre. These results highlight the ability of our method for estimating wheat grain yield before harvest and determining spatial factors that influence yield at the regional scale.

**Keywords:** remote sensing; agriculture; crop monitoring; Sentinel-2; wheat

## 1. Introduction

Bread wheat (*Triticum aestivum* L.) is an important grain in Spain, representing 3% of the total of European production [1]. The majority of wheat production is located in central and southern Spain, which corresponds to the Castille and Andalusia regions. Navarre's wheat croplands in northern Spain, where this paper focuses, also have significant production. Furthermore, Navarre hosts several Spanish agroclimates and a wide range of water regimes in a relatively compact spatial area due to



its topographical heterogeneity, which can serve as a good proxy for helping to understand wheat performance across all Spanish agrosystems. Estimating wheat grain yield, mapping wheat crop lands, and studying spatial interactions that affect grain yields can be of great help for farmers and agricultural institutions. This is especially relevant in the European Union, as the Common Agricultural Policy (CAP) subsidies requirements are linked to cropland uses and member states have to verify the declared field area, the sowed crop type and harvest [2].

Regarding wheat grain yield, there are two main ways to estimate it that take advantage of remote sensing data; on the one hand empirical models based on actual field data and spectral-based vegetation indices, which have been successfully used with Sentinel-2 data and wheat grain yield estimation [3,4], and on the other hand growth models [5] involving physiological parameter estimations based on, for instance, radiative transfer models. The availability of data [6] and computer processing requirements needed for crop growth modelling make empirical models more versatile to use in practice for limited datasets, which is the case of this study. Furthermore, in order to scale field-level grain yield predictions regionally, crop type mapping is essential as it can target the fields on which to apply the grain yield prediction models. Both field-level classifications and total wheat cropland area estimations are of great interest for improving agricultural management.

Wheat grain yield estimation at the field level before harvest plays an important role in easing farmers' management challenges and livelihoods. Studying field grain yield at the regional scale also allows for analysis of the most suitable sites for wheat cropland growth by exploring which spatial factors affect crop performance. Moreover, it can potentially be useful to estimate local and regional taxes and subsidies on agricultural production. More often than not, estimating grain yield at field level has been laborious and complex, also regarding the available remote sensing technologies in temporal frequency and spatial and spectral resolution. One challenging factor is the availability of the information regarding field-level agricultural yields, which may be considered sensitive information and hard to obtain due to market interests and other socioeconomic factors, as well as due to the limited ground measures capacity of researchers. This can jeopardize the training and validation dataset availability for crop yield models at a field level in actual agricultural contexts.

Moreover, besides crop growing conditions, such models depend on the local geomorphological contexts (i.e., topographic features) [7]. Defourny et al. [8] also stated that a significant agro-climatic gradient spanning over a large territory gradually shifts the cropping calendar of most crops, as well as the crop type distribution, and accordingly affects crop performances. Thus, regional specific approaches need to be considered when analyzing field-level performances, such as in the study here presented, where phenology is matched by combining different satellite dates for each separate ecozone in a geographically diverse region that may otherwise be captured by a single satellite scene. Several global and country/region-level initiatives assessing crop yields have been developed [9–15]. However, there is a reduced number of studies working with accurate field-level crop yield estimates [16] due to the difficult accessibility of actual crop yield data and its uncommon combination with field-level crop type classifications, both of which are very dependent on ground validation measurements.

Regarding the available remote sensing technologies for estimating grain yield, the launches of Sentinel-2 a and b satellites brought significant improvements towards functional wheat grain yield estimation, thanks to its improved temporal, spectral, and spatial resolution and open accessibility. Several studies showed Sentinel-2 data capacity for building crop yield models in wheat [4,17–19], as well as crop yields prediction in other crops such as corn [20,21] or sorghum [16,22]. Nonetheless, despite the promising results, such models have been trained and validated based on one single or, at most, a dozen fields in one or two-season periods [4,18,20,21], with clearly constrained data availability and limited regional representativeness or adaptation to geographic variability. It is also common that governments only make crop yields at the regional or district level publicly available [17,19], resulting in limitations for training field-level grain yield estimation models. Various models have been trained and validated with partial in-field grain yield measurements (i.e., partial field crop cuts) [16,22], which are limited regarding pixel errors (reduced harvested area vs. satellite

spatial resolution). Added limitations in relation to the lack of representativeness of the crop cut in comparison to the whole field-level grain yield are expected in addition to difficulties when applying such models at a broader, namely regional scale. If such models aim to be applied to actual crop fields, addressing field-level yields, capturing field level variability, and obtaining such accurate data in coordination with farmers while stressing the mutual benefits of such an approach is pivotal.

Regarding cropland and crop type classification, Sentinel-2 imagery has also been used to successfully classify and map agricultural sites [23–27]. Moreover, the Sentinel-2 improved multispectral and temporal frequency can champion grain yield estimation [28] and crop type classification in comparison with other openly accessible satellites [29]. Thus, the use of Sentinel-2 imagery and field-level wheat grain yield data in coordination with the local farmers brings a novel approach for estimating wheat performance and allows for exploring the context effects (i.e., topographic features and rainfall) on wheat grain yield at the field-level.

Mapping wheat grain yield regionally can also contribute to furthering our understanding of spatial factors that impact crop performance. In agriculture, a relatively important limiting factor is topography, as many topographically influenced landscape attributes affect grain yield [30]. Examples of topographical features that are likely to affect yields are slope [31], altitude and aspect [32]. Besides topography, water is often the main limiting factor in agricultural performance impeding optimal yields. Therefore, understanding the spatial dynamics of rainfall can help in understanding grain yield gaps, especially in rain-fed wheat agrosystems like those prevailing in Navarre that also represent the great majority of wheat cultivation environments in Spain [33].

Field-level yield data for training and validation is difficult to obtain as it is often collected in sample plots (crop cuts) or obtained from governmental regional/district level datasets. The objective of this study was to use actual field-level grain yield data in coordination with the local farmers in order to develop empirical models for predicting bread wheat grain yield at the field-level using European Space Agency (ESA) Sentinel-2 imagery in Navarre (Northern Spain). We also aimed to explore different spatial factors influencing wheat grain yield regionally. In order to achieve these objectives, we developed a combination of techniques to take maximal advantage of the still novel aspects of the recent full operability of Sentinel-2 a + b constellation in terms of its temporal, spatial, and spectral resolution. We aimed to maximize the benefits of improved Sentinel-2 temporal resolution by matching satellite data with phenology over a diverse geographical region. Increased spatial resolution improvements were used for an adequate accuracy of areal crop estimates, and also used for improving empirical models with zonal statistical summaries (maximum, minimum, mean, and median) that take advantage of the increased number of total pixels per field in small- and medium-scale agricultural parcels. We also took advantage of the improved spectral characteristics, which are pivotal for a refined crop monitoring, with stepwise multilinear equations to explore zonal statistic combinations of ten different vegetation indices calculated per field. Furthermore, wheat-growing parcels in the region were mapped using the random forest classification algorithm and the most suitable Sentinel-2 spectral reflectance bands. Finally, the model developed for estimating grain yield was applied to each of the classified wheat parcels, and then spatial topographic factors (aspect, altitude, and slope) and rainfall effects on grain yield were investigated for their correlations with grain yield across all bread wheat fields for the entire Navarre region of Spain.

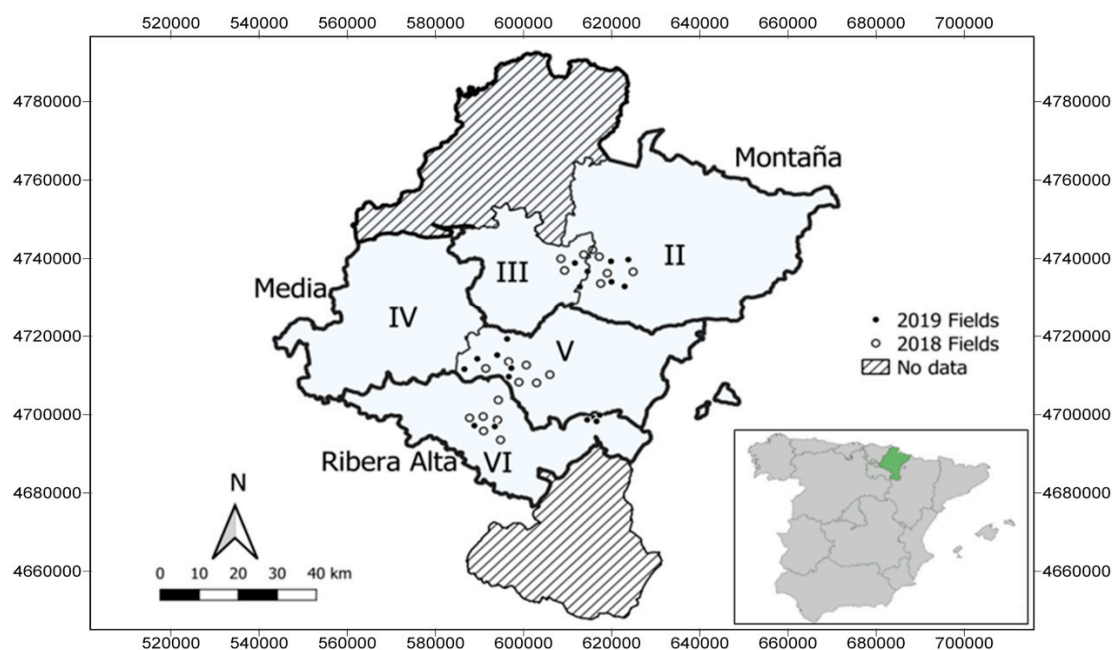
## 2. Materials and Methods

### 2.1. Data

#### 2.1.1. Study Site

The study area is the Chartered Community of Navarre in Northern Spain. Navarre's sub-regions are divided into agrarian zones, as shown in Figure 1. Montaña is divided into two zones, as well as Media, while Ribera Alta is one zone by itself. The region of study is very diverse and incorporates several of the most common Spanish agro-climates, with rainfalls averaging from 800 to 1000 mm

in the northern areas, between 800 to 300 mm in the middle and under 300 mm in the southern areas. The two most cultivated rain-fed winter crops are bread wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.), which occupy the 83% of winter croplands in Navarre [34]. The cultivation of bread wheat is ubiquitous throughout Navarre, except for the Northern Pyrenean and Atlantic areas with more rugged topography, lower temperatures, and excess precipitation, as well as in southern bottom-end semi-arid areas, which are too arid and thus unsuitable for growing commercially rain-fed wheat. The majority of crop lands are middle to small holdings, which has made them previously difficult to map adequately with satellite data.



**Figure 1.** In the bottom right corner is shown a map of Spain with Navarre highlighted in green. On the main map the fields of study are point-marked in white for the seasons 2018 and in black for 2019. The three Navarre's sub-regions: the northern (Montaña), the middle (Media), and the Southern (Ribera Alta) are also indicated. The agrarian zones are indicated with roman numerals: in Montaña zones II and III, in Media zones V and IV, and in Ribera Alta zone VI. UTM (Universal Transverse Mercator) coordinates are expressed in meters.

The region has an advanced agricultural intensification with high agronomical standards in terms of cultivation (e.g., certified seeds and widespread access to fertilizers). The regional three-season crop rotation guidelines for bread wheat are fallow, followed by fava beans, oat or sunflower, and bread wheat again in the following third season. Rotation practices are widely encouraged by agrarian institutions. Wheat is sown between the middle to the end of October depending on the agrarian region within Navarre, harvesting happens between the end of June and the middle of July. Grains are mainly commercialized to markets through various agrarian cooperatives distributed throughout the region.

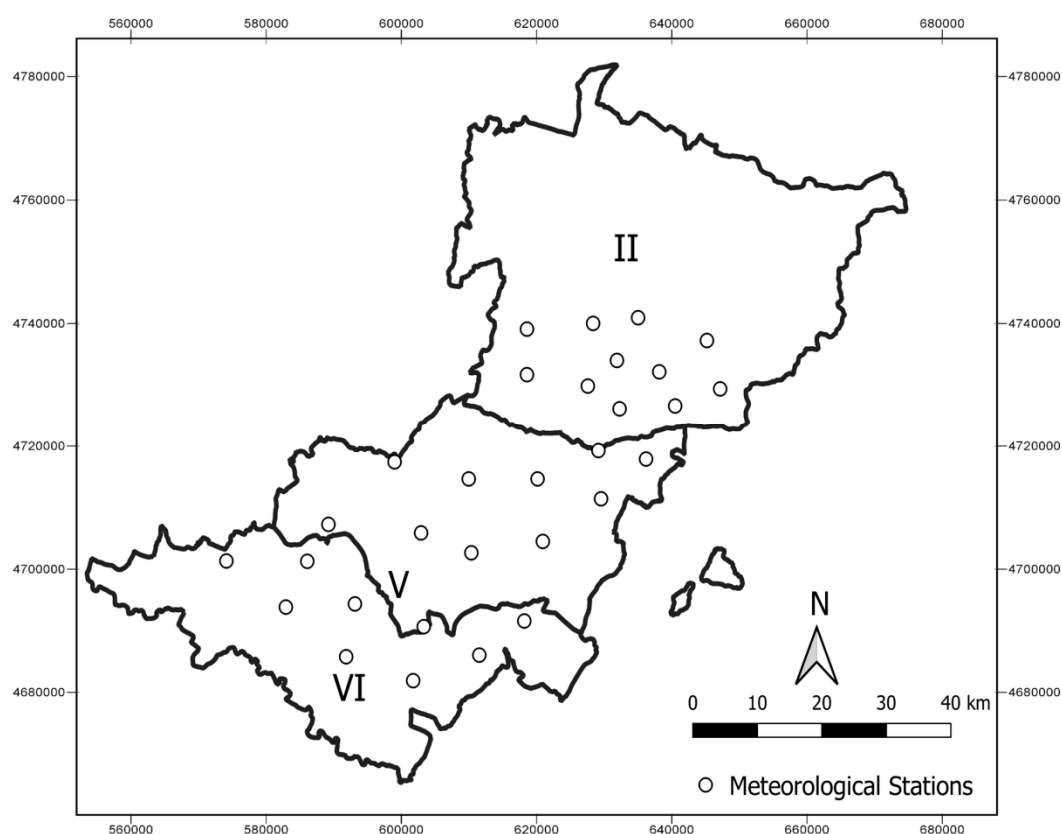
### 2.1.2. Field Data

During two growing seasons (2017–2018 and 2018–2019), 39 farmer-managed agricultural fields growing either the Marcopolo or Camargo varieties of bread wheat were studied. The fields were different for each season. Both Marcopolo and Camargo are winter, long-cycle genotypes. For both years, 15 fields were selected in the northern region (Montaña), 12 in the middle region (Media), and 12 in the southern region (Ribera Alta) (Figure 1). In 2017–2018 growing season, 8 fields were located in Montana, 6 in Media, and 6 in Ribera Alta; whereas in 2018–2019 season, 7 fields were located in Montana, 6 in Media, and 6 in Ribera Alta. Periodical field visits were programmed in order to

follow the phenological stage of the crop and report any specific growth-related issues. As reported by the farmers the analyzed fields received a basal fertilization with either superphosphate 45% or pig slurry, top-dressing fertilization was also applied with granulated urea at the majority of fields. In the Montaña region wheat was sown around the 20th October, in the Media region around the 28th October, and in the Ribera Alta around the 30th October. The harvest was collected around the middle of July in Montaña and during the first week of July in Media and Ribera Alta. The grain yield ( $\text{Kg}\cdot\text{ha}^{-1}$ ) per field was reported to the researchers by the participating farmers and standardized in terms of humidity.

### 2.1.3. Meteorological Data

Temperature data, in order to estimate the crop phenological stages through the calculation of growing degree days (GDD) in zones II, V, and VI, was obtained from 30 openly accessible regional government meteorological stations ([www.meteo.navarra.es/estaciones](http://www.meteo.navarra.es/estaciones)), shown in Figure 2. Rainfall data was also obtained from the same 30 meteorological stations distributed throughout the three agricultural zones under study. The GDDs for each of the two varieties for rain-fed management in Navarre is shown in Table 1 [35].



**Figure 2.** Meteorological station locations from which temperature and rainfall data was obtained for the three zones (II, V and VI). UTM coordinates are expressed in meters.

**Table 1.** Growing degree days (GDD) and Zadoks scale [36] for the two analyzed varieties.

Phenology (Zadoks Scale)	Camargo (GDD)	Marcopolo (GDD)
Tillering (26–30)	598	563
Heading (55–59)	1060	1150
Ripening (75–99)	1783	1819

Following average minimum and maximum mean temperatures of the 30 stations at the three zones, the accumulated GDD was calculated (Equation (1)) following calculations by Arnold [37] to estimate zonal phenological dates for 2018 and 2019 seasons. The reported dates that adjust to the phenological stage regarding GDD (Table 1) are reported in Table 2; as Camargo and Marcopolo had slightly different GDD for each stage, the average of both was considered to estimate the zonal phenological date.

$$GDD = \sum_{jun}^{nov} \frac{T_{max} + T_{min}}{2} - T_{base} \quad (1)$$

where GDD is the growing degree days,  $\sum_{jun}^{nov}$  indicates the sum throughout the season, November to June, of daily maximum and minimum temperatures ( $T_{max}$  and  $T_{min}$ ) divided by 2, minus the base temperature, which in this case was considered to be 0 °C.

**Table 2.** Estimated phenological day of year regarding GDD average (Table 1) for the seasons analyzed (2018 and 2019).

	Montaña (II)	Media (V)	Ribera Alta (VI)
<b>Tillering</b>	30-03-2018; 28-03-2019	03-04-2018; 29-03-2019	18-03-2018; 22-03-2019
<b>Heading</b>	10-05-2018; 08-05-2019	17-05-2018; 15-05-2019	26-04-2018; 29-04-2019
<b>Ripening</b>	25-06-2018; 27-06-2019	19-06-2018; 20-06-2019	05-06-2018; 06-06-2019

#### 2.1.4. Sentinel-2 Imagery and Phenology

The specific dates of Sentinel-2 images for 2018 were selected as those closest to the estimated phenological date (regarding GDD calculations) per zone as detailed in Table 2 and with the minimal cloud cover thresholds; the specific Sentinel-2 images used in this study are shown in Table 3. The Sentinel-2 a + b satellite constellation's combined global coverage every 5 days enables this level of phenology matching for the first time ever since its second satellite was launched in 2017.

**Table 3.** Closest Sentinel-2 images available to the estimated phenological dates in 2018 (Table 2), used for building the prediction model.

	Montaña (II)	Media (V)	Ribera Alta (VI)
<b>Tillering</b>	30-03-2018; 14-04-2018	30-03-2018; 14-04-2018	15-03-2018; 30-03-2018
<b>Heading</b>	09-05-2018; 19-05-2018	19-05-2018; 24-05-2018	24-04-2018; 09-05-2018
<b>Ripening</b>	18-06-2018; 23-06-2018	18-06-2018; 23-06-2018	03-06-2018; 18-06-2018

Sentinel-2 multispectral bands (Table 4) were downloaded from Copernicus Open Access Hub (<https://scihub.copernicus.eu/>). Despite the high 5-day temporal resolution of Sentinel-2 a + b constellation, some images were discarded due to the image-wide cloud cover areas exceeding 40%. Four Sentinel-2 tiles cover the region under study. The images were downloaded for the 2018 (15-03, 30-03, 14-04, 24-04, 09-05, 19-05, 24-05, 03-06, 18-06 and 23-06) and 2019 (15-03, 20-03, 25-03, 30-03, 09-04, 29-04, 09-05, 14-05, 08-06, 18-06 and 28-06) seasons. Images were downloaded as an L1C product and were corrected to level 2A using the Sen2Cor tool on SNAP (Sentinel Application Platform), obtaining Bottom-Of-Atmosphere (BOA) and cirrus corrected reflectance images. Moreover, a cloud and cloud shadow mask were also applied with this tool.

From the Sentinel-2 multispectral data bands (Table 4), the following spectral reflectance vegetation indices (VIs), which are simple mathematic combinations of spectral bands, were calculated: the chlorophyll index green (CI green) and the chlorophyll index red edge (CI red edge), related to chlorophyll content [38]; the normalized difference water index (NDWI), related to water content [39]; the modified soil-adjusted vegetation index (MSAVI) [40] and the optimized soil adjusted vegetation index (OSAVI) [41], associated with vegetation cover; and the renormalized difference vegetation index



(RDVI) [42], the ratio vegetation index (RVI) [43], the modified simple ration (MSR) [44], the normalized difference vegetation index (NDVI) [45], and the green normalized difference vegetation index (GNDVI) [46] sensitive to biomass (Table 5).

**Table 4.** Spectral bands and spatial resolutions of the Sentinel-2 Multispectral Instrument (MSI). Novelty in spectral coverage includes three red-edge spectral bands and the vegetation red-edge (RE) at 20 m as well as improved SWIR (Short-Wave InfraRed) coverage at 20 and 60 m spatial resolutions. Broadband spectral coverage of the visible and near infrared are provided at 10 m spatial resolution.

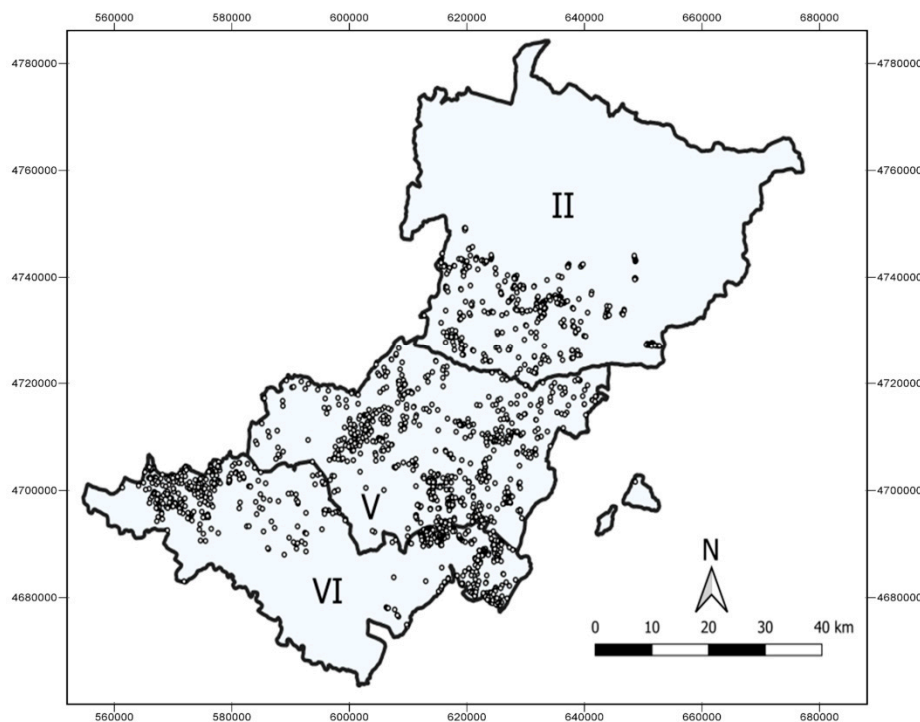
MSI Band	Spatial Resolution (m)	Central Wavelength (nm)
B1: Coastal Aerosol	60	443
B2: Blue	10	490
B3: Green	10	560
B4: Red	10	665
B5: Red-Edge	20	705
B6: Red-Edge	20	740
B7: Red-Edge	20	783
B8: NIR	10	842
B8A: Vegetation RE	20	865
B9: Water Vapour	60	945
B10: SWIR Cirrus	60	1375
B11: SWIR	20	1610
B12: SWIR	20	2190

**Table 5.** Vegetation indices calculation regarding Sentinel-2 bands, B refers a specific band, among those reported in Table 4. The acronyms refer to: chlorophyll index green (CI green), chlorophyll index red edge (CI red edge), normalized difference water index (NDWI), the modified soil-adjusted vegetation index (MSAVI), optimized soil adjusted vegetation index (OSAVI), renormalized difference vegetation index (RDVI), ratio vegetation index (RVI), modified simple ration (MSR), normalized difference vegetation index (NDVI), and the normalized difference vegetation index (GNDVI).

	S-2 Formula	Measures	Reference
CI Green	$(B7/B3) - 1$	Chl content	Gitelson et al. (2003)
CI Red Edge	$(B7/B5) - 1$	Chl content	Gitelson et al. (2003)
NDWI	$(B8 - B11)/(B8 + B11)$	Water content	McFeeters et al. (1996)
OSAVI	$1.16 \cdot (B8 - B4)/(B8 + B4 + 0.16)$	Vegetation cover	Baret et al. (1993)
MSAVI	$[2 \cdot B8 + 1 - \sqrt{(2 \cdot (B8 + 1)^2 - 8 \cdot (B8 - B4))}] / 2$	Vegetation cover	Qi et al. (1994)
MSR	$(B8/B4 - 1)/[\sqrt{(B8/B4 + 1)}]$	Biomass	Chen et al. (1996)
NDVI	$(B8 - B4)/(B8 + B4)$	Biomass	Rouse Jr. (1974)
GNDVI	$(B8 - B3)/(B8 + B3)$	Biomass	Louhaichi (2001)
RDVI	$(B8 - B4)/\sqrt{(B8 + B4)}$	Biomass	Roujean (1995)
RVI	$B8/B4$	Biomass	Jordan (1969)

### 2.1.5. Polygons

The agricultural parcels delineations were obtained from Sistema de Información Geográfica de Parcelas Agrícolas (SIGPAC), a Spanish governmental application (<http://sigpac.mapa.gob.es/>) that allows identifying specific declared agricultural fields subjected to Common Agricultural Policy (CAP) subsidies. GPS coordinates were recorded to delineate the 39 fields used for training and validating the grain yield model. Crop type points were coordinated in fields throughout the three zones (Figure 3) by obtaining GPS coordinates from each of the classified types (wheat, bare soil, and other crops). The whole regional agricultural parcels polygons were downloaded for the attributes herbaceous crops and rain-fed for the three agrarian areas under study within Navarre. Furthermore, fields were 10-m-buffered with the ArcGis Pro 2.3.0 *buffer* tool in order to reduce edge effects.



**Figure 3.** Three stratified areas corresponding to three agrarian zones, II in Montaña, V in media, and VI in Ribera Alta. The stratification for crop type classification and the corresponding training points. UTM coordinates are expressed in meters.

## 2.2. Models and Analysis

### 2.2.1. Grain Yield Estimation

At each of the three studied phenological stages (tillering, heading, and ripening), the previously described vegetation indices were calculated per field on ArcGis Pro 2.3.0 using the downloaded Sentinel-2 imagery. Four different statistical summary extractions of each vegetation index (maximum, minimum, mean, and median) were obtained for each field at each different phenological stage, considering the regional differences outlined in Tables 2 and 3. Mean and median were calculated with the closest available images to the phenological stage using zonal statistics, while minimum and maximum were respectively the averages of the lowest and highest pixel values between the two closest dates to the phenological stage. These calculations were performed on ArcGis Pro 2.3.0 and Microsoft Excel. First, the *Raster to Point* and *Spatial Join* tools on ArcGis were used in order to list, per field, all the pixel values; second, with the corresponding attribute tables the maximum and minimum vegetation indices pixel values between the dates were calculated with Microsoft Excel. Finally, the maximum and minimum were calculated on the points' attribute table were converted into two rasters in ArcGis in order to perform zonal statistics and obtain maximum and minimum averages per field over the specific periods. Pearson regressions against grain yield were calculated on R studio ("cor.test") for each phenological stage and summary (min, max, mean, and median). Following this, in order to find the combination of the best vegetation indices and statistical summary to estimate grain yield, multilinear stepwise regressions were studied at the most suitable phenological stage in R Studio (library "MASS"). The selected best model was chosen on the basis of Akaike Information Criterion, computed with the stepAIC function from the MASS library. The normal distribution of residuals and collinearity of the independent variables were studied with the library "olsrr", the variance inflator factor (VIF) was measured as it is an efficient method to asses collinearity [47–49], which strongly limits stepwise selection method. Furthermore, the residuals normality Shapiro-Wilk test was calculated.

The 39 monitored fields were separated into two groups, one for training and the other one for validation of the multilinear stepwise equations. The 20 fields from 2018 season were used for modeling and the 19 fields from 2019 were used for validation. For the 2019 season, only one phenological stage, heading, was eventually analyzed, as it appeared to be the most suitable for grain yield estimation (Table 7), with the following images: Montaña, 09-05 and 14-05; Media, 14-05 and 08-06; and Ribera Alta, 29-04 and 09-05.

### 2.2.2. Crop Type Classification and Regional Per-Parcel Grain Yield Estimation

We performed a pixel-based supervised random forest (RF) classification to differentiate wheat fields from fields of other crops. The classified area was stratified in three zones: II, V, and VI, corresponding to three of Navarre's agricultural zones shown in Figure 1. Each one located in Montaña, Media, and Ribera Alta areas, respectively. Following Colditz [50] and Noi and Kappas [27], which state that for RF classifications the training sample should correspond to around 0.25% of the total area, 513 points of parcels of the ground dataset in Zone II (33,345 pixels), 927 in Zone V (58,401 pixels), and 1050 in Zone VI (60,900 pixels) formed the training dataset. Figure 3 indicates the points of the parcels from which the pixels from the various analyzed classes were obtained. The points collected were classified into three classes: wheat, bare soil, and other crops. As mentioned before, the target classification areas were delimited with the available rain-fed and herbaceous crops polygons mask from SIGPAC.

Considering that the majority of cropland under study (83%) is either bread wheat or barley, we took advantage of the later ripening of bread wheat in order to differentiate it from other crops. After considering both preliminary results and phenological characteristics, differences between wheat and barley were clearer late in the season. This is a consequence of the shorter crop duration of barley, which makes it reach maturity and change color before wheat. Hence, late season Sentinel-2 images (24-05-2018, 03-06-2018 and 18-06-2018) were used for the classification as the moment of greatest contrast.

Due to their better spatial resolution (10 m) and contrasting capacities in comparison with its other bands, Sentinel-2 B3 (green), B4 (red), and B8 (near infrared) bands were used for the crop type classification. Crop classifications [26,51] and field boundary delimitation [52] were successfully mapped using B3-B4-B8 images in previous studies with Sentinel-2 data. The trained datasets were divided as follows, 2/3 for training the model and 1/3 for validation. The number of trees was 500 [53]. RF classification and object filters were computed with ArcGis Pro 2.3.0. We applied majority voting of the classified pixels within the agricultural parcels (SIGPAC), as field crops are exclusively grown in monocultures in the region of Navarre, which we confirmed during farmer visits. The pixel's majority from the zonal statistics tool was then used to object filter the classification. With the validation data and the output of classification, a confusion matrix was computed at parcel level in order to obtain classification accuracies. At parcel level, the Fscore of each class was calculated with the Equation (2), where PA refers to producer's accuracy while UA refers to user's accuracy.

$$Fscore = \frac{2 * PA * UA}{PA + UA} \quad (2)$$

The most suitable multilinear stepwise regression model was extended into all the classified parcels of bread wheat in order to estimate grain yield per field for the three agricultural sub-regions. The resulting wheat cropland and grain yield average per zone were compared to the Navarre regional government's official statistics ([www.navarra.es](http://www.navarra.es)).

### 2.2.3. Rainfall Interpolation and Topographic Models

A 2 m resolution digital elevation model (DEM) was obtained from Navarre's regional Government at <ftp://ftp.cartografia.navarra.es/>. The downloaded tiles were mosaicked and slope and aspect were calculated from the DEM using the *slope* and *aspect* tools.



The sum of the accumulated rainfall (mm) during the wheat growing season, from November to May, at every station was gathered in a georeferenced-points dataset. This dataset was used to interpolate rainfall using two strategies: inverse distance weighting (IDW) and kriging. All DEM, topographic feature and subsequent GIS processing were completed on ArcGis Pro 2.3.0.

#### 2.2.4. Ordinary Least Square and Geographically Weighted Regression

The ordinary least square (OLS) was used in the first term to explore the relationship between grain yield and the explanatory variables. In OLS the existence of local variation is not taken into account in the regression, hence the regression coefficient remains constant for each variable (Equation (3)).

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \quad (3)$$

where  $y_i$  is the dependent variable (grain yield),  $\beta_0$  is the intercept;  $\beta_k$  is the coefficient of the independent variable ( $x_{ik}$ ) and the random error is  $\varepsilon_i$ .

Alternatively, GWR 4.0 software, developed by Nakaya et al. [54], was used. A geographically weighted regression (GWR) model was also calculated in order to explore the relationship between grain yield and the explanatory variables (Equation (4)) taking into account the spatial factor, which is an essential feature when dealing with spatial heterogeneity, something that is especially central to the objectives of this study. GWR uses the least square method given the location as a weighting factor. The optimal bandwidth, namely the optimum number of neighbors, was determined by the Akaike Information Criterion (AIC). A multiple comparison of AIC to find the best bandwidth (the one with the lowest AIC) was computed. For OLS and GWR comparison, an ANOVA was performed to compare the accuracy of both levels. Furthermore, for both OLS and GWR, the spatial autocorrelation (Moran's I) of residuals were tested.

$$y_i = \beta_0(\mu_i, v_i) + \sum_k \beta_k(\mu_i, v_i) x_{ik} + \varepsilon_i \quad (4)$$

where  $y_i$  is the dependent variable (grain yield),  $(\mu_i, v_i)$  are the coordinates (x,y) at the location  $i$ ,  $\beta_0$  is the intercept,  $\beta_k$  is the coefficient of the independent variable ( $x_{ik}$ ) at the specific weighted location, and  $\varepsilon_i$  is the random error.

A geographic variability test was performed in order to determine if the explanatory variables were spatially heterogeneous. In order to do so, two models were developed, one that considers all of the variables (slope, altitude, height, and rainfall) to be spatially heterogeneous, and on the other hand one that considers all of the variables to be spatially homogenous. If the variables were indeed spatially heterogeneous determinants of grain yield, i.e., if the coefficients varied significantly in space, then the AIC size of the second model should be larger and result in a negative diff-criterion value for this test (GWR4 manual) [55].

### 3. Results

#### 3.1. Grain Yield Estimation

In Table 6 we show the Pearson's correlation between the calculated vegetation indices and the actual grain yield at the three studied phenological stages (tillering, heading, and ripening) for the various extracted zonal statistics (min, max, mean, and median). The correlations were substantially higher at heading in comparison with tillering and ripening. The resulting correlations against grain yield were similar among the chlorophyll content, water content, vegetation cover, and biomass-sensitive vegetation indices, with no obvious results to determine better correlating indices with the Pearson's correlation. While at heading the correlations averaged between 0.79 to 0.84 for the mean, 0.76 to 0.84 for the median, 0.72 to 0.81 for the minimum, and 0.82 to 0.84 for the maximum, at tillering the results were between 0.41 and 0.74 for the mean, 0.40 and 0.71 for the median, 0.38 and 0.63 for the minimum, and 0.43 and 0.65 for the maximum. Regarding ripening,

the correlations yielded between 0.32 and 0.49 for the mean, 0.31 and 0.49 for the median, 0.38 and 0.45 for the minimum, and 0.34 and 0.45 for the maximum. Furthermore, the statistical significance of the correlations at heading was greater than that at both tillering and ripening, which demonstrated lesser ( $<0.05$ ) or no significance differences in the majority of the cases.

**Table 6.** Pearson correlation coefficients (R). Relationship of vegetation indices (VIs) and grain yield at three phenological stages and four statistical approaches. Statistical significance is indicated by \*\* and \* at 99% and 95%, respectively.

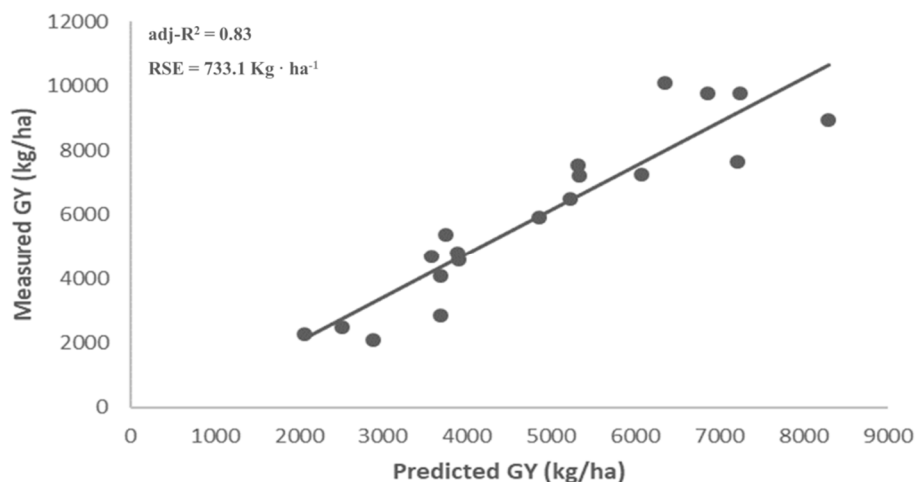
MEAN	Tillering	Heading	Ripening	MEDIAN	Tillering	Heading	Ripening
CI Green	0.41	0.81 **	0.32	CI Green	0.40	0.80 **	0.31
CI Red	0.61 *	0.79 **	0.46 *	CI Red	0.59 *	0.76 **	0.44 *
Edge				Edge			
GNDVI	0.63 *	0.82 **	0.41	GNDVI	0.60 *	0.80 **	0.41
MSAVI	0.74 **	0.83 **	0.44	MSAVI	0.71 **	0.81 **	0.43
MSR	0.61 *	0.82 **	0.49 *	MSR	0.59 *	0.82 **	0.47 *
NDVI	0.65 *	0.83 **	0.48 *	NDVI	0.61 *	0.82 **	0.49 *
NDWI	0.69 **	0.84 **	0.49 *	NDWI	0.66 **	0.84 **	0.49 *
OSAVI	0.66 **	0.83 **	0.47 *	OSAVI	0.64 **	0.81 **	0.48 *
RDVI	0.68 **	0.83 **	0.42	RDVI	0.67 **	0.82 **	0.44
RVI	0.59 *	0.81 **	0.48 *	RVI	0.57 *	0.81 **	0.47 *
MIN	Tillering	Heading	Ripening	MAX	Tillering	Heading	Ripening
CI Green	0.38	0.72 **	0.40	CI Green	0.43	0.82 **	0.34
CI Red	0.59 *	0.75 **	0.41 *	CI Red	0.61 *	0.83 **	0.41 *
Edge				Edge			
GNDVI	0.56 **	0.75 **	0.41	GNDVI	0.63 *	0.83 **	0.41
MSAVI	0.63 **	0.81 **	0.40 *	MSAVI	0.63 *	0.84 **	0.40 *
MSR	0.61 *	0.77 **	0.39	MSR	0.60 *	0.83 **	0.39
NDVI	0.62 **	0.77 **	0.38 *	NDVI	0.65 *	0.81 **	0.42 *
NDWI	0.63 **	0.79 **	0.39 *	NDWI	0.65 **	0.83 **	0.41 *
OSAVI	0.62 **	0.79 **	0.41	OSAVI	0.64 *	0.84 **	0.45
RDVI	0.63 *	0.80 **	0.45 *	RDVI	0.62 **	0.84 **	0.45 *
RVI	0.53 *	0.77 **	0.38	RVI	0.56 *	0.83 **	0.42

At heading, the results from the multilinear stepwise regression (Table 7) show the most suitable combination of vegetation indices in order to estimate grain yield. In this sense, the best-suited equation uses minimum averages of MSAVI pixels and maximum averages of RDVI pixels. VIF was 2.16 for both variables. The histogram of the residuals is shown in Figure A1 in the Appendix A together with the Shapiro-Wilk normality test results, the null hypotheses of which assumes that the data is normal while the alternative assumes it is not. The  $p$  value of the test was 0.297 and the statistic 0.982.

The resulting multilinear stepwise equation (Table 7) was tested at heading for the studied fields for the following year 2019 using the corresponding calculated vegetation indices and summary statistics. The equations worked accurately as shown in Figure 4, with an  $R^2$  of 0.83 ( $RSE = 733.1 \text{ kg}\cdot\text{ha}^{-1}$ ).

**Table 7.** Multilinear stepwise equation estimating grain yield (GY) using two (MIN MSAVI and MAX RDVI) variables. Stepwise equation estimate,  $p$ -value and variance inflation factor (VIF) are shown together with the summarized equation. The model was developed with data from the 2018 wheat growing season at the indicated fields shown in Figure 1.

Parameter	Estimate	$p$	VIF
Intercept	−1290	$<0.05$	
MIN MSAVI	7085	$<0.05$	2.16
MAX RDVI	10805	$<0.01$	2.16
<b>Model Summary</b>			
$GY = -1290 + 7085 \cdot \text{MIN\_MSAVI} + 10805 \cdot \text{MAX\_RDVI}$			



**Figure 4.** Stepwise equation validation, correlation of predicted against actual 2019 grain yield (GY). The adjusted  $R^2$  correlation of the estimated GY ( $\text{kg} \cdot \text{ha}^{-1}$ ) with the stepwise equation developed for the previous season against reported grain yields in 2019.

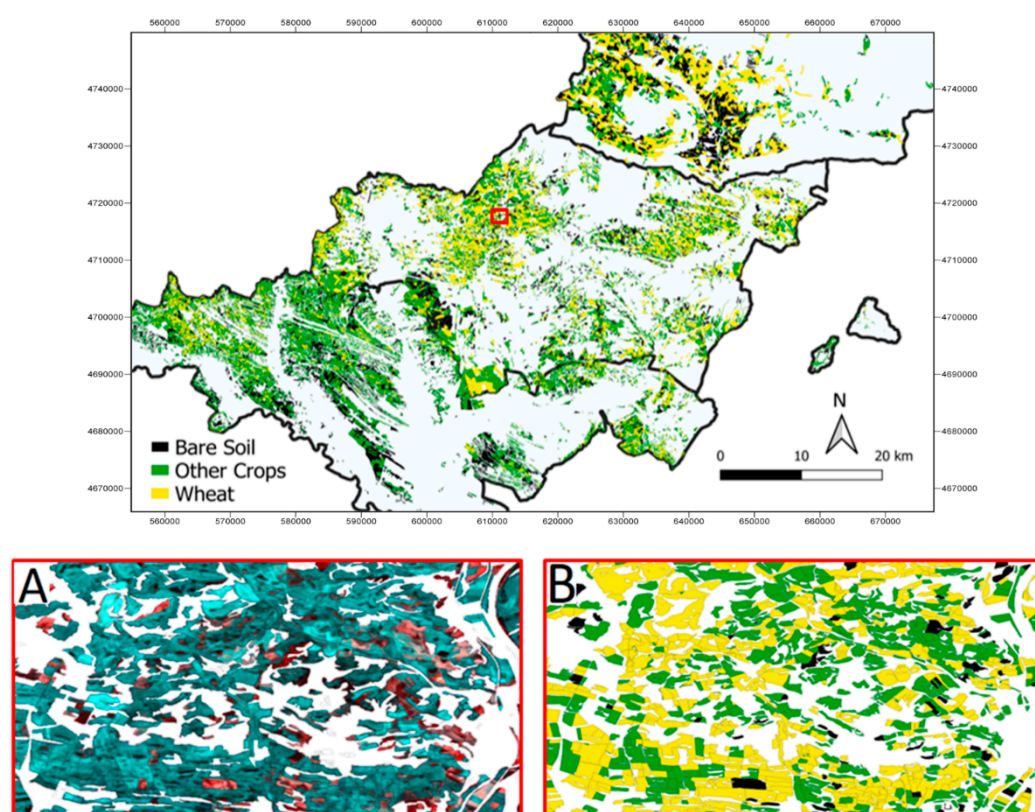
### 3.2. Crop Type Classification

The overall classification accuracy reached 86.55%, 89.32%, and 83.14% for the II, V, and VI zones, respectively (Table 8). Wheat was well-mapped at the three classified sites with an Fscore of 88% at the agrarian sub-regions II, 91.90% at zone V, and 84.05% at zone VI. On the other hand, bare soil was reliably classified with an Fscore of 91.43% at zone II, 93.26% at zone V, and 84.06% at zone VI. The class “other crops” was the least well classified, the corresponding Fscores at the three areas where 78.15%, 78.26%, and 81.82% for the zones II, V and VI, respectively. Figure 5 shows the crop type classified map.

**Table 8.** Confusion matrixes for the pixel-based crop type classification at the three stratified agrarian regions under study. Zone II in the northern Montaña region, zone V in the Media region, and zone VI in the Ribera Alta Southern region. The confusion matrix was calculated after the object filtering of fields. Thus, the number indicates the classification accuracy of fields and not pixels. UA indicates users’ accuracy and PA producers’ accuracy. The overall accuracy at the bottom right of each zone is highlighted in bold.

ZONE II	Bare Soil	Other Crops	Wheat	Total	UA [%]
Bare Soil	48	5	2	55	87.27
Other Crops	0	34	4	38	89.47
Wheat	2	10	66	78	84.62
Total	50	49	72	171	
PA [%]	96.00	69.38	91.67		<b>86.55</b>
ZONE V	Bare Soil	Other Crops	Wheat	Total	UA [%]
Bare Soil	90	8	1	99	90.91
Other Crops	2	54	5	61	88.53
Wheat	2	15	132	149	88.59
Total	94	77	138	309	
PA [%]	95.74	70.13	95.65		<b>89.32</b>
ZONE VI	Bare Soil	Other Crops	Wheat	Total	UA [%]
Bare Soil	87	23	3	113	76.99
Other Crops	3	117	10	130	90.00
Wheat	4	16	87	107	81.30
Total	94	156	100	350	
PA [%]	92.55	75.00	87.00		<b>83.14</b>

The estimated grain yield average for zone II corresponded to 103% of the official data, while zone V and VI correspond to 95% and 94%, respectively (Table 9). Regarding wheat crop land, the classified wheat field area can be an indicator of the crop type area. In this sense, the estimated wheat crop land area for the whole region corresponded to 85.77% of the estimate in the official statistics. Regarding each zone, the surface corresponded to 81.41%, 81.15%, and 123.16% with the zones I, V, and VI respectively. Regarding zone VI, the estimated wheat cropland was overestimated if compared with official statistics, while the opposite happened at sites I and V. We trust the area estimation indicator as the polygons used for the object-filtering are official trusted vectors, as well as the mask used for the classification itself. All Spanish cereal producers are required to declare their croplands, and thus the polygons used in this study should be fairly accurate.



**Figure 5.** Crop type map output obtained from Sentinel-2 imagery and RF classification for the three segregated zones. The zoom-in shows the false color B3-B4-B8 image (A) and a closer classification output (parcels filtered) (B). UTM coordinates are expressed in meters.

**Table 9.** Statistics for grain yield (GY), total wheat crop land surface studied, and farmer field average surface from both the official regional Government of Navarre ([www.navarra.es](http://www.navarra.es)) and the calculated results for the season 2017–2018 at the regions Montaña (zone II), Media (zone V), and Ribera Alta (zone VI).

Zone	Estimated GY Average (kg·ha <sup>-1</sup> )	Official GY Average (kg·ha <sup>-1</sup> )	Wheat Crop Land Surface Estimation (ha)	Official Wheat Crop Land Surface (ha)	Field Surface Average (ha)
II	5142 ± 1223	4992	6906	8483	7.37 ± 3.82
V	4553 ± 1304	4804	9589	11,817	5.24 ± 2.27
VI	3697 ± 1093	3932	3020	2452	4.92 ± 2.92
<b>Total</b>			19,515	22,752	

### 3.3. Topographic Features and Rainfall Effects on Performance

OLS and GWR are both models that seek to explain the relationship between a dependent variable, in this case the calculated grain yield at all the classified wheat parcels, and explanatory independent variables, topographic features (aspect, slope, and altitude), and rainfall for this study. Table 10 presents GWR as the most suitable model for explaining these effects. In this case AIC, sigma and both  $R^2$  and adjusted  $R^2$  improved with GWR in comparison with OLS for the three zones (II, V, and VI) (Table 11). For instance, OLS yielded an adjusted  $R^2$  of 0.05, 0.06, and 0.07 for zones II, V, and VI respectively. Meanwhile, for the statistically significant ( $p < 0.05$ ) independent variables (rainfall and altitude in zone II; rainfall, slope, and altitude in zone V; and rainfall, slope, and altitude in zone VI) OLS can explain a 5%, a 6%, and a 7% of grain yield variability at each zone (II, V, and VI). On the other hand, GWR presents an adjusted  $R^2$  of 0.20, 0.11, and 0.20 for zones II, V, and VI respectively; therefore 20%, 11%, and 20% of GY variability at each zone can be explained with GWR. The local  $R^2$  of each wheat field is shown in Figure 6. Furthermore, in Table 10, the ANOVA results support significant improvement ( $p < 0.05$ ) when using GWR in all the three studied zones compared to the global model (OLS). The GWR model explains the effects of topographic factors (aspect, altitude, and slope) and rainfall on grain yield significantly better than the OLS model for all three zones by adding information on the spatial variation of grain yield with these local phenomena.

The autocorrelation test of both GWR and OLS for the three zones (Table 12) presents Moran's I values closer to the expected for GWR in comparison with OLS. The complete spatial randomness of GWR residuals was accepted as the  $p$ -value is not statistically significant, making the null hypothesis, that there is not significant spatial autocorrelation, correct. In contrast with this,  $p$ -value is statistically significant for OLS and therefore shows spatial autocorrelation of OLS residues.

Regarding the rainfall interpolation, kriging interpolation outperformed IDW (RMSE = 91.90 mm and 105.08 mm respectively), hence the previous was used. Grain yield correlated positively with rainfall, namely when rainfall increased, so did grain yield. Although in GWR models the regression coefficients vary locally, the fact that in zone II the mean coefficient is 13.87 (Table 11) suggests the strong effect of the rainfall gradient distribution on wheat grain yield in this northern area. Regarding rainfall at the southern region (Zone VI), the mean coefficient is also positive and relatively high, 12.32 (Table 11). Meanwhile, the mean regression coefficient in zone V was lower at 1.73. Concerning the topographic variables, they were spatially heterogeneous except for aspect in the middle zone.

**Table 10.** ANOVA test of GWR improvements over OLS for each agrarian zone: zone II in the northern Montaña region, zone V in the Media region, and zone VI in the Ribera Alta Southern region. GWR: geographically weighted regression and OLS: ordinary least squares. F refers to the F-test value.

ZONE-II	Sum of Square	Degree of Freedom	Mean Square	F	p-Value
Global Residuals	7,843,015,749.31	3147.00			
GWR Improvements	1,747,830,060.01	239.72	7,291,131.57		
GWR Residuals	6,095,185,689.30	2907.28	2,096,525.17	3.48	<0.05
ZONE-V	SS	DF	MS	F	
Global Residuals	1,137,248,334.91	2439.15			
GWR Improvements	401,137,210.77	338.37	1,185,498.75		
GWR Residuals	736,111,124.14	2100.78	350,529.11	3.38	<0.05
ZONE-VI	SS	DF	MS	F	
Global Residuals	1,023,577,071.92	2013.50			
GWR Improvements	300,131,748.19	277.42	1,081,867.74		
GWR Residuals	723,445,323.73	1736.08	416,711.97	2.59	<0.05

**Table 11.** Comparison between OLS and GWR for the four studied factors (aspect, altitude, slope, and rainfall) at the three agrarian zones: zone II in the northern Montaña region, zone V in the Media region, and zone VI in the Ribera Alta Southern region. GWR: geographically weighted regression and OLS: ordinary least squares. SE refers to standard error.

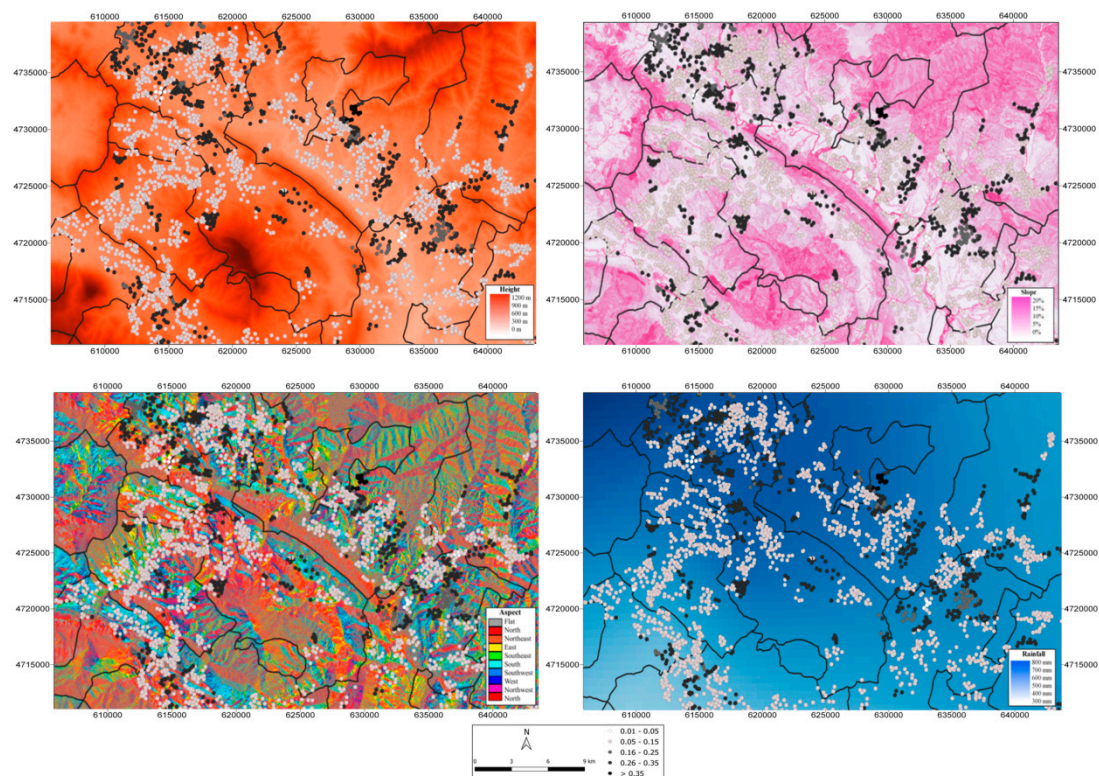
Zone II		OLS			GWR	
Variable	Estimate	SE	t (Est/SE)	p-value	Mean Coeff.	SE
Intercept	4103.53	546.3	7.51	0.00	2984.72 *	393.62
Rainfall	0.91	0.73	1.25	0.00	13.87 *	5.47
Slope	−7.03	6.5	−1.08	0.07	−14.45 *	78.07
Altitude	−3.02	0.47	−6.43	0.00	0.69 *	1.03
Aspect	−0.7	0.33	2.12	0.14	0.14 *	1.65
R-squared			0.07		0.29	
Adj			0.05		0.20	
R-Squared						
Sigma		1647.83			1433.43	
AIC		70,931.92			67,218.33	
Zone V		OLS			GWR	
Variable	Estimate	SE	t (Est/SE)	p-value	Mean Coeff.	SE
Intercept	7883.91	1028.72	7.66	0.00	5035.23 *	372.81
Rainfall	0.83	0.28	2.96	0.00	1.73 *	0.71
Slope	−16.18	7.02	−2.31	0.05	−31.22 *	9.04
Altitude	−0.88	0.33	−2.67	0.00	1.09 *	0.47
Aspect	−0.39	0.23	−1.70	0.07	−1.02	0.33
R-squared			0.08		0.15	
Adj			0.06		0.11	
R-Squared						
Sigma		823.04			676.04	
AIC		65,362.14			5772.26	
Zone VI		OLS			GWR	
Variable	Estimate	SE	t (Est/SE)	p-value	Mean Coeff.	SE
Intercept	3031.82	1521.12	1.99	0.04	7812.44 *	897.18
Rainfall	2.83	0.26	10.88	0.01	12.32 *	6.56
Slope	−61.82	12.81	−4.83	0.01	−43.79 *	16.43
Altitude	1.7	0.59	2.88	0.00	−0.97 *	11.56
Aspect	−0.02	0.19	−0.11	0.11	0.25 *	0.95
R-squared			0.09		0.30	
Adj			0.07		0.20	
R-Squared						
Sigma		847.87			698.28	
AIC		60,831.84			33,865.58	

\* Indicates that the coefficients vary significantly throughout the space (geographically variability test).

**Table 12.** Moran's I spatial autocorrelation for both OLS and GWR residuals at the three zones: Zone II in the northern Montaña region, zone V in the Media region, and zone VI in the Ribera Alta Southern region. GWR: geographically weighted regression and OLS: ordinary least squares.

RESIDUAL ZONE-II	OLS	GWR	RESIDUAL ZONE-V	OLS	GWR	RESIDUAL ZONE-VI	OLS	GWR
Moran's I	0.133	−0.015	Moran's I	0.111	−0.002	Moran's I	0.071	−0.023
Expected I	−0.002	−0.002	Expected I	−0.004	−0.004	Expected I	−0.005	−0.005
z-score	27.209	4.644	z-score	9.033	3.909	z-score	8.877	3.909
p-value	<0.01	>0.05	p-value	<0.01	>0.05	p-value	<0.01	>0.05





**Figure 6.** GWR local  $R^2$  distribution showed on the four analyzed parameters (altitude, slope, aspect, and rainfall maps, in order from top to bottom) zoomed in to zone II (north). GWR: geographically weighted regression. UTM coordinates are expressed in meters.

#### 4. Discussion

Testing the applicability of empirical vegetation index models in different years of data allows for evaluating the reliability and reproducibility of that model through time. In this case the stepwise multilinear model split between the training dataset, season 2018, and validation dataset, season 2019, revealed the efficiency of this approach for at least a two-year period in Navarre. This is uncommon, as generally such empirical models are developed and validated with same-year data [16,21], something that likely happens due to the difficulty to obtain data from various growing seasons, and thus gives robustness to the techniques and analyses that we have presented here.

The full operability of Sentinel-2 constellation reached in 2018 allowed focusing on phenological stages thanks to its improved frequency, as the results presented here show. Until this recent date, phenology-based grain yield prediction was unlikely to be assessed with alike satellites (i.e., Landsat 8). Our results suggest that heading is the most suitable phenological stage in order to study empirical relations of vegetation indices with grain yield for rain-fed wheat in Navarre. This is consistent with other ground spectral measurements for estimating wheat grain yield under rain-fed conditions as reported by Fernández-Gallego et al. [56]. Moreover, this stage was also found optimal for grain yield prediction with proximal remote sensing (i.e., UAV) and equivalent growing conditions [57]. We argue that the results here obtained showed that focusing on phenology for grain yield estimation is now possible using Sentinel-2 imagery and not only applicable with ground and proximal sensing. Furthermore, the combination of climatic (i.e., temperature, growing degree days) and satellite data has resulted in an adequate approach for estimating phenology, as previous studies have suggested [58,59], and wheat grain yield.

The minimum average of pixel values at heading stage of MSAVI and the maximum of RDVI provided the most optimal combination of vegetation indices in the stepwise multilinear regression produced in this study. The improved features of Sentinel-2 regarding spectral and spatial resolution

have allowed increasing the VIs able to be calculated and the most optimal statistic summary, which with coarser spectral and spatial resolutions would be very limited. The improved capacities of these indices could be explained due to the reduced soil background influences, regarding MSAVI, and, as detailed by Rougean and Breon [42], the capacity to minimize background reflectance effects of RDVI. In this case, MSAVI together with RDVI show the ability to maintain sensitivity to total vegetation biomass for fully developed canopies corresponding to heading, while other vegetation indices might be saturated at this phenological stage. These vegetation indices have already been successfully correlated with wheat performance using remote sensing data. In the scientific literature, we find examples of RDVI, which is used for green biomass estimation as it is sensitive to vegetation coverage [60] and has also been successfully used in empirical models estimating wheat yields [61]. Around the phenological stage of heading, Bao et al. [62] reported a correlation of 0.79 ( $R^2$ ) between RDVI and wheat yield, while Zhao et al. [63] found a correlation of 0.76 ( $R^2$ ), which show the relative high correlation ability of RDVI with wheat yield. Examples are also available for MSAVI, which has shown relative high correlation with wheat yields [64,65]. Regarding the goodness of the model, the selected vegetation indices showed absence of collinearity, as a VIF of 2.16, obtained for both cases, was lower than the collinearity threshold value of 10 [66]. Moreover, the normality of the residuals was assumed, as the Shapiro-Wilk test p value was over 0.05 (0.297) and the null hypotheses, which assumes that the data is normal, was accepted. The plot of residuals also suggested it as shown in Figure A1 of Appendix A.

The fact that maximum and minimum pixel value averages were the most suitable statistical summary combination shows that, for studying grain yield, exploring crop canopy evolution through time may provide an improved output in comparison with single date image data. However, it shows that focusing on the phenological period of heading can provide relatively higher correlations with grain yield, explaining 83% of variability in this case. Namely, segregating phenological periods is a useful approach in order to calculate empirical grain yield estimations in contrast with following whole season pixel values without explicitly considering phenological periods itself. Estimating the phenological stage with the GDD and focusing on the specific period, heading in this case, eases processing demands and yields accurate results, while clearly diminishing costs. Sentinel-2 spatial resolution allows for the fine evaluation of the statistical summaries and consequently more precise estimations.

Using late-season imagery in order to take advantage of barley's earlier senescence proved to be efficient for wheat crop land classification, again, taking maximal advantage of the S2 return interval temporal frequency to identify the optimal date for the separation of like cereal crops. Our results are congruent with the Navarre regional government's official statistics, as Table 9 demonstrates. The average grain yield obtained after applying the grain yield stepwise equation from Table 7, for all of the classified wheat fields, showed similar results with official (i.e., after harvest) grain yield averages. Sentinel-2 spatial resolution allowed accurate mapping and estimation of grain yield, while other satellites would not have yielded such precise results regarding field delimitations, for instance. The grain yield estimation model was applied to the wheat classified parcels without considering genotypic differences within the region (both concerning classification and grain yield modelling issues) due to the fact that for the 2017–2018 season 80% of the fields across the whole of Navarre used either Camargo or Marcopolo [67]. Therefore, a general wheat genotypic homogeneity was supposed in the whole region.

In order to be able to use the yield estimation based on spectral data from Sentinel-2 to estimate the yield across the entire study area, it is necessary to also develop a pixel or better field-level image classification. The slightly lower Fscore of wheat classification accuracy in the Ribera Alta southern sub-region might be due to the limited surface area devoted to this crop type in comparison with the other two sites—a higher difficulty in accurately classifying wheat fields was observed with 13 misclassifications with a proportionally smaller number of ground data points for wheat at this site (Table 8). Moreover, it could be a consequence of the smaller size of the fields in this southern zone. Bare soil did present substantial misclassification with “other crops” in this southern zone; this likely



occurred due to the major presence of barley in zone VI and the earlier senescence it features, especially considering that this zone is substantially drier. Relevant misclassifications happened with wheat and bare soil in all of the three zones. Unfortunately, ground reference points in which the cultivar was specified were only obtained for wheat, the other available ground points could not be specified among the other cultivated crops, namely barley and other minor crops such as oats were not well differentiated. Despite targeting wheat crop lands, specifically classifying each of the various crop types in the region could have improved the classification results overall. Nonetheless, the lack of this exact information did not jeopardize the classification of only the wheat crop lands too much, which showed relatively high Fscores for all of the zones and was the central aim of this study.

Topographic features and rainfall together accounted for an average of 11% to 20% of the observed spatial variation in grain yield in Navarre; this percentage range is consistent with other studies dealing with these factors. Regarding topography effects on wheat at the field level in central Europe, slope and altitude have been estimated to account for an average of 5% to 42.5% variation in grain yield [68], meanwhile in North America those same topographic attributes were estimated to account for an average of 39% to 65% of the variation in grain yield [69]. Moreover, Basso et al. showed that in the Mediterranean context aspect, altitude, and slope are directly linked with wheat grain yield [70] and that rainfall distribution is a major limiting factor for wheat yields [71]. It has been observed that topographic features affect grain yield more dramatically in dry years or in areas with naturally irregular rainfalls (i.e., Mediterranean rain-fed wheat croplands) [72], which explains the relatively wide range of variability of the rainfall and topographic effects on grain yield in this study as well as in the scientific literature. Thus, an increased effect of topographic features might be expected in Spanish wheat agrosystems in a climate change scenario, and these features consequently deserve attention. In this sense, the results of this study demonstrate novelty as the relations between topography and grain yield found in the scientific literature have been, on the one hand, studied in few or single fields [73], presumably due to the difficulty in obtaining precise grain yield predictions for multiple fields from both traditional methods (i.e., time consuming field-work) and from remote sensing techniques (i.e., access restrictions or coarser resolutions to correlate with grain yields and delineate fields). On the other hand, when focusing on regional-scale topographic and climatic data, previous studies have defined suitability zones solely within the agroecological landscape [74–76]; none have completed an estimation of the effects of those factors directly on grain yield. We here took advantage of both topographic and climatic data of Navarre, Spain to explore the potential effects on grain yield in various regions, while estimating precisely the grain yield per field at a regional level using to the best of our advantage all of the Sentinel-2 improved characteristics. We believe that studying climatic and topographic impacts on grain yield at a larger scale can provide new insights that in turn will be of interest for managing croplands systemically while advancing towards improved agroecological landscape assessments.

Regarding the previously-mentioned GWR results, approximately 80%, 89%, and 80% of grain yield variability are yet to be explained in each region respectively, as the tested variables cannot explain the whole of the variability. Factors such as soil type, evapotranspiration, radiation, and fertilization or farming practices, among others, could not be evaluated with the remote sensing techniques employed in this study, and might help to explain the remaining variability. Be that as it may, the mean coefficient of regression is positive for rainfall at all three zones. This suggests a relatively strong effect of the rainfall gradient distribution on wheat grain yield in the northern zone. This might be explained with the varying precipitation patterns of the area due to the rugged topography and the numerous valleys that the area includes and that may consequently affect wheat grain yield. Regarding zone VI, the result links the drier features of the southern climate and its positive strong dependence on rainfall.

In zone II, the slope showed a negative coefficient or inverse relationship (the steeper the terrain the lower the GY), while altitude and aspect were positive, hence suggesting that southern-oriented fields at higher altitudes experienced improved performance for that season. Regarding the middle zone, the regression coefficient for the slope was also negative, while the altitude coefficient was

positive. In zone VI, the slope and altitude regression coefficients were negative, while aspect was positive. This suggests that at this site that grain yield decreases the higher and steeper the parcel is; while orientation, towards the South in this case, affects grain yield positively. The lower adjusted average of  $R^2$  results in zone V suggest only a moderate influence of the tested factors in comparison with the northern zone II and southern zone VI, agricultural areas. This might be due to the milder weather with sufficient rainfall, in comparison with the southern zone, and the fairly flatter ground in comparison with the rugged northern zone.

## 5. Conclusions

We believe that Sentinel-2's improved temporal frequency benefitted the grain yield prediction models by adapting to phenology across ecozones in the geographically diverse region of Navarre and improved the spectral separation of bread wheat from other crop types by using precisely timed images at a moment of phenological differences. The improved Sentinel-2 spatial detail was also optimally harnessed by capturing field level variability with the use of zonal statistics as model input parameters, and contributed to improving total estimates related to both yield and crop types. Furthermore, the use of field-level grain yield data for both training and validation of the wheat grain yield estimation model proved to be an efficient approach. The matching of crop type classification and field-level grain yield estimation models are scarce in the scientific literature, as often studies either focus on classification [23–26] or grain yield estimation [17–22] alone. Notwithstanding, the combination of both as presented in this study provided improved data for a more robust model (i.e., comparison to official cropland area and regional yield statistics) and for testing its applicability in actual agricultural contexts. Moreover, another novel aspect that this research addresses is the study of large-scale field-level topographic effects on grain yield, which has not been studied extensively, as single fields or reduced-scale approaches are used in the majority of studies on this topic [73].

In addition, we have applied here various techniques to take maximal advantage of several different types of openly accessible data sources for the Navarre region in northern Spain, where no such study had been previously reported. It is especially relevant, as in 2020 the CAP in the European Union will be reformulated aiming to optimize resources and advance towards an integrated regional management. The CAP accounts for around 30% of the total European budget [77], and member states are responsible of supervising the declared croplands and harvests. The results obtained and the methodology here developed can be easily used and reproduced for both Navarre's regional government and other European regions with equivalent data availability and frameworks, easing their expensive field works. We here used polygon databases from agricultural fields declared in the CAP subvention system, public topographic and environmental data, as well as European Union and European Space Agency-launched Sentinel-2 openly accessible imagery. Furthermore, we believe that in order to obtain field-level crop yields in actual agricultural contexts, coordination with local farmers is central to this endeavor.

In summary, this research successfully showed the ability of our relatively straightforward and potentially operational method for estimating wheat grain yield at field-level before harvest and determining the spatial factors that influence grain yield. Concerning grain yield estimation, two observations were deemed most pertinent: on the one hand this approach allows grain yield at approximately two months before harvesting (at the phenological stage of heading) to be estimated correctly, and on the other hand, it is applicable across at least a two-season period. The combination of our grain yield estimation model and crop type mapping also achieved accurate results with regards to official statistics. The techniques and models presented can be said to have successfully mapped wheat croplands and calculated regional grain yields. Furthermore, the observed spatial variation caused by topographic features (slope, aspect, and altitude) and rainfall could be attributed to explain on average 11% to 20% of spatial grain yield variation. Additional work is warranted in relation to the other genotype/variety, environmental, or management (GxExM) factors that may account for regional grain yield variability across local to regional scales, such as those presented here for Navarre, Spain.

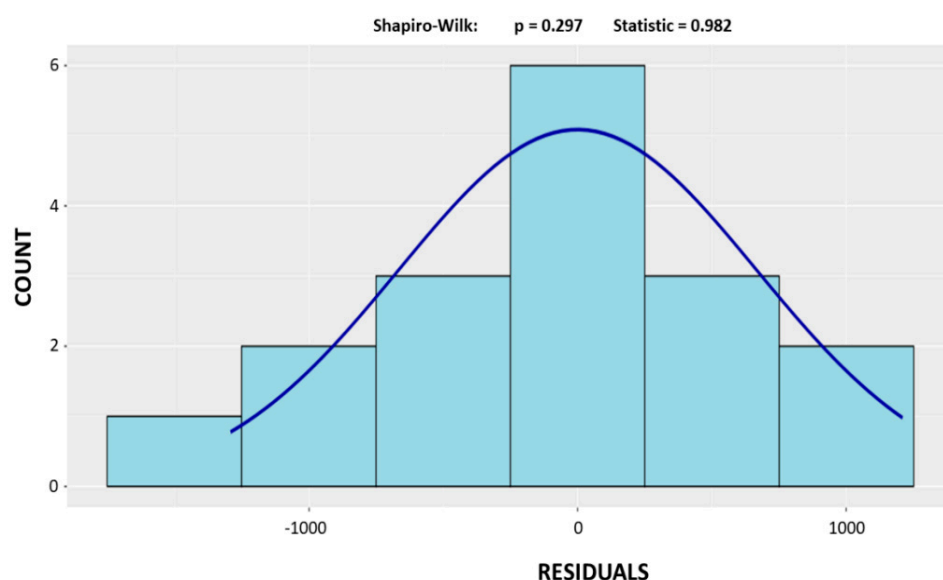
**Author Contributions:** J.L.A. and S.C.K. designed and had the original idea of the study. J.S. processed the Sentinel-2 images, mapped the crop types, developed the performance estimation models and analyzed the topographic and rainfall effects on it under the supervision of J.L.A. and S.K. J.G.-T. and I.A. managed the field work in Navarre and organized the datasets. J.S. and S.C.K. wrote the paper.

**Funding:** This study was supported by the Spanish projects PID2019-106650RB-C21 from the Ministerio de Ciencia e Innovación and the IRUEC PCIN-2017-063 from the Ministerio de Economía y Competitividad and by the project “Nuevas aplicaciones de agricultura de precisión para la monitorización del rendimiento del cultivo de trigo en Navarra” from the Government of Navarre, Spain (0011-1365-2018-000213/0011-1365-2018-000150).

**Acknowledgments:** We would like to acknowledge all the anonymous farmers working with wheat in Navarre that participated in this study. We would also like to thank the anonymous Reviewers for their comments and suggestions, which contributed significantly to the improvement of the revised version of the manuscript. J.L. Araus also acknowledges the support of Catalan Institution for Research and Advanced Studies (ICREA, Generalitat de Catalunya, Spain), through the ICREA Academia Program.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A



**Figure A1.** Histogram of the residuals of the stepwise model to assess the normality assumption.

## References

- Calatrava, C.A.; Spiegelberg, P.S.; Díaz, I.B.; Piferrer, S.J. *Avances Nacionales de Superficies y Producciones Agrícolas*; Ministerio de Agricultura, Pesca y Alimentación del Gobierno de España: Madrid, Spain, 2018.
- EU Regulation (EU) 1306/2013 of the European Parliament and of the Council on the financing, management and monitoring of the common agricultural policy. *Off. J. Eur. Union* **2013**, *347*, 549–607.
- Hunt, M.L.; Blackburn, G.A.; Carrasco, L.; Redhead, J.W.; Rowland, C.S. High resolution wheat yield mapping using Sentinel-2. *Remote Sens. Environ.* **2019**, *233*, 111410. [[CrossRef](#)]
- Toscano, P.; Castrignanò, A.; Di Gennaro, S.F.; Vonella, A.V.; Ventrella, D.; Matese, A. A Precision Agriculture Approach for Durum Wheat Yield Assessment Using Remote Sensing Data and Yield Mapping. *Agronomy* **2019**, *9*, 437. [[CrossRef](#)]
- Spitters, C. Crop Growth Models: Their Usefulness and Limitations. *Acta Hort.* **1990**, 349–368. [[CrossRef](#)]
- Burke, M.; Lobell, D.B. Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2189–2194. [[CrossRef](#)]
- Matsushita, B.; Yang, W.; Chen, J.; Onda, Y.; Qiu, G. Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Case Study in High-Density Cypress Forest. *Sensors* **2007**, *7*, 2636–2651. [[CrossRef](#)]

8. Defourny, P.; Bontemps, S.; Bellemans, N.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Nicola, L.; Rabaute, T.; et al. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sens. Environ.* **2019**, *221*, 551–568. [\[CrossRef\]](#)
9. Whitcraft, A.K.; Becker-Reshef, I.; Justice, C.O. A Framework for Defining Spatially Explicit Earth Observation Requirements for a Global Agricultural Monitoring Initiative (GEOGLAM). *Remote Sens.* **2015**, *7*, 1461–1481. [\[CrossRef\]](#)
10. Salmon, J.M.; Friedl, M.A.; Frolking, S.; Wisser, D.; Douglas, E.M. Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. *Int. J. Appl. Earth Obs. Geoin.* **2015**, *38*, 321–334. [\[CrossRef\]](#)
11. Zhang, X.; Zhang, Q. Monitoring interannual variation in global crop yield using long-term AVHRR and MODIS observations. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 191–205. [\[CrossRef\]](#)
12. Azzari, G.; Jain, M.; Lobell, D.B. Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. *Remote Sens. Environ.* **2017**, *202*, 129–141. [\[CrossRef\]](#)
13. Guan, K.; Wu, J.; Kimball, J.S.; Anderson, M.; Frolking, S.; Li, B.; Hain, C.R.; Lobell, D.B. The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields. *Remote Sens. Environ.* **2017**, *199*, 333–349. [\[CrossRef\]](#)
14. Rembold, F.; Meroni, M.; Urbano, F.; Csak, G.; Kerdiles, H.; Perez-Hoyos, A.; Lemoine, G.; Leo, O.; Negre, T. ASAP: A new global early warning system to detect anomaly hot spots of agricultural production for food security analysis. *Agric. Syst.* **2019**, *168*, 247–257. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Sakamoto, T. Incorporating environmental variables into a MODIS-based crop yield estimation method for United States corn and soybeans through the use of a random forest regression algorithm. *ISPRS J. Photogramm. Remote Sens.* **2020**, *160*, 208–228. [\[CrossRef\]](#)
16. Lambert, M.-J.; Traore, P.C.S.; Blaes, X.; Baret, P.; Defourny, P. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali’s cotton belt. *Remote Sens. Environ.* **2018**, *216*, 647–657. [\[CrossRef\]](#)
17. Skakun, S.; Vermote, E.; Franch, B.; Roger, J.-C.; Kussul, N.; Ju, J.; Masek, J. Winter Wheat Yield Assessment from Landsat 8 and Sentinel-2 Data: Incorporating Surface Reflectance, Through Phenological Fitting, into Regression Yield Models. *Remote Sens.* **2019**, *11*, 1768. [\[CrossRef\]](#)
18. Fieuzal, R.; Bustillo, V.; Collado, D.; Dedieu, G. Combined Use of Multi-Temporal Landsat-8 and Sentinel-2 Images for Wheat Yield Estimates at the Intra-Plot Spatial Scale. *Agronomy* **2020**, *10*, 327. [\[CrossRef\]](#)
19. Dedeoğlu, M.; Başayığit, L.; Yüksel, M.; Kaya, F. Assessment of the vegetation indices on Sentinel-2A images for predicting the soil productivity potential in Bursa, Turkey. *Environ. Monit. Assess.* **2019**, *192*, 16. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Kayad, A.G.; Sozzi, M.; Gatto, S.; Marinello, F.; Pirotti, F. Monitoring Within-Field Variability of Corn Yield using Sentinel-2 and Machine Learning Techniques. *Remote Sens.* **2019**, *11*, 2873. [\[CrossRef\]](#)
21. Schwalbert, R.; Amado, T.J.C.; Nieto, L.; Varela, S.; Corassa, G.M.; Hörbe, T.; Rice, C.W.; Peralta, N.R.; Ciampitti, I.A. Forecasting maize yield at field scale based on high-resolution satellite imagery. *Biosyst. Eng.* **2018**, *171*, 179–192. [\[CrossRef\]](#)
22. Habyarimana, E.; Piccard, I.; Catellani, M.; De Franceschi, P.; Dall’Agata, M. Towards Predictive Modeling of Sorghum Biomass Yields Using Fraction of Absorbed Photosynthetically Active Radiation Derived from Sentinel-2 Satellite Imagery and Supervised Machine Learning Techniques. *Agronomy* **2019**, *9*, 203. [\[CrossRef\]](#)
23. Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2018**, *204*, 509–523. [\[CrossRef\]](#)
24. Song, X.; Yang, C.; Wu, M.; Zhao, C.; Yang, G.; Hoffmann, W.C.; Huang, W. Evaluation of Sentinel-2A Satellite Imagery for Mapping Cotton Root Rot. *Remote Sens.* **2017**, *9*, 906. [\[CrossRef\]](#)
25. Wang, M.; Liu, Z.; Baig, M.H.A.; Wang, Y.; Li, Y.; Chen, Y. Mapping sugarcane in complex landscapes by integrating multi-temporal Sentinel-2 images and machine learning algorithms. *Land Use Policy* **2019**, *88*, 104190. [\[CrossRef\]](#)
26. Cai, Y.; Lin, H.; Zhang, M. Mapping paddy rice by the object-based random forest method using time series Sentinel-1/Sentinel-2 data. *Adv. Sp. Res.* **2019**, *64*, 2233–2244. [\[CrossRef\]](#)
27. Noi, P.T.; Kappas, M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors* **2017**, *18*, 18. [\[CrossRef\]](#)

28. Waldner, F.; Canto, G.S.; Defourny, P. Automated annual cropland mapping using knowledge-based temporal features. *ISPRS J. Photogramm. Remote. Sens.* **2015**, *110*, 1–13. [\[CrossRef\]](#)
29. Segarra, J.; Buchailot, M.L.; Araus, J.L.; Kefauver, S.C. Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications. *Agronomy* **2020**, *10*, 641. [\[CrossRef\]](#)
30. Da Silva, J.M.; Silva, L.L. Evaluation of the relationship between maize yield spatial and temporal variability and different topographic attributes. *Biosyst. Eng.* **2008**, *101*, 183–190. [\[CrossRef\]](#)
31. Changere, A.; Lal, R. Slope Position and Erosional Effects on Soil Properties and Corn Production on a Miamian Soil in Central Ohio. *J. Sustain. Agric.* **1997**, *11*, 5–21. [\[CrossRef\]](#)
32. Wang, K.; Huggins, D.R.; Tao, H. Rapid mapping of winter wheat yield, protein, and nitrogen uptake using remote and proximal sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *82*, 101921. [\[CrossRef\]](#)
33. Galán-Martín, Á.; Vaskan, P.; Antón, A.; Esteller, L.J.; Guillén-Gozálbez, G. Multi-objective optimization of rainfed and irrigated agricultural areas considering production and environmental criteria: A case study of wheat production in Spain. *J. Clean. Prod.* **2017**, *140*, 816–830. [\[CrossRef\]](#)
34. Balance de la Campaña de Cereal en Navarra 2017/2018. Available online: <https://uagn.es/balance-de-la-campana-de-cereal-en-navarra-2017-2018> (accessed on 14 July 2020).
35. Goñi, J. Nuevas variedades de cereal. *Navarra Agrar.* **2014**, 16–31.
36. Zadoks, J.C.; Chang, T.T.; Konzak, C.F. A decimal code for the growth stages of cereals. *Weed Res.* **1974**, *14*, 415–421. [\[CrossRef\]](#)
37. Charles, A. The determination and significance of the base temperature in a linear heat unit system. *Proc. Am. Soc. Hortic. Sci.* **1959**, *74*, 3.
38. Gitelson, A.; Viña, A.; Arkebauer, T.J.; Rundquist, D.; Keydan, G.; Leavitt, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* **2003**, *30*. [\[CrossRef\]](#)
39. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [\[CrossRef\]](#)
40. Qi, J.; Chehbouni, A.; Huete, A.; Kerr, Y.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [\[CrossRef\]](#)
41. Baret, F.; Jacquemoud, S.; Hanocq, J.F. The soil line concept in remote sensing. *Remote Sens. Rev.* **1993**, *7*, 65–82. [\[CrossRef\]](#)
42. Roujean, J.L.; Bréon, F.M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [\[CrossRef\]](#)
43. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [\[CrossRef\]](#)
44. Chen, J.M. Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Can. J. Remote Sens.* **1996**, *22*, 229–242. [\[CrossRef\]](#)
45. Rouse Jr., J.W.; Haas, R.; Schell, J.; Deering, D. Monitoring vegetation systems in the great plains with erts. *NASA Spec. Publ.* **1974**, 351.
46. Louhaichi, M.; Borman, M.M.; Johnson, D.E. Spatially Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat. *Geocarto Int.* **2001**, *16*, 65–70. [\[CrossRef\]](#)
47. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Márquez, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J.; et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2012**, *36*, 27–46. [\[CrossRef\]](#)
48. Harrell, F.E. *Regression modeling Strategies*; Springer: New York, NY, USA, 2001.
49. Meloun, M.; Militký, J.; Hill, M.; Brereton, R.G. Crucial problems in regression modelling and their solutions. *Analyst* **2002**, *127*, 433–450. [\[CrossRef\]](#)
50. Colditz, R.R. An Evaluation of Different Training Sample Allocation Schemes for Discrete and Continuous Land Cover Classification Using Decision Tree-Based Algorithms. *Remote Sens.* **2015**, *7*, 9655–9681. [\[CrossRef\]](#)
51. Ashourloo, D.; Shahrabi, H.S.; Azadbakht, M.; Aghighi, H.; Nematollahi, H.; Alimohammadi, A.; Matkan, A.A. Automatic canola mapping using time series of sentinel 2 images. *ISPRS J. Photogramm. Remote. Sens.* **2019**, *156*, 63–76. [\[CrossRef\]](#)
52. Watkins, B.; Van Niekerk, A. A comparison of object-based image analysis approaches for field boundary delineation using multi-temporal Sentinel-2 imagery. *Comput. Electron. Agric.* **2019**, *158*, 294–302. [\[CrossRef\]](#)



53. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Olmo, M.C.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote. Sens.* **2012**, *67*, 93–104. [\[CrossRef\]](#)
54. Nakaya, T.; Fotheringham, A.S.; Charlton, M.; Brunsdon, C. Semiparametric Geographically Weighted Generalised Linear Modelling in GWR 4.0. *Geocomputation* **2009**. Available online: [http://mural.maynoothuniversity.ie/4846/1/MC\\_Semiparametric.pdf](http://mural.maynoothuniversity.ie/4846/1/MC_Semiparametric.pdf) (accessed on 14 July 2020).
55. Nakaya, T.; Charlton, M.; Lewis, P.; Brunsdon, C.; Yao, J.; Fotheringham, S. *GWR4 User Manual*; Windows Applications for Geographically Weighted Regressions Modelling. GWR4 Dev. Team. 2016. Available online: [https://sgsup.asu.edu/sites/default/files/SparcFiles/gwr4manual\\_409.pdf](https://sgsup.asu.edu/sites/default/files/SparcFiles/gwr4manual_409.pdf) (accessed on 14 July 2020).
56. Fernandez-Gallego, J.A.; Kefauver, S.C.; Vatter, T.; Gutiérrez, N.A.; Nieto-Taladriz, M.; Araus, J.L. Low-cost assessment of grain yield in durum wheat using RGB images. *Eur. J. Agron.* **2019**, *105*, 146–156. [\[CrossRef\]](#)
57. Hassan, M.A.; Yang, M.; Rasheed, A.; Yang, G.; Reynolds, M.; Xia, X.; Xiao, Y.; He, Z. A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant. Sci.* **2018**, *282*, 95–103. [\[CrossRef\]](#)
58. Boken, V.K.; Shaykewich, C.F. Improving an operational wheat yield model using phenological phase-based Normalized Difference Vegetation Index. *Int. J. Remote Sens.* **2002**, *23*, 4155–4168. [\[CrossRef\]](#)
59. Song, Y.; Wang, J.; Yu, Q.; Huang, J. Using MODIS LAI Data to Monitor Spatio-Temporal Changes of Winter Wheat Phenology in Response to Climate Warming. *Remote Sens.* **2020**, *12*, 786. [\[CrossRef\]](#)
60. Rodrigues, F.; Ortiz-Monasterio, I.; Zarco-Tejada, P.; Schulthess, U.; Gerard, B. High resolution remote and proximal sensing to assess low and high yield areas in a wheat field. *Prec. Agric.* **2015**, *15*, 191–198. [\[CrossRef\]](#)
61. Wang, L.; Zhou, X.; Zhu, X.; Dong, Z.; Guo, W. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop J.* **2016**, *4*, 212–219. [\[CrossRef\]](#)
62. Bao, Y.; Gao, W.; Gao, Z. Estimation of winter wheat biomass based on remote sensing data at various spatial and spectral resolutions. *Front. Earth Sci. China* **2009**, *3*, 118–128. [\[CrossRef\]](#)
63. Zhao, C.; Bao, Y.; Cheng, M.; Huang, W.; Liu, L. Use of Landsat TM and EOS MODIS imaging technologies for estimation of winter wheat yield in the North China Plain. *Int. J. Remote Sens.* **2011**, *33*, 1029–1041. [\[CrossRef\]](#)
64. Liaqat, M.U.; Cheema, M.J.M.; Huang, W.; Mahmood, T.; Zaman, M.; Khan, M.M. Evaluation of MODIS and Landsat multiband vegetation indices used for wheat yield estimation in irrigated Indus Basin. *Comput. Electron. Agric.* **2017**, *138*, 39–47. [\[CrossRef\]](#)
65. Xie, Q.; Huang, W.; Dash, J.; Song, X.; Huang, L.; Zhao, J.; Wang, R. Evaluating the potential of vegetation indices for winter wheat LAI estimation under different fertilization and water conditions. *Adv. Space Res.* **2015**, *56*, 2365–2373. [\[CrossRef\]](#)
66. Hair, J.F.; Anderson, R.E.; Tatham, R.L.; Black, W.C. *Multivariate Data Analysis*, 3rd ed.; Macmillan: New York, NY, USA, 1995.
67. Goñi, J.; Caballero, A. Nuevas variedades de cereal. *Navarra Agrar.* **2017**, 11–21.
68. Kumhálová, J.; Moudrý, V. Topographical characteristics for precision agriculture in conditions of the Czech Republic. *Appl. Geogr.* **2014**, *50*, 90–98. [\[CrossRef\]](#)
69. Green, T.R.; Salas, J.D.; Martinez, A.; Erskine, R.H. Relating crop yield to topographic attributes using Spatial Analysis Neural Networks and regression. *Geoderma* **2007**, *139*, 23–37. [\[CrossRef\]](#)
70. Basso, B.; Cammarano, D.; Chen, D.; Cafiero, G.; Amato, M.; Bitella, G.; Rossi, R.; Basso, F. Landscape Position and Precipitation Effects on Spatial Variability of Wheat Yield and Grain Protein in Southern Italy. *J. Agron. Crop. Sci.* **2009**, *195*, 301–312. [\[CrossRef\]](#)
71. Basso, B.; Fiorentino, C.; Cammarano, D.; Cafiero, G.; Dardanelli, J. Analysis of rainfall distribution on spatial and temporal patterns of wheat yield in Mediterranean environment. *Eur. J. Agron.* **2012**, *41*, 52–65. [\[CrossRef\]](#)
72. Ferrara, R.M.; Trevisiol, P.; Acutis, M.; Rana, G.; Richter, G.; Baggaley, N. Topographic impacts on wheat yields under climate change: Two contrasted case studies in Europe. *Theor. Appl. Clim.* **2009**, *99*, 53–65. [\[CrossRef\]](#)
73. Kravchenko, A.N.; Bullock, D.G. Correlation of corn and soybean grain yield with topography and soil properties. *Agron. J.* **2000**, *92*, 75–83. [\[CrossRef\]](#)
74. Motuma, M.; Suryabagavan, K.V.; Balakrishnan, M. Land suitability analysis for wheat and sorghum crops in Wogdie District, South Wollo, Ethiopia, using geospatial tools. *Appl. Geomat.* **2016**, *8*, 57–66. [\[CrossRef\]](#)

75. Pilevar, A.R.; Matinfar, H.R.; Sohrabi, A.; Sarmadian, F. Integrated fuzzy, AHP and GIS techniques for land suitability assessment in semi-arid regions for wheat and maize farming. *Ecol. Indic.* **2020**, *110*, 105887. [[CrossRef](#)]
76. Fekadu, E.; Negese, A. GIS assisted suitability analysis for wheat and barley crops through AHP approach at Yikalo sub-watershed, Ethiopia. *Cogent Food Agric.* **2020**, *6*, 1–21. [[CrossRef](#)]
77. Comission, E. EU Budget: The Common Agricultural Policy beyond 2020. Available online: [https://ec.europa.eu/commission/presscorner/detail/en/MEMO\\_18\\_3974](https://ec.europa.eu/commission/presscorner/detail/en/MEMO_18_3974) (accessed on 14 July 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).





## **Capítol 4.** Avaluació multiescala de dades multiespectrals a nivell de sòl, dron i satèl·lit pel monitoratge del contingut de nitrogen en gra de blat

**Chapter 4.** Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content

Joel Segarra <sup>a,b</sup>, Fatima Zahra Rezzouk <sup>a,b</sup>, Nieves Aparicio <sup>c</sup>, Jon González-Torralba <sup>d</sup>, Iker Aranjuelo <sup>e</sup>, Adrian Gracia-Romero <sup>a,b</sup>, Jose Luis Araus <sup>a,b</sup>, Shawn C. Kefauver<sup>a,b,\*</sup>

<sup>a</sup>Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, Barcelona 08028, Spain

<sup>b</sup>AGROTECNIO (Center for Research in Agrotechnology), Lleida 251981, Spain

<sup>c</sup>Agro-technological Institute of Castilla y León (ITACyL), Valladolid 47071, Spain

<sup>d</sup>Grupo AN, Tajonar, 31192, Navarre, Spain

<sup>e</sup>Instituto de Agrobiotecnología (IdAB), CSIC- Gobierno de Navarra, Mutilva Baja 31192, Pamplona, Navarre, Spain

\*Correspondance

Publicat a la revista Information Processing in Agriculture



# Multiscale assessment of ground, aerial and satellite spectral data for monitoring wheat grain nitrogen content



Joel Segarra<sup>a,b</sup>, Fatima Zahra Rezzouk<sup>a,b</sup>, Nieves Aparicio<sup>c</sup>, Jon González-Torralba<sup>d</sup>, Iker Aranjuelo<sup>e</sup>, Adrian Gracia-Romero<sup>a,b</sup>, Jose Luis Araus<sup>a,b</sup>, Shawn C. Kefauver<sup>a,b,\*</sup>

<sup>a</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, Barcelona 08028, Spain

<sup>b</sup> AGROTECNIO (Center for Research in Agrotechnology), Lleida 251981, Spain

<sup>c</sup> Agro-technological Institute of Castilla y León (ITACyL), Valladolid 47071, Spain

<sup>d</sup> Grupo AN, Tajonar 31192, Navarre, Spain

<sup>e</sup> Instituto de Agrobiotecnología (IdAB), CSIC- Gobierno de Navarra, Mutilva Baja 31192, Pamplona, Navarre, Spain

## ARTICLE INFO

### Article history:

Received 3 December 2020

Received in revised form

5 May 2022

Accepted 17 May 2022

Available online 21 May 2022

### Keywords:

Wheat

Remote sensing

Sentinel-2

Grain nitrogen content

Phenotyping

## ABSTRACT

Wheat grain quality characteristics have experienced increasing attention as a central factor affecting wheat end-use products quality and human health. Nonetheless, in the last decades a reduction in grain quality has been observed. Therefore, it is central to develop efficient quality-related phenotyping tools. In this sense, one of the most relevant wheat features related to grain quality traits is grain nitrogen content, which is directly linked to grain protein content and monitorable with remote sensing approaches. Moreover, the relation between nitrogen fertilization and grain nitrogen content (protein) plays a central role in the sustainability of agriculture. Both aiming to develop efficient phenotyping tools using remote sensing instruments and to advance towards a field-level efficient and sustainable monitoring of grain nitrogen status, this paper studies the efficacy of various sensors, multispectral and visible red-green-blue (RGB), at different scales, ground and unmanned aerial vehicle (UAV), and phenological stages (anthesis and grain filling) to estimate grain nitrogen content. Linear models were calculated using vegetation indices at each sensing level, sensor type and phenological stage. Furthermore, this study explores the up-scalability of the best performing model to satellite level Sentinel-2 equivalent data. We found that models built at the phenological stage of anthesis with UAV-level multispectral cameras using red-edge bands outperformed grain nitrogen content estimation ( $R^2 = 0.42$ , RMSE = 0.18%) in comparison with those models built with RGB imagery at ground and aerial level, as well as with those built with widely used ground-level multispectral sensors. We also demonstrated the possibility to use UAV-built multispectral linear models at the satellite scale to determine grain nitrogen content effectively ( $R^2 = 0.40$ , RMSE = 0.29%) at actual wheat fields.

\* Corresponding author at: Permanent postal address at: Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, Avinyuda Diagonal, Barcelona, Spain.

E-mail address: [sckefauver@ub.edu](mailto:sckefauver@ub.edu) (S.C. Kefauver).

Peer review under responsibility of China Agricultural University.

<https://doi.org/10.1016/j.inpa.2022.05.004>

2214-3173 © 2022 China Agricultural University. Production and hosting by Elsevier B.V. on behalf of KeAi.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

One of United Nations' Sustainable Development Goals is achieving improved nutrition while promoting sustainable agriculture [1]. Wheat (*Triticum*) together with maize (*Zea mays*) and rice (*Oryza sativa*) is one of the three most-cultivated cereals globally [2]. Moreover, wheat in the form of bread, as reported by Peña [3], has historically provided more nutrients to the world population than any other single food source. Nonetheless, in the last decades cereal breeder focus on improving grain yield has resulted in a genetic erosion of quality traits [4]. Meanwhile, products made with low-quality grains have been associated with nutritious and health issues for consumers [5–7]. Protein content, together with grain hardness or starch properties, among other factors, are grain characteristics that affect the end-use quality of wheat-based products [8]. Furthermore, reductions in grain protein content (GPC) have been closely linked to reductions in grain nitrogen content (GNC) after breeding programs focused on high-yielding ideotypes [9–11]. Vogel et al. [12], after analysing 12 600 genotypes at the USDA World Wheat Collection, reported a range of GPC between 7 and 22%, corresponding to 1.2–3.7% of nitrogen in grains, which accounts for an average value of around 16% of grain protein composition [14]. Specifically in Spain, where this study takes place, Sánchez-García et al. [11] observed a decrease of 0.21% per year in GPC during the XX century and Acreche and Slafer [13] observed that GNC was reduced in modern varieties after breeding for increased grain yields.

GPC is determined by genetic and environmental factors (notably the availability of nitrogen fertilization and water [14–16]) and is directly linked to GNC, which can be monitored with non-destructive remote sensing approaches at ground, unmanned aerial vehicle (UAV) and satellite levels. GNC cannot be directly observed as there are resolution (grain size) and spectral sensing limitations (nitrogen content in plants is not discernible because nitrogen absorption features are obscured by liquid water in the crop canopy [17–19]). However, chlorophyll spectral features can be sensed within the visible and red-edge parts of the spectrum and chlorophyll-related vegetation indices can be used as proxies of crop nitrogen concentration [20], which in turn may be related to GNC. In this sense, a precise assessment of the spectral features could likely allow differentiating nitrogen-based proteins from other constituents at canopy level [21].

Accessible and effective ways to estimate GNC can provide optimized strategies for both phenotyping high-quality lines at experimental fields and to assist farmers with crops management at field level. High-quality bread wheat lines will experience an increasing demand due to current nutritious and health challenges [22,23]. Therefore, low-cost and accessible quality-oriented phenotyping is pivotal. Furthermore, developing an improved monitoring of nitrogen accumulation

in wheat grains at field-level is central for crops management. It is a key aspect regarding both the importance of environmental factors on GNC and the aim to advance towards high-quality grains within a sustainable agriculture [24].

Various studies related with GNC monitoring and remote sensing have been published in recent years. At the ground level, different authors [25–28] showed significant positive correlations between canopy spectral indices and nitrogen accumulation in wheat grains. At the UAV level, several studies have successfully correlated vegetation indices, calculated with multispectral instruments, with nitrogen accumulated in wheat grains [29–31]. At the satellite level, until recently GPC or GNC monitoring was limited due to the spatial, spectral and temporal resolutions of the orbiting satellites [32]. Several GNC related studies, showed potentialities regarding regional/field-level GNC monitoring, for instance with the Landsat satellite [33–35] or its combination with MODIS satellite [36]. Notwithstanding, Sentinel-2's recent fully operational and improved spectral, spatial and temporal resolutions have opened up opportunities [37], such as for the case of GNC monitoring [38].

Besides focusing on single sensing levels, few studies have dealt with various platforms such as satellite/ground [39] or UAV/satellite [40] spectral data to monitor GNC. The study here presented goes a step forward and works across phenotypes with remotely sensed data on three sensing scales (ground, UAV and satellite). Moreover, it discusses the most efficient sensors, either visible red–green–blue (RGB) or multispectral, and phenological stages for GNC monitoring. Vegetation indices sensitive to chlorophyll content, greenness and biomass were calculated at each level, and linear models were developed to find the most optimal GNC estimators. This research paper aims to develop accessible and reliable GNC empirical models with remote sensing data that can support cereal phenotyping and crop management. In order to achieve this, in experimental plots growing various wheat lines, RGB and multispectral sensors were deployed at both ground and UAV levels. Vegetation indices calculated from the sensed spectral reflectance data were correlated with GNC using linear models. Regarding the satellite level, the most suitable GNC estimation model was applied to actual farmer fields with equivalent Sentinel-2 data over two years. For each farmer's field, carbon isotopic composition ( $\delta^{13}\text{C}$ ) was obtained as an indicator for photosynthetic performance, in this study we aim to use this indicator as a proxy to characterize water conditions across agroclimates which may eventually affect GNC. Thus, water stress causes a decrease in stomatal conductance which subsequently will increase the  $\delta^{13}\text{C}$  of photoassimilates and then of grains. Nonetheless, carbon isotopic composition was not used in the models as we focus on spectral information in modelling. The research was structured around 4 questions to investigate which sensors, sensing levels and phenological stages were the most efficient for GNC monitoring in wheat:

- (i) Which imagery, RGB vs Multispectral, can contribute the most to estimate wheat GNC?
- (ii) Which sensing level, UAV vs ground images, is more effective to generate models for GNC monitoring?
- (iii) How relevant is phenology for GNC estimation?
- (iv) Are models built with ground or UAV images up-scalable to equivalent Sentinel-2 images?

This study concludes by testing the most optimal GNC model to estimate field-level GNC on an agro-ecosystem scale.

## 2. Materials and methods

### 2.1. Study site and field data

The study area was located at two sites. The experimental plots were located at the research station of Zamadueñas in Valladolid (Spain) and farmer field trials were located in three climatic regions within the community of Navarre in Northern Spain (Fig. 1). Valladolid has annual rainfalls ranging from 400 to 500 mm. Regarding Navarre, it incorporates several agroclimates, with rainfall averaging from 800 to 1 000 mm in the northern areas, 300 to 800 mm in the middle and around 300 mm in the southern areas.

At the Zamadueñas research station a set of 38 post green revolution and advanced wheat lines were grown in experimental plots. The design consisted of three randomized blocks with one replicate per genotype in each block, two plots (one in each corner) were sown to avoid edge effects. The cultivars had support irrigation due to 2016–2017 severe droughts in the region (a total of 129 mm of rainfall plus 60 mm of support irrigation). At the panel, 31 bread wheat

(*Triticum aestivum* L.) lines and two durum wheat (*Triticum turgidum* L. subsp. *durum* (Desf) Husn.) cultivars were utilized (Table 1). Five genotypes were no further utilized for GNC estimation in this study due to their unevenness and unsuitability for modelling. The 33 genotypes utilized are described in Table 1. The basal fertilization in the plots consisted in 300 kg/ha of 8–15–15 (NPK) in addition to top-dressing re-fertilization of 300 kg/ha of Nitrosylsulphuric acid 27% at the phenological stage of stem elongation. Plots were 6 m long and 1.5 m wide and were sown on December 2nd, 2016. At maturity, grains were sampled for further analysis. On July 20th, 2017, the plots were harvested. The experimental plots were visited at two phenological stages, in middle-May 2017, at anthesis, and in the second week of June 2017 at grain filling.

In Navarre, during 2017–2018 and 2018–2019 growing seasons 19 rain-fed farmers' fields were monitored. The agricultural fields were growing either Marcopolo (RAGT) or Camargo (Disasem) commercial lines of bread wheat. As reported by the farmers, the analysed fields received a basal fertilization with either superphosphate 45% or pig slurry, top-dressing fertilization was also applied with granulated urea at most fields. In the northern region wheat was sown around the 20th of October, in the middle region around the 28th of October, and in the southern around the 30th of October. The harvest was collected around the middle of July in the North and during the first week of July in the Middle and South.

### 2.2. Nitrogen content and carbon isotope composition in grains

Three half-square-meter representative areas were manually harvested at each farmers' field in Navarre (Spain), the grains

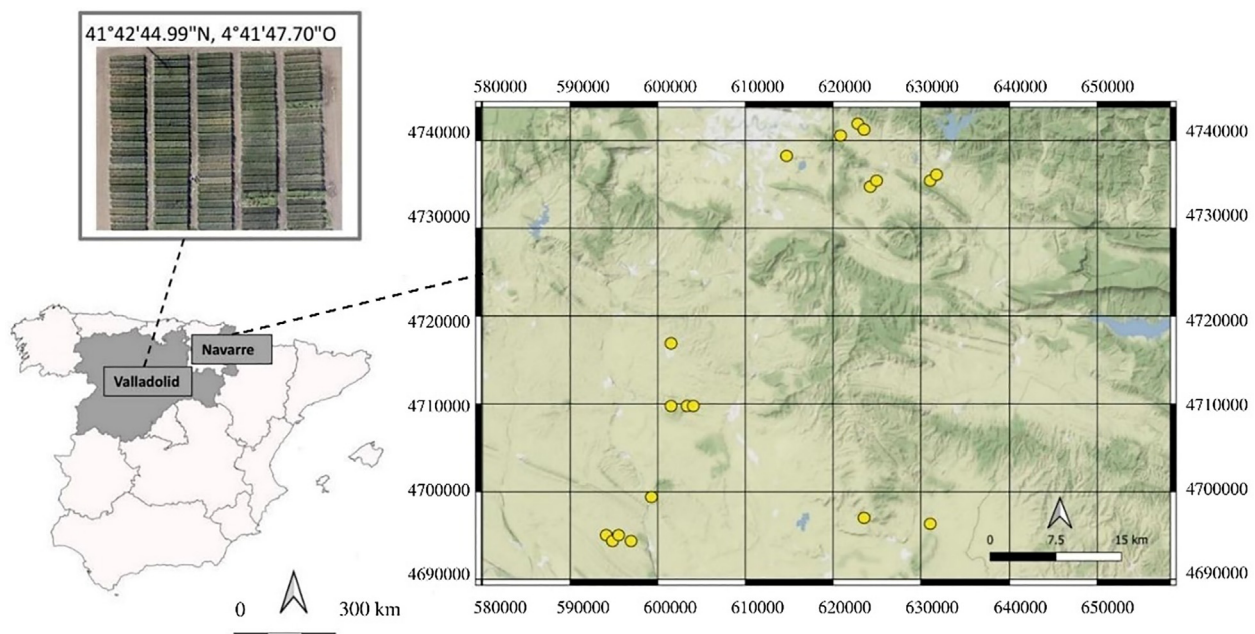


Fig. 1 – Maps of the study sites, Zamadueñas research station in Valladolid, and farmers' fields in Navarre (EPSG: 25,830 (UTM 30N, ETRS89) are indicated with a yellow dot. The map of whole Spain is also shown. The background terrain map is based on OpenStreetMap data.

**Table 1 – The utilized genotypes to develop the grain nitrogen content estimation models, the breeders of the lines are referenced.**

Name	Breeder	Species
Atomo	LG Seeds	<i>Triticum aestivum</i>
Bisanzio	MAS Seeds	
Galera	LG Seeds	
Togano	SARL Raoul Rolly	
Eneas	Dafisa	
08THES1262	Batlle	
Albertus	Saatzucht Donau	
Algoritmo	RAGT	
Bologna	Batlle	
Dolly	SARL Raoul Rolly	
Forcalli	KWS	
Ingenio	CC Benoist	
Mecano	Secobra	
Rebelde	Agri-Obtentions	
Rimbaud	Secobra	
Tribat	Batlle	
Chambo	LG Seeds	
Complice	Florimond Desprez	
Cosmic	Lemaire Deffontaines	
Ippon	Florimond Desprez	
Nemo	Secobra	
Oregrain	Florimond Desprez	
PR22R58	Pioneer Hi-Bred	
Soberbio	Caussade	
Soisson	Florimond Desprez	
MH1307	KWS	
MH1341	KWS	
MH1444	KWS	
MH1411	KWS	
Craklin	LG Seeds	
Marcopolo	RAGT	
Mimmo	PRO.SE.ME	<i>Triticum durum</i>
Credit	PRO.SE.ME	

were subsequently mixed. The 19 rain-fed farmers' fields used in this study ranged from 3.14 to 11.19 ha. At Zamadueñas research station in Valladolid (Spain), one sample was taken for each harvest of the experimental plots. The harvested mature grains were dried at 60 °C for 48 h and grinded (Mixer Mill MM 400; Retsch GmbH, Haan; Germany). After this, 1 mg of each sample was weighed in tin capsules to analyse carbon and nitrogen stable isotope signatures. Total nitrogen content was expressed as the percentage (%) of total nitrogen on dry matter basis. The carbon isotopic composition ( $\delta^{13}\text{C}$ ) was expressed following the Eq. (1):

$$\delta^{13}\text{C} (\text{‰}) = [(R_{\text{sample}}/R_{\text{standard}}) - 1] \times 1000 \quad (1)$$

where  $R_{\text{sample}}$  is the  $^{13}\text{C}/^{12}\text{C}$  ratio of the sample, while  $R_{\text{standard}}$  is the molar abundance ratio of the secondary standard calibrated against the primary standard Pee Dee Belemnite ( $\delta^{13}\text{C}$ ) [41]. Different secondary standards were used for carbon (IAEA-CH7, IAEA-CH6 and IAEA-600, and USGS 40) isotope analyses. For further information see Rezzouk et al. [42]. Isotopes and elemental analyses were performed employing an elemental analyser operating in a continuous flow mode with a mass spectrometer (Delta C IRMS; ThermoFinnigan, Bremen; Germany), at the Scientific and Technical facilities of the University of Barcelona (Centres

Científics i Tecnològics de la Universitat de Barcelona (CCiTUB)).

### 2.3. Instruments and indices

At the ground level, RGB images were taken zenithally at 80 cm above the canopy with a Sony ILCE-QX1 (Sony Europe Limited, Brookland, United Kingdom), which has a resolution of 20.1 megapixels, is equipped with a 23.2 mm × 15.4 mm sensor (type CMOS Exmor HD), and uses a 16 mm focal lens with an exposure time of 1/60 s. Also at ground level, the normalized difference vegetation index (NDVI) was measured at each plot using a GreenSeeker (Trimble, Sunnyvale, CA, USA), which is a hand held spectroradiometer with an active self-illuminated sensor in red (660 ± 10 nm) and near infrared (780 ± 15 nm) [43]. As described by Barmeier and Schmidhalter [44], the measurements were taken at a constant height of 60 cm from above the soil with a perpendicular position to the canopy.

At aerial level, RGB images were taken with a Lumix GX7 (Panasonic, Osaka, Japan) at 50 m, which has a resolution of 16 megapixels with an image sensor size of 17.3 × 13.0 mm (type Live MOS) and uses a 20 mm focal lens with an exposure time of 1/8000 s. The camera was mounted on a manually



controlled UAV (Drone Mikrokopter OktoXL 6S12, Moormerland, Germany). Aerial multispectral images were taken with a Tetracam micro-MCA at 50 m a.g.l. with the same UAV used with the RGB camera. The multispectral sensors of the Tetracam micro-MCA consists of 11 bands with  $450 \pm 40$  nm,  $550 \pm 10$  nm,  $570 \pm 10$  nm,  $670 \pm 10$  nm,  $700 \pm 10$  nm,  $720 \pm 10$  nm,  $780 \pm 10$  nm,  $840 \pm 10$  nm,  $860 \pm 10$  nm,  $900 \pm 20$  nm and  $950 \pm 40$  nm of spectral resolution (band centre and FWHM of the filter spectral response curve). Moreover, the camera possesses one sensor dedicated to calibration (Incident Light Sensor, ILS), which provides band-by-band reflectance calibration in real-time, correcting the 11 bands to reduce atmospheric and sun angle effects and provide image by image correction to at-sensor reflectance.

Aerial multispectral images were spatially aligned and radiometrically calibrated using PixelWrench 0.2 version 1.2.2.2 (Tetracam, Chatsworth, CA, USA). Aerial RGB and multispectral image mosaics were reconstructed using Agisoft Photoscan Pro (Agisoft LLC, St. Petersburg, Russia, <https://www.agisoft.com>) [45]. This software overlaps images (a minimum of 30 images with at least 80% overlapping is required) and removes UAV flight effects to produce accurate orthomosaics. Once the whole experimental field image was resampled, the plots were cropped and processed using the MosaicTool software (developed by S. C. Kefauver and others, for details, <https://integrativecropecophysiology.com/software-development/mosaictool/>, freely available at <https://gitlab.com/sckefauver/MosaicTool/>, University of Barcelona, Barcelona, Spain) integrated as a plugin for the open source image analysis platform FIJI (Fiji is Just ImageJ; <https://fiji.sc/Fiji>) [46].

RGB vegetation indices collected from both ground and aerial platforms were obtained using an updated version of the original Breedpix 2.0 software [47]. From the CIE (Commission Internationale de l'Eclairage; the International Commission on Illumination), CIELab colour space was used to calculate lightness component,  $a^*$  and  $b^*$  dimensions; and CIELuv colour space was used to calculate  $u^*$  and  $v^*$  coordinates. The  $b^*$  and  $v^*$  express the blue to yellow spectrum, while  $a^*$  and  $u^*$  represent the green to red spectrum. HIS colour space, referring to the components Hue, Saturation, and Intensity, was also used. Hue is described as the chroma traversing the visible spectrum in the form of an angle between  $0^\circ$  and  $360^\circ$ , where  $0^\circ$  and  $360^\circ$  are decrypted into red,  $60^\circ$  into yellow,  $120^\circ$  into green and  $180^\circ$  into cyan. Derived from the Hue, the indices Green Area (GA) and Greener Area (GGA) were described as the fraction area presented by green pixels in the image, and which Hue ranges from  $60^\circ$  to  $180^\circ$  (GA) and from  $80^\circ$  to  $180^\circ$  (GGA). While GA gives a broader perception to canopy greenness, GGA excludes yellowish green pixels [47,48]. The following RGB indices were also calculated: the triangular green index (TGI) [49], which estimates chlorophyll concentration in canopies; the normalized green–red difference index (NGRDI) [50], which compares the differences between the green and red bands in a calculation similar to NDVI but with less marked differences and less signal saturation; and the crop senescence index (CSI) [51,52], which combines GA and GGA to provide an index of canopies greenness. The instruments and indices calculated are summarized in Table 2.

Regarding multispectral indices, 14 vegetation indices (VIs) were calculated (summarized in Table 2): CCCI (Canopy Chlorophyll Content Index) [53], CCI (Chlorophyll/Carotenoid Index) [54], CI Green (Chlorophyll Index Green) [55], CI Red Edge (Chlorophyll Index Red Edge) [55], CVI (Chlorophyll Vegetation Index) [56], EVI (Enhanced Vegetation Index) [57], MCARI (Modified Chlorophyll Absorption Ratio Index) [90], NDVI (Normalized Difference Vegetation Index) [59], OSAVI (Optimized Soil Adjusted Vegetation Index) [60], RDVI (Renormalized Difference Vegetation Index) [61], SAVI (Soil Adjusted Vegetation Index) [62], TCARI (Transformed Chlorophyll Absorption Index) [58], TCARI/OSAVI [58] and TCI (Triangular Chlorophyll Index) [63]. These indices were chosen due to its sensitiveness to chlorophyll content and capacities to monitor plants' nitrogen related features. At aerial level, VIs were calculated on Microsoft Excel (2010) with the obtained reflectance from the Tetracam micro-MCA mounted on a UAV.

#### 2.4. Grain nitrogen content estimation






In order to see the GNC gradient in the genotypes grown at the experimental site, a post hoc analyses was calculated on R studio [64] ("agricolae"). After, stepwise multilinear regressions and simple regressions between GNC and VIs were calculated in order to obtain empirical models for GNC estimation. The VIs means were extracted for each field. Preliminary, simple R-Pearson regressions of all the VIs against GNC were calculated with the function "cor.test" on R studio [64] for each phenological stage (anthesis and grain filling), sensing level (ground and aerial) and sensor type (RGB and multispectral). In order to represent these results understandably, they were plotted in a heatmap, the obtained R-Pearson values were added in a matrix and plotted on R studio [64] ("ggplot2").

At aerial and ground level, stepwise multilinear models were calculated with RGB and multispectral indices against GNC with the exception of ground multispectral data, which only had one parameter (NDVI) and thus a simple regression was calculated. In all cases a 70% of each dataset (RGB ground, RGB aerial, multispectral ground and multispectral aerial) was used for training the model while 30% was used for validation. In order to find the most suitable combination of parameters (indices) for the stepwise multilinear regression the library "caret" was used on Rstudio. With the functions "trainControl" and "train" a 10-fold cross validation was used to estimate the average predicting error, root mean square (RMSE), and select the most optimal model. With the best models, the variance inflator factor (VIF) was measured as it is an efficient method to assess collinearity [65–67], which strongly limits stepwise selection method. With the 30% of the dataset that was not used for the training the models were tested.

In order to ease the reading of the models' validation results, a bar chart was plotted on Rstudio ("ggplot2") to show R square, RMSE and the significance of the tested linear models regarding the two phenological stages (anthesis and grain filling), type of sensor (RGB and multispectral images), and sensing level (ground and aerial level). The original graphs of validation are shown in the Appendix B.



**Table 2 – Summary of sensing levels, sensors type, instruments used, and vegetation indices calculated for each case. The spatial resolution of the instruments is also indicated. The reference (Ref.) is specified too.**

Level	Spectral data	Sensor	Index	Description	Ref.
Ground	Multispectral	<b>GreenSeeker</b> (60 cm spatial resolution) 	<b>NDVI</b>	(B840-B670)/ (B840 + B670)	[59]
	RGB	<b>Sony ILCE-QX1</b> (0.01 mm/pixel) 	<b>Intensity</b> <b>Saturation</b> <b>Lightness</b> <b>a*</b> <b>b*</b> <b>u*</b> <b>v*</b>	HIS color space HIS color space CIElab color space CIElab color space CIElab color space CIEluc color space CIEluc color space	[48] [48] [48] [48] [48] [48] [48]
UAV	RGB	<b>Lumix GX7</b> (5 cm/pixel) 	<b>Hue</b> <b>GA</b> <b>GGA</b> <b>CSI</b> <b>NGRDI</b>  <b>TGI</b>	HIS color space % of pixels in hue from 60° to 180° % of pixels in hue from 80° to 180° $((GA-GGA)/GA) \times 100$ $(GREEN-RED)/(GREEN + RED)$ $((GREEN-0.39) \times (RED-0.61)) \times BLUE$	[48] [48,49] [48,49] [51,52] [50] [49]
	Multispectral	<b>Tetracam micro-MCA</b> (5 cm/pixel) 	<b>CCCI</b> <b>CCI</b> <b>CI Green</b> <b>CI red edge</b> <b>CVI</b> <b>EVI</b>  <b>MCARI</b>	$((B840B700)/(B840 + B700))/((B840-B670)/(B840 + B670))$ $(B550-B570)/(B550 + B670)$ $(B840/B560)-1$ $(B840/B700)-1$ $(B840/B560) \times (B670/B560)$ $(2.5 \times (B840-B670))/(B840+(6 \times B670)-(7.5 \times B450) + 1)$ $(B700-B670) - 0.2 \times (B700 - B550) \times (B700/B670)$	[53] [54] [55] [55] [56] [57] [90]
Satellite	Multispectral	<b>Sentinel-2 a + b</b> (10 and 20 m/pixel) 	<b>NDVI</b> <b>OSAVI</b> <b>RDVI</b> <b>SAVI</b> <b>TCARI</b>  <b>TCARI/OSAVI</b> <b>TCI</b>	$(B840-B670)/(B840 + B670)$ $(1.16) \times (B780 - B670)/(B780 + B670 + 0.16)$ $(B840-B670)/\sqrt{(B840 + B670)}$ $(B840-B670)/(B840 + B670)$ $3 \times (B700 - B670) - 0.2 \times (B700-B550) \times (B700/B670)$ $TCARI/OSAVI$ $1.2 \times (B700 - B550) - 1.5 \times (B670 - B550) \times \sqrt{(B700/B670)}$	[59] [60] [61] [62] [58] [58] [63]

### 2.5. Sentinel-2 imagery

The specific dates of Sentinel-2 images for 2018 and 2019 were selected as those closest to the most suitable phe-

nological stage for GNC monitoring (Table 3). At regional scale the phenological stage was estimated with growing degree days (GDD) calculations as described by Arnold [68] (Eq. (2)).

**Table 3 – Anthesis phenological stage estimated with GDD and Sentinel-2 closest date during 2018 and 2019 seasons.**

	North	Middle	South
Anthesis (GDD)	10-05-18 / 08-05-19	17-05-18 / 15-05-19	26-04-18 / 29-04-19
Sentinel-2 (dates)	09-05-18 / 09-05-19	19-05-18 / 14-05-19	24-04-18 / 29-04-19

$$GDD = \sum_{jun}^{Nov} \frac{T_{max} + T_{min}}{2} - T_{base} \quad (2)$$

where GDD is the growing degree days and  $\sum_{jun}^{Nov}$  indicates the sum throughout the season, November to June, of daily maximum and minimum temperatures ( $T_{max}$  and  $T_{min}$ ) divided by 2, minus the base temperature ( $T_{base}$ ), which in this case was considered 0 °C.

This data is used following Segarra et al. [69] study for the case of Navarre. The average temperature data, in order to estimate the crop phenological stages through the calculation of GDD in each zone (North, Middle and South), was obtained from 30 openly accessible regional government meteorological stations (<https://www.meteo.navarra.es/estaciones>), GDD and phenology for the wheat lines was obtained from the genotypes information as described by Goñi [70], corresponding to 1105 accumulated GDD for anthesis as the average for both (Camargo and Marcopolo wheat lines).

In Table 3 the estimated phenological date and the closest Sentinel-2 image is detailed. The Sentinel-2 images were downloaded from Copernicus Open Access Hub (<https://sci-hub.copernicus.eu/>) as a 2A product (Bottom of Atmosphere reflectance images) and with a maximum cloud cover threshold of 30%.

Sentinel-2 bands are shown in Table 4, with these bands the following most suitable VIs from the model, CI red-edge [56] (Eq.(3)), TCARI [63] (Eq.(4)), and EVI [58] (Eq.(5)), were calculated on ArcGIS Pro 2.3.0 with the corresponding Sentinel-2 bands:

$$CI_{red-edge} = \frac{(B8)}{(B5)} - 1 \quad (3)$$

$$TCARI = 3 \times (B5 - B4) - 0.2 \times (B5 - B3) \times \frac{B5}{B4} \quad (4)$$

$$EVI = 2.5 \times \frac{(B8 - B4)}{(B8 + (6 \times B4) - (7.5 \times B2) + 1)} \quad (5)$$

### 3. Results

The GNC among the utilized wheat lines ranged between 2.07% and 2.74% (Fig. 2). The highest GNC was in the group

formed by the genotypes Dolly, Mimmo, Galera, 08THES1262, Albertus and Rebelde, while the group with the lowest GNC was formed by PR22R58, Nemo, MH1411, Cosmic, Rimbaud, Oregain, Marcopolo, Mecano, Chambo, Complice, Soberbio, Craklin and Togano (see the appendix Table A1 for the post hoc test and the averaged GNC per wheat line). The genotypes with the highest GNC ranged between 2.74% and 2.60%, while the lowest ranged from 2.26% to 2.07%.

In comparison with grain filling, which presented lower values of R-Pearson, the descriptive results presented in the heat map (Fig. 3) show that the highest R-Pearson values of the simple regressions between GNC and multispectral and RGB indices were observed at the phenological stage of anthesis. At anthesis and aerial level, the multispectral vegetation indices showed a positive direct relationship with GNC. Regarding the RGB indices at aerial level, positive correlations were observed. However,  $a^*$  and  $u^*$  indices showed to be more negative with greener vegetation due to the nature of the colour space conversion. Hence, an inverse relationship was observed in this case. Counterintuitively, this was not observed at RGB ground level.

As shown in Fig. 4, the less negative  $\delta^{13}C$  values were found, in general terms, in fields located in the southern and middle areas of Navarre, while the most negative values were mainly found in fields in the North.

In Fig. 5, the obtained GNC estimation accuracy ( $R^2$ ) of the validated models at all the sensing levels and sensor type is shown. At anthesis, the multispectral sensors at UAV level were the ones with the most precise GNC estimation capacities with a significant  $R^2$  of 0.42 (Fig. 6). At grain filling, the UAV-multispectral sensor effectiveness at estimating GNC was reduced to a  $R^2$  of 0.29. At ground level, the validation of the model built with RGB data to estimate GNC was significant at the phenological stage of anthesis with a  $R^2$  of 0.19. None of the validation of the rest of models (UAV RGB and ground multispectral at anthesis; and ground RGB, UAV RGB and ground multispectral at grain filling) showed significant results.

The most suitable GNC estimation model obtained from multispectral imagery at UAV level at the phenological stage of anthesis is summarized in Table 5. The most optimal veg-

**Table 4 – Bands and resolution of Sentinel-2 Multispectral Instrument (MSI).**

MSI Band	Spatial Resolution(m)	Central Wavelength(nm)
B1: Coastal Aerosol	60	443
B2: Blue	10	490
B3: Green	10	560
B4: Red	10	665
B5: Red-Edge	20	705
B6: Red-Edge	20	740
B7: Red-Edge	20	783
B8: NIR	10	842
B8A: Vegetation RE	20	865
B9: Water Vapour	60	945
B10: SWIR Cirrus	60	1375
B11: SWIR	20	1610
B12: SWIR	20	2190

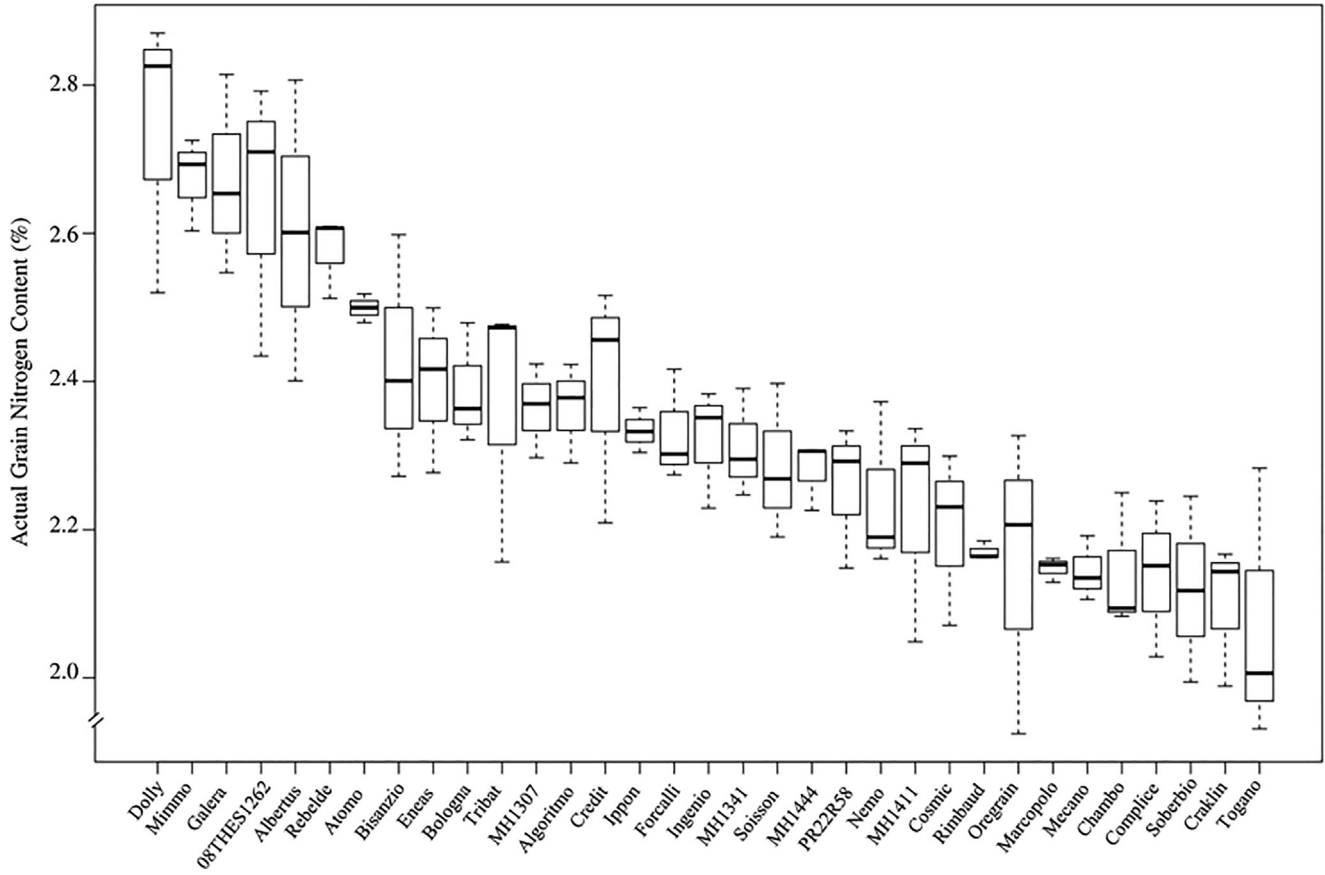


Fig. 2 – Posthoc analyses of the utilized wheat lines at the experimental station of Zamadueñas, the significance letters are shown in the appendix, Table A1.

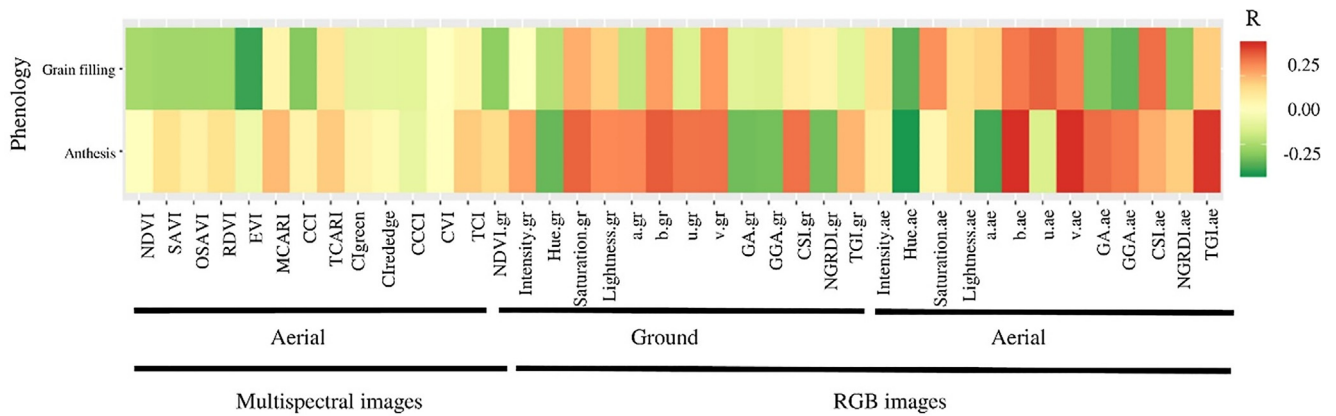


Fig. 3 – Heatmap descriptive representation of R Person correlations between GNC and various multispectral and RGB indices at aerial and ground levels. In the colour scale, red represents the highest positive R value, while green is the lowest negative R value.

etation indices as estimators were TCARI, Cl red edge and EVI. The parameters were significant and had VIF values of 9.0 for TCARI, 9.7 for Cl red edge and 1.6 for EVI. Regarding the goodness of the model the best performing model showed absence of collinearity as all VIF values (9 for TCARI, 9.7 for Cl red edge and 1.6 for EVI) were closer to the collinearity threshold value of 10 [71].

The model built with UAV data that best suited GNC estimation was applied at satellite level with Sentinel-2 imagery. Over the two years the validation of the model with satellite data was significant. The estimation of GNC with satellite data reached an  $R^2 = 0.40$  and  $RMSE = 0.29\%$  (Fig. 7), alike the one obtained with the UAV multispectral imagery ( $R^2 = 0.42$  and  $RMSE = 0.18\%$ , Fig. 6).

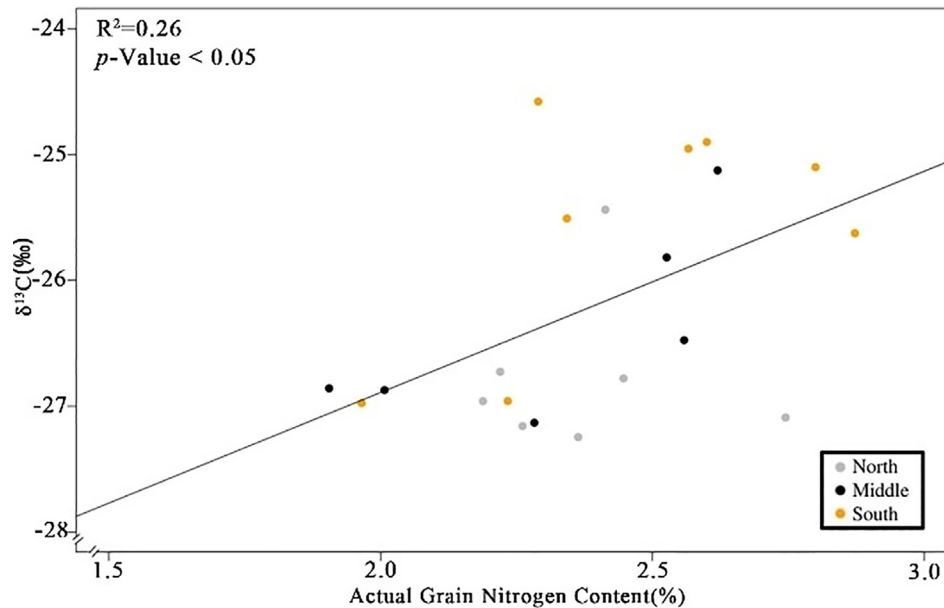


Fig. 4 – Carbon isotopic composition ( $\delta^{13}\text{C}$ , ‰) against grain nitrogen content (%) to assess C isotopic discrimination and characterize hydric conditions across the agroclimates (in North, Middle and South of Navarre, Spain) where the farmers' fields are located.

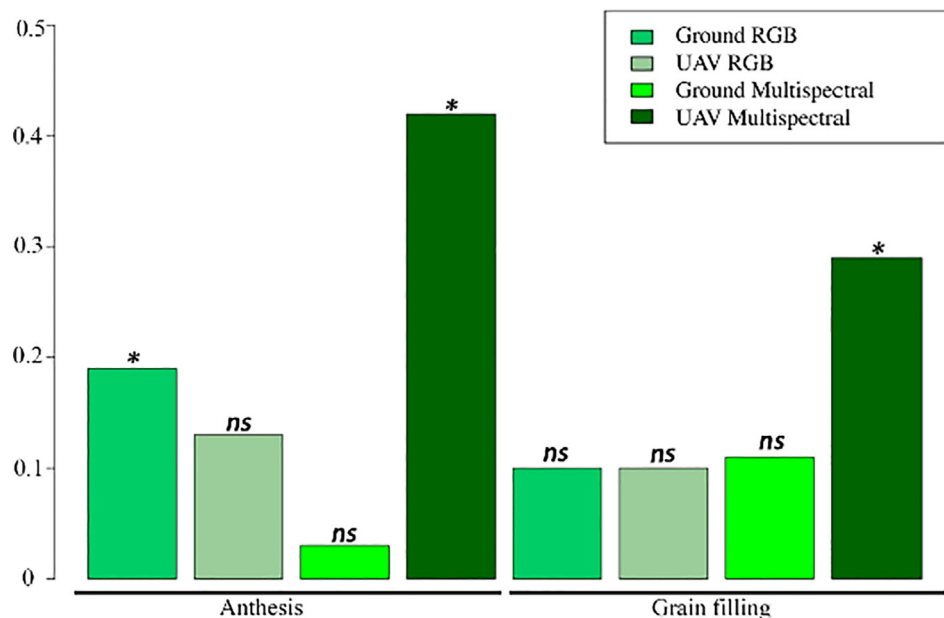
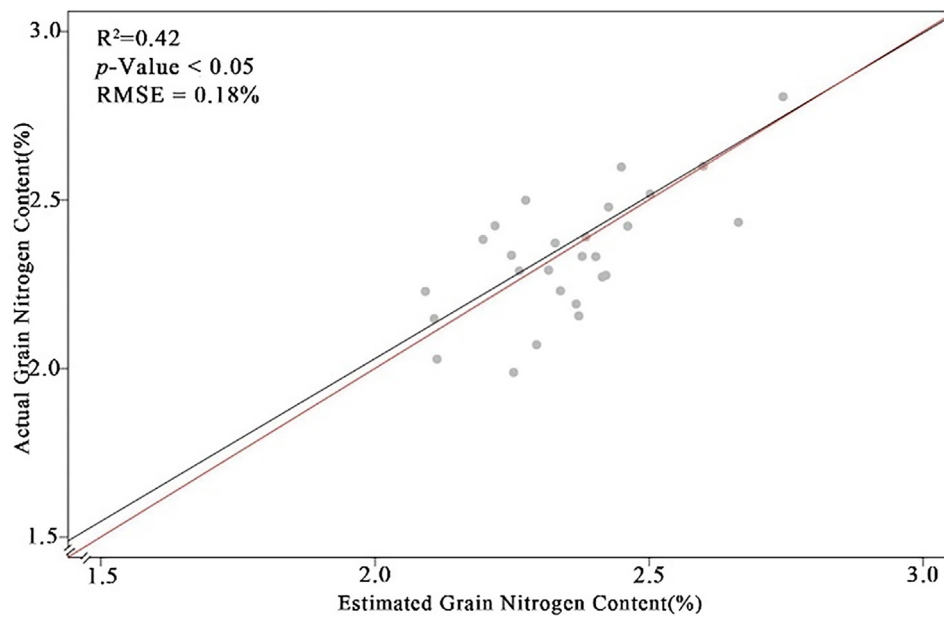


Fig. 5 – Summary of the obtained  $R^2$  at validation for GNC estimation with RGB and multispectral sensors at ground and aerial levels, the \* indicates a significance of p-value < 0.05 while ns refers to no significance. The two phenological stages studied are shown (anthesis and grain filling).

#### 4. Discussion

The models developed in this study using UAV multispectral sensors outperformed RGB models (at ground and aerial level) as well as models built with a widely used multispectral ground-level sensor (GreenSeeker). Coming back to the questions one (which imagery, RGB vs Multispectral, can contribute the most to estimate wheat GNC?) and two (which sensing level, UAV vs ground images, is more effective to

generate models for GNC monitoring?) the following was observed. When validating the GNC estimation models with multispectral data at UAV level the best results showed (Fig. 5) an  $R^2$  of 0.42 and RMSE of 0.18% at anthesis, which aligns to the range reported by Zhao et al. [38]. No significant correlations were observed with models built with RGB data at ground and aerial level, except for the low yielding ( $R^2 = 0.19$ , RMSE = 0.19%) of the ground RGB model at anthesis (Fig. 5). High spectral resolution in the visible photosyntheti-



**Fig. 6 – Validation of the stepwise multilinear regression models for UAV multispectral data at anthesis as the most suitable model for GNC estimation, 30% of the data set of the experimental plots were used for validation. The equation of the model is shown in Table 5. The  $R^2$ ,  $p$ -value and root mean square (RMSE) are shown. The red line indicates 1:1.**

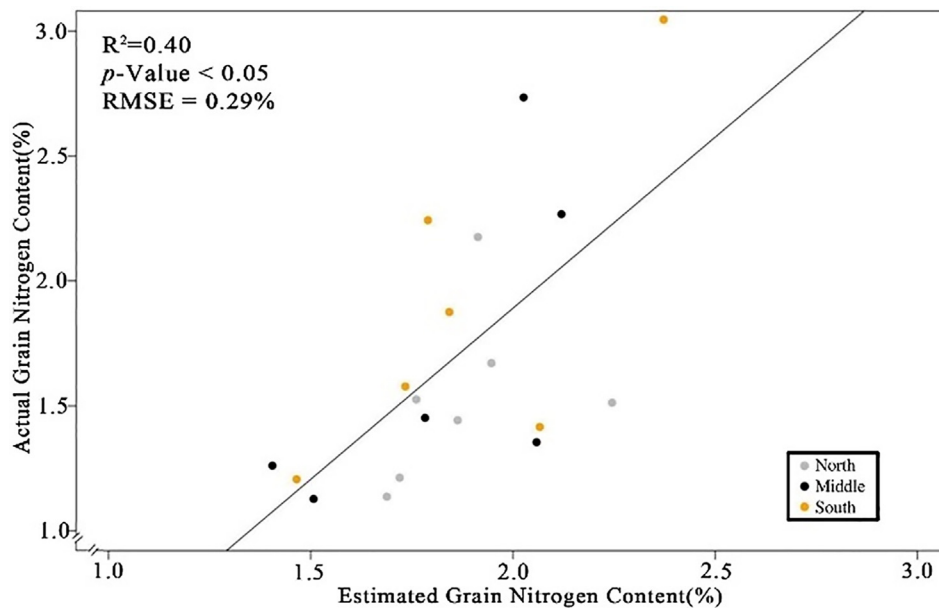
**Table 5 – Summary of stepwise multilinear results for the UAV multispectral model at anthesis, it shows the equation, the selected vegetation indices and the VIF.**

Parameter	Estimate	$p$	VIF
Intercept	−1.09	<0.01	
TCARI	31.50	<0.01	9.0
Cl <sub>rededge</sub>	0.60	<0.01	9.7
EVI	−1.74	<0.01	1.6
<b>Model Summary</b>			
Grain N content = $-1.09 + 31.5 \times \text{TCARI} + 0.60 \times \text{Cl}_{\text{rededge}} - 1.74 \times \text{EVI}$			

cally active wavelength regions (RGB) has been shown to be capable of assessing crop nitrogen content [72], nonetheless we observed that it is not assessable when using broad band reflectance measurements provided by commercial RGB cameras, regardless of their high spatial resolution and reliable colour calibration [73]. Hence, we can argue that in this study the model built with multispectral data at UAV level contributed the most to GNC estimation. The multispectral instrument used at the ground level only measures red and near-infrared reflectance (computed as NDVI) and showed no significant correlations. This might be related, first to the lack of relationship between a biomass sensitive index such as NDVI and GNC, which is related to N content in the plant. Moreover, the saturation of NDVI, which is reached in high vegetation density covers such as anthesis, does not contribute to generate an indicator of the plant that could be likely linked to GNC. In addition, Greenseeker has spatial resolution limitations in comparison with UAV sensors (Table 2). These results demonstrate the superior capacities of multiple bands and multispectral information itself, as well as spatial resolution, for the case of GNC estimation.

Due to the nature of the data, empirical models were chosen in contrast with other approaches. Radiative transfer

models, for instance, offer an improved understanding of the structure and physiological features of the crop canopy although are very dependent on large datasets and computationally demanding algorithms. Empirical models, such as VIs regressions, sometimes experience limitations, as are more affected by site and sensor conditions, but are more versatile when having reduced datasets and widely used and understood. In this case, 14 VIs were calculated at UAV level using Tetracam multispectral data while only one multispectral index was calculated at ground level (NDVI). Although the Greenseeker is a low-cost and widely used ground multispectral instrument for phenotyping [46,74,75], the spectral resolution it features is insufficient for GNC estimation. On the other hand, this contrasts with the generally more costly and logistically demanding (i.e. a minimum 5 m long-poles or UAV) multispectral cameras. Nonetheless, in the last years multispectral cameras prices have decreased and their use in phenotyping and research-oriented proposes is increasing. Hence, nowadays multispectral cameras mounted on UAVs are common, as the reviewed scientific literature from the last years shows [76], and may further contribute to phenotyping quality related traits.



**Fig. 7 – Application of the most suitable stepwise multilinear model with Sentinel-2 data at actual farmers' fields in Navarre, Spain, across three agroclimatic regions (North, Middle and South). The  $R^2$ ,  $p$ -value and root mean square (RMSE) are shown.**

The UAV remotely sensed multispectral data in this study has several advantages over the RGB data. We observed that among the three selected vegetation indices (TCARI, CI red-edge and EVI) as the most suitable for the best performing stepwise multilinear model (Table 5), EVI uses RGB and NIR bands, while TCARI and  $CI_{red-edge}$ , besides those, use a band in the red-edge part of the spectrum ( $700 \pm 10$  nm). The advantages of indices using red-edge bands for determining nitrogen up-taken into grains have been observed in various studies [28,77,78], as the study here presented also corroborates. EVI is an index sensitive to biomass, which has nonetheless been successfully used in plant nitrogen concentration estimations [79], also in the estimation of grain protein content in wheat [32]. EVI has improved features to prevent atmospheric and soil influences, which make it more responsive to canopy variations in comparison with other indices [80]. In this sense, the results suggest that the amount of spectral information available from multi-band multispectral sensors increases GNC estimation effectiveness in comparison with the RGB visible three-band data, red-green-blue, as it expands the sensing spectrum further to red (>670 nm). The red-edge part of the spectrum is highly sensitive to chlorophyll changes and therefore can be used as a nitrogen proxy [20,81]. TCARI has been successfully used at estimating grain protein content in cereals [82], as well as  $CI_{red-edge}$  at UAV level in the case of wheat with a reported  $R^2$  of 0.50 [30]. We can argue that the spectral resolution (bands in the red-edge) is especially relevant when the sensing aim, nitrogen in grains, is indirectly discernible and highly dependent on chlorophyll-related spectral information.

Regarding the third question (how relevant is phenology for GNC estimation?) we observed that phenology was central for GNC estimation (Figs. 3 and 5). At anthesis, the UAV multispectral GNC estimation model reached an  $R^2$  of 0.42, while

at grain filling it decreased to 0.29. This aligns with the conceptualization that wheat nitrogen accumulation in grains start happening at grain filling. Namely, plants nitrogen concentration starts descending at the grain filling stage. Thus, the results here presented suggest that VIs sensitive to nitrogen concentration at anthesis can be an indicator for the final GNC in wheat, as was also observed by Dupont and Altenbach [83]. In wheat, between 60 and 95% of the nitrogen is remobilized from leaves and stems into grains, being the most important sources of GNC [84,85] and thereby of protein in wheat grains. The results of the study show that it is possible to monitor GNC through these vegetation indices sensitive to chlorophyll if plant nitrogen content is sensed around anthesis. This phenological stage corresponds to the maximum nitrogen accumulation [86,87] in wheat plants, before it is remobilized into grains.

At the agro-ecosystem scale, the higher (i.e. less negative) carbon isotope composition ( $\delta^{13}C$ ) values of mature grains in the Middle and Southern parts of Navarre compared with the Northern area (Fig. 4) suggests the occurrence of more drought stress in the central and southern areas of Navarra. This is in agreement with their lower rainfall, higher temperatures, and lower grain yields observed for the fields in this study in Middle and Southern Navarre [69]. In general, grain nitrogen content (GNC) was higher in the Middle and Southern areas than in the Northern area, probably related with a smaller grain size, caused by drought stress during grain filling, in the former areas. In that sense  $\delta^{13}C$  and GNC across sites and areas were positively correlated (Fig. 4) and served to explain agroclimatic features linked to wheat physiology and GNC.

This study demonstrates that the stepwise multilinear model built at the UAV level could be up-scaled with Sentinel-2 data and cover field-level GNC estimations. Hence,



regarding the fourth question (Are models built with ground or UAV images up-scalable to equivalent Sentinel-2 images?), we observed that the applied model over two seasons reached an  $R^2$  coefficient of 0.40 and RMSE of 0.29% (Fig. 7), which is in the same range as the one obtained with the validation at the experimental plots with the UAV multispectral data ( $R^2$  of 0.42 and RMSE of 0.18%, Fig. 6). The RMSE is slightly higher when the model is applied at the field-level; nonetheless, these results align with other studies that concomitantly used UAV obtained spectral data to build models and apply them with Sentinel-2 data [88,89], as well for the case of GNC estimation [38] where several N sensitive stepwise models showed a RMSE range of 0.19–0.53%. Building models with actual field-level to regional data requires vast field work, covering large spatial extensions. On the other hand, building the models at smaller experimental sites eases the process. The study here discussed confirms that the simultaneous use of UAV multispectral cameras and Sentinel-2 data with equivalent spectral bands allows GNC estimation.

This study has been able to define the most optimal approaches for estimating GNC with applications at both experimental plots and agricultural fields. The novel comparison presented here among different sensing platforms and phenological stages, sets the basis for an improved quality-oriented phenotyping and crop monitoring. Often, quality traits have been under-considered in breeding programs [9,10] and the development of models able to separate high GNC phenotypes from others could help identify the most suitable lines in wheat populations exhibiting genetic variability. In this sense, the genotypic range used for this study was 2.07% to 2.74% of GNC, it could be broader (1.2–3.7%) when adding a bigger diversity of genotypes, from modern varieties to landraces. Therefore, the model is specific for a range of GNC in the instance of our study. Despite the achieved  $R^2$  is inferior to 0.5, the significance of the models and the results at validation proves its effectiveness in phenotyping at experimental plots and monitoring GNC at actual agricultural fields, while selecting the subset of genotypes with the desired range values of N content. Currently, these models cannot determine GNC with potential applications in precise certification of grain protein content before harvest. For this aim destructive measurements would still be needed. Nonetheless, the obtained models can orient farmers on grain quality after a quick UAV flight before harvest and could also potentially be used in smart farming to spot low and high GNC concentration areas in a field. This information could contribute to optimize the use of resources by prioritizing increased top-dressing fertilization in poor GNC zones, and thus helping to advance to agricultural stability in terms of GNC. In addition, at the agroecosystem level with the use of Sentinel-2 imagery, the current reliability of the models might provide support for developing regional GNC estimation maps which can contribute to understand environmental and soil features affecting GNC, as well as to help farmers and institutions to manage croplands sustainably. The Sentinel-2 spatial resolution makes their direct application as a phenotyping platform in breeding programs unsuitable, nonetheless for the application in crop phenotyping the model developed in this study could be implemented with satellite platforms with higher resolutions.

## 5. Conclusion

This study demonstrates that UAV multispectral models can provide accurate ( $R^2 = 0.42$  and RMSE = 0.18%) GNC estimations across phenotypes when the sensing period happens at anthesis. Moreover, it corroborates the central use of red-edge bands ( $700 \pm 10$  nm) to calculate effective VIs for GNC estimation. The application of UAV-built multispectral models with equivalent Sentinel-2 real data effectively estimated GNC over two seasons at field-level ( $R^2 = 0.40$  and RMSE = 0.29%). RGB true colour images and the widely used ground multispectral phenotyping sensor GreenSeeker were not effective in GNC estimation. The models here developed show the potentialities of GNC estimation across phenotypes with UAV multispectral data as it allows to differentiate groups of genotypes regarding their GNC. Moreover, we also demonstrated the potentialities of the application of these models to accessible satellite data (Sentinel-2) for an improved GNC monitoring at field-scale. It takes advantage to the novel Sentinel-2 features with its red-edge bands and high revisit time. To our knowledge, only one other study using UAV-level data (hyperspectral in this case) to build models and monitor GNC at field-level with equivalent Sentinel-2 data has been published [38]. The study here presented shows two main novelties. On the one hand, it demonstrates that multispectral UAV-mounted cameras can produce effective GNC estimation models when having equivalent bands in the red-edge though additional improvements in this field may be reached by including hyperspectral information in GNC estimation in the future. On the other hand, it is the first paper that together reviews remote sensing techniques at various levels to determine the most optimal strategy for both GNC phenotyping and field-level monitoring.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This study was supported by the projects PID2019-106650RB-C21 (Ministerio de Ciencia e Innovación, MICINN, Spain) and 0011-1365-2018-000213/0011-1365-2018-000150 (Government of Navarre, Spain). J.S. is recipient of a FPI doctoral fellowship (Grant: PRE2020-091907) from MICINN, Spain. J.L.A. acknowledges support from ICREA Academia, Generalitat de Catalunya, Spain. S.C.K. is supported by the Ramon y Cajal RYC-2019-027818-I research fellowship from MICINN, Spain. The processing of satellite images was supported by the European Cooperation in Science and Technology Action CA17134 SENSECO.

## Appendix A. Wheat lines differences in GNC

See Table A1.



**Table A1 – Average GNC in % for each genotype at the research site of Zamadueñas. The superscript letter indicates differences among the line after post hoc test.**

Genotypes	N % in grain	Genotypes	N <sub>o</sub> in grain
Dolly	2.74 <sup>a</sup>	MH1341	2.31 <sup>efghijk</sup>
Mimmo	2.74 <sup>ab</sup>	Soisson	2.28 <sup>efghijkl</sup>
Galera	2.67 <sup>ab</sup>	MH1444	2.26 <sup>efghijklm</sup>
08THES1262	2.64 <sup>ab</sup>	PR22R58	2.24 <sup>efghijklm</sup>
Albertus	2.60 <sup>abc</sup>	Nemo	2.25 <sup>efghijklm</sup>
Rebelde	2.56 <sup>abcd</sup>	MH1411	2.23 <sup>efghijklm</sup>
Atomo	2.50 <sup>bcde</sup>	Cosmic	2.20 <sup>efghijklm</sup>
Bisanzio	2.42 <sup>cdef</sup>	Rimbaud	2.17 <sup>efghijklm</sup>
Eneas	2.40 <sup>defg</sup>	Oregrain	2.15 <sup>efghijklm</sup>
Bologna	2.39 <sup>defg</sup>	Marcopolo	2.15 <sup>efghijklm</sup>
Tribat	2.37 <sup>defg j</sup>	Mecano	2.14 <sup>efghijklm</sup>
MH1307	2.36 <sup>defgh</sup>	Chambo	2.14 <sup>efghijklm</sup>
Algoritmo	2.33 <sup>defgh</sup>	Complice	2.14 <sup>efghijklm</sup>
Credit	2.33 <sup>defgh</sup>	Soberbio	2.12 <sup>efghijklm</sup>
Ippon	2.33 <sup>efghij</sup>	Craklin	2.10 <sup>efghijklm</sup>
Forcalli	2.33 <sup>efghij</sup>	Togano	2.07 <sup>m</sup>
Ingenio	2.32 <sup>efghij</sup>		

## Appendix B. Results and graphs of the calculated models.

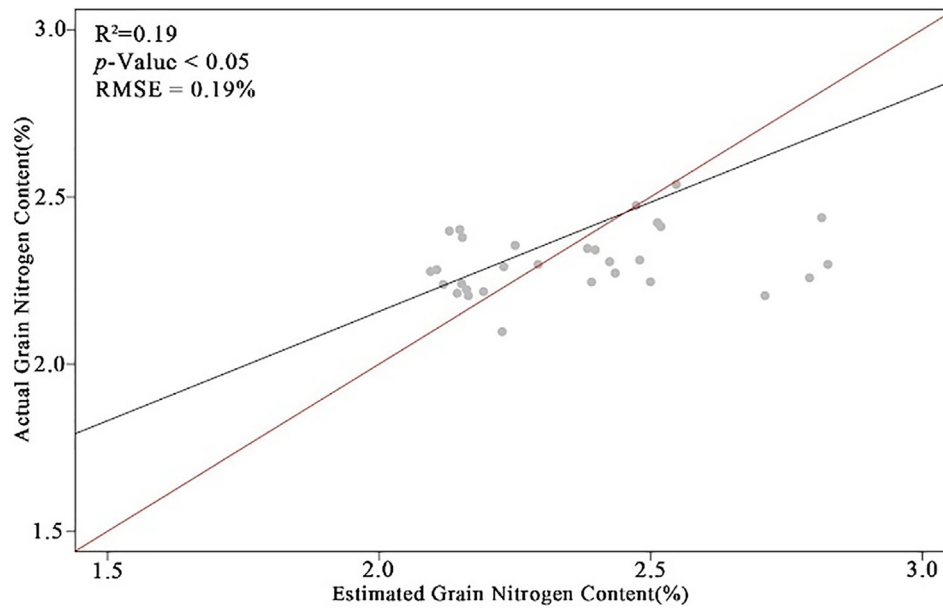
See [Tables B1 and B2](#) and [Figs. B1–B7](#).

**Table B1 – Summary of stepwise multilinear results for the ground RGB models at anthesis It shows the equation, the selected vegetation indices and the VIF.**

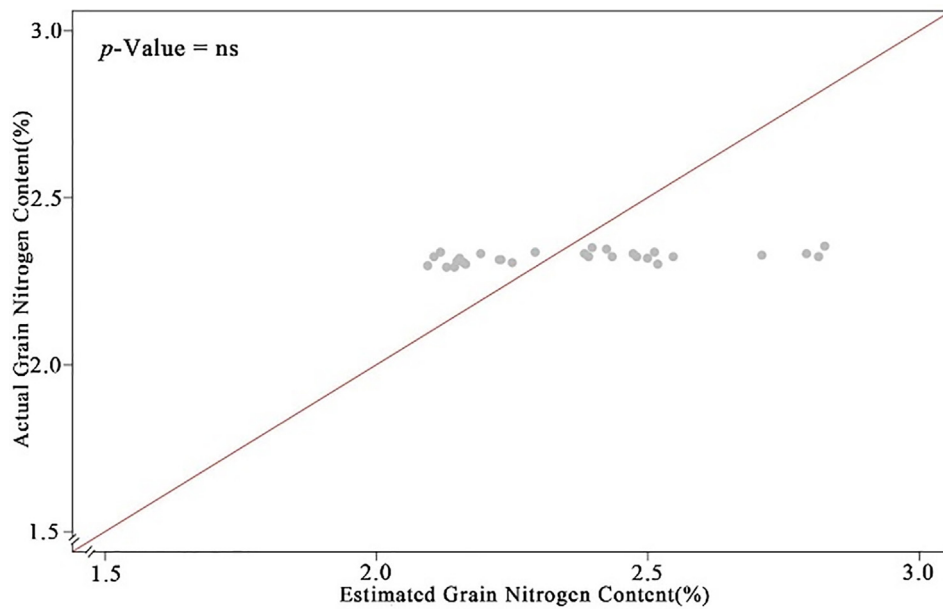
Parameter	Estimate	p	VIF
Intercept	3.38	<0.01	
Hue	−0.01	<0.01	5.1
v	0.01	<0.01	1.3
GA	−0.6	<0.01	5.2
<b>Model Summary</b>			
Grain N content = $3.38 - 0.01 \times \text{Hue} + 0.01 \times v - 0.60 \times \text{GA}$			

**Table B2 – Summary of stepwise multilinear results for the UAV multispectral models at grain filling. It shows the equation, the selected vegetation indices and the VIF.**

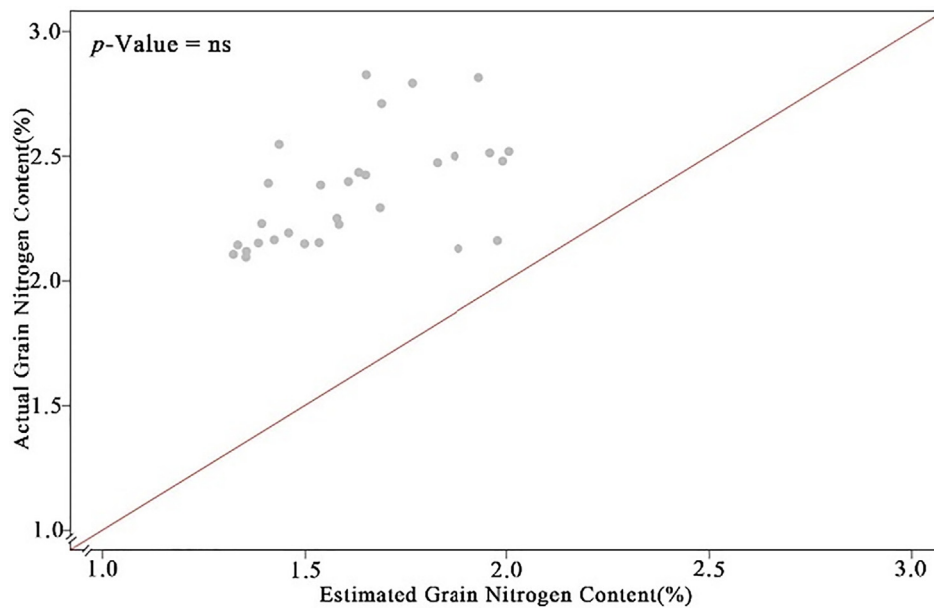
Parameter	Estimate	p	VIF
Intercept	3.3	<0.01	
OSAVI	−6.1	ns	69.0
RDVI	8.1	<0.01	80.8
EVI	−1.3	<0.01	2.8
<b>Model Summary</b>			
Grain N content = $3.3 - 6.1 \times \text{OSAVI} + 8.1 \times \text{RDVI} - 1.3 \times \text{EVI}$			



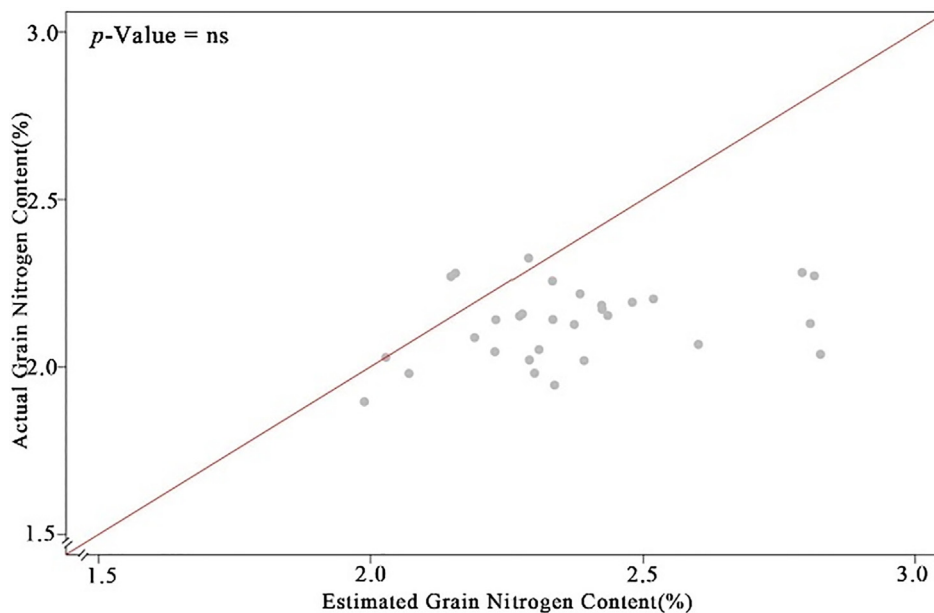
**Fig. B1** – Validation of the stepwise multilinear regression models for ground RGB data at anthesis, 30% of the data set of the experimental plots were used for validation. The equation of the model is shown in [Table B.1](#). The  $R^2$ ,  $p$ -value and root mean square (RMSE) are shown. The red line indicates 1:1.



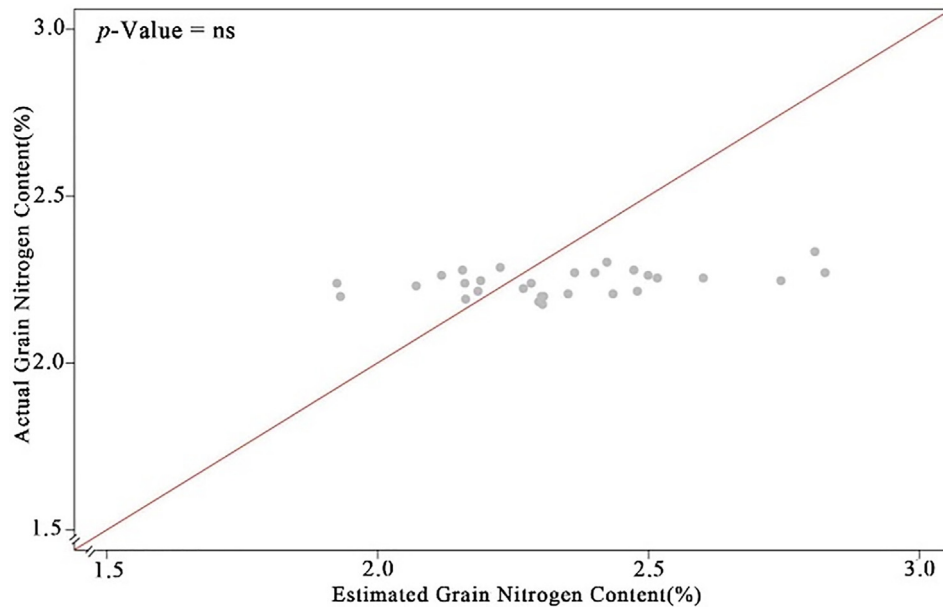
**Fig. B2** – Validation of the stepwise multilinear regression models for ground multispectral data at anthesis, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.



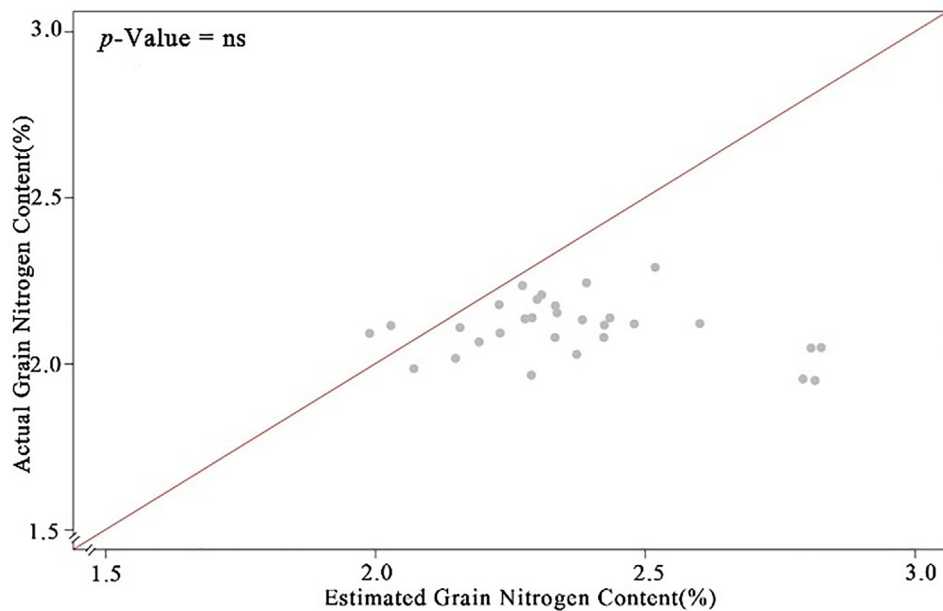
**Fig. B3** – Validation of the stepwise multilinear regression models for UAV RGB data at anthesis, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.



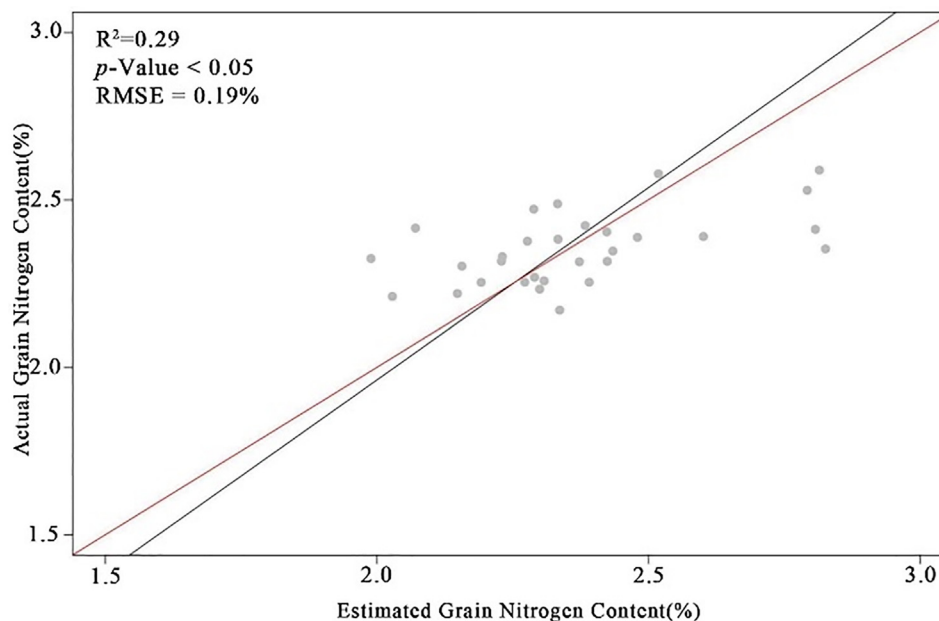
**Fig. B4** – Validation of the stepwise multilinear regression models for ground RGB data at grain filling, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.



**Fig. B5** – Validation of the stepwise multilinear regression models for ground multispectral data at grain filling, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.



**Fig. B6** – Validation of the stepwise multilinear regression models for UAV RGB data at grain filling, 30% of the data set of the experimental plots were used for validation. The red line indicates 1:1.



**Fig. B7 – Validation of the stepwise multilinear regression models for UAV multispectral data at grain filling, 30% of the data set of the experimental plots were used for validation. The equation of the model is shown in Table B.1. The  $R^2$ ,  $p$ -value and root mean square (RMSE) are shown. The red line indicates 1:1.**

## Appendix C. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.inpa.2022.05.004>.

## REFERENCES

- [1] Assembly G. Transforming our world: the 2030 Agenda for Sustainable Development. New York, United States: United Nations; 2015.
- [2] Awika JM. Major cereal grains production and use around the world. In: Awika JM, Piironen V, Bean S, editors. *Advances in Cereal Science: Implications to Food Processing and Health Promotion*. USA: ACS Symp Ser; 2011. p. 1–13.
- [3] Peña RJ. Wheat for Bread and Other Foods. Rome, Italy: Food and Agriculture Organization of the United Nations; 2002.
- [4] Esquinas-Alcázar J. Protecting crop genetic diversity for food security: Political, ethical and technical challenges. *Nat Rev Genet* 2005;6(12):946–53.
- [5] Peña RJ. Current and future trends of wheat quality needs. In: Buck HT, Nisi JE, Salomón N, editors. *Developments in plant breeding wheat production in stressed environments*. Dordrecht: Springer Netherlands; 2007. p. 411–24.
- [6] Dinu M, Whittaker A, Pagliai G, Benedettelli S, Sofi F. Ancient wheat species and human health: Biochemical and clinical implications. *J Nutr Biochem* 2018;52:1–9.
- [7] Hazard B, Trafford K, Lovegrove A, Griffiths S, Uauy C, Shewry P. Strategies to improve wheat for human health. *Nat Food* 2020;1(8):475–80.
- [8] Regina A, Guzman C. Starch and starch-associated proteins: impact on wheat grain quality. In: Igrejas G, Ikeda TM, Guzmán C, editors. *Wheat quality for improving processing and human health*. Berlin, Heidelberg, Germany: Springer; 2020. p. 39–73.
- [9] Zhu J, Huang S, Khan K, Brien L. Relationship of protein quantity, quality and dough properties with Chinese steamed bread quality. *J Cereal Sci* 2001;33(2):205–12.
- [10] Pronin D, Börner A, Weber H, Wheat SKA. (*Triticum aestivum* L.) breeding from 1891 to 2010 contributed to increasing yield and glutenin contents but decreasing protein and gliadin contents. *J Agric Food Chem* 2020;68(46):13247–56.
- [11] Sanchez-Garcia M, Álvaro F, Peremarti A, Martín-Sánchez JA, Royo C. Changes in bread-making quality attributes of bread wheat varieties cultivated in Spain during the 20th century. *Eur J Agron* 2015;63:79–88.
- [12] Vogel KP, Johnson VA, Mattern PJ. Protein and Lysine Content of Grain, Endosperm, and Bran of Wheats from the USDA World Wheat Collection. *Crop Sci* 1976;16(5):655–60.
- [13] Acreche MM, Slafer GA. Variation of grain nitrogen content in relation with grain yield in old and modern Spanish wheats grown under a wide range of agronomic conditions in a Mediterranean region. *J Agric Sci* 2009;147(6):657–67.
- [14] Lloveras J, Lopez A, Ferran J, Espachs S, Solsona J. Bread-Making Wheat and Soil Nitrate as Affected by Nitrogen Fertilization in Irrigated Mediterranean Conditions. *Agron J* 2001;93(6):1183–90.
- [15] Daniel C, Triboí E. Changes in wheat protein aggregation during grain development: Effects of temperatures and water stress. *Eur J Agron* 2002;16(1):1–12.
- [16] Xue LH, Cao WX, Yang LZ. Predicting Grain Yield and Protein Content in Winter Wheat at Different N Supply Levels Using Canopy Reflectance Spectra. *Pedosphere* 2007;17(5):646–53.
- [17] Serrano L, Peñuelas J, Ustin SL. Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data. *Remote Sens Environ* 2002;81(2):355–64.
- [18] Kokaly RF. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sens Environ* 2001;75(2):153–61.
- [19] Fourty T, Baret F, Jacquemoud S, Schmuck G, Verdebout J. Leaf optical properties with explicit description of its biochemical

- composition: Direct and inverse problems. *Remote Sens Environ* 1996;56(2):104–17.
- [20] Guerif M, Houles V, Baret F. Remote sensing and detection of nitrogen status in crops. Application to precise nitrogen fertilization. In: 4th int symp intell inf technol agric ISIITA. Beijin, China; 2007. p. 19–28.
- [21] Féret JB, Berger K, de Boissieu F, Malenovsky Z. PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents. *Remote Sens Environ* 2021;252:112173–88.
- [22] Ibrahim A, Csúr-Varga A, Jolánkai M, Mansour H, Hamed A. Monitoring some quality attributes of different wheat varieties by infrared technology. *Agric Eng Int CIGR J* 2018;20(1):201–10.
- [23] Cheng A. Review: Shaping a sustainable food future by rediscovering long-forgotten ancient grains. *Plant Sci* 2018;269:136–42.
- [24] Arzani A, Ashraf M. Cultivated Ancient Wheats (*Triticum* spp.): A Potential Source of Health-Beneficial Food Products. *Compr Rev Food Sci Food Saf* 2017;16(3):477–88.
- [25] Moges SM, Raun WR, Mullen RW, Freeman KW, Johnson GV, Solie JB. Evaluation of green, red, and near infrared bands for predicting winter wheat biomass, nitrogen uptake, and final grain yield. *J Plant Nutr* 2004;27(8):1431–41.
- [26] Wright DL, Rasmussen VP, Ramsey RD, Baker DJ, Ellsworth JW. Canopy reflectance estimation of wheat nitrogen content for grain protein management. *GIScience Remote Sens* 2004;41(4):287–300.
- [27] Saleem MF, Ma BL, Voldeng H, Wang TC. Nitrogen nutrition on leaf chlorophyll, canopy reflectance, grain protein and grain yield of wheat varieties with contrasting grain protein concentration. *J Plant Nutr* 2010;33(11):1681–95.
- [28] Prey L, Schmidhalter U. Temporal and spectral optimization of vegetation indices for estimating grain nitrogen uptake and late-seasonal nitrogen traits in wheat. *Sensors* 2019;19(21):1–27.
- [29] Geipel J, Link J, Wirwahn JA, Claupein W. A programmable aerial multispectral camera system for in-season crop biomass and nitrogen content estimation. *Agric* 2016;6(1):1–19.
- [30] Zhou X, Kono Y, Win A, Matsui T, Tanaka TST. Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches. *Plant Prod Sci* 2020;24(2):137–51.
- [31] Quemada M, Pancorbo JL, Alonso-Ayuso M, Gabriel JL, López-Herrera J, Pérez-Martín E. Vegetation indices from remote sensing imagery as proxies for yield and grain N in wheat. In: Precision Agriculture'19 Wageningen Academic Publishers; Montpellier, France; 2019. p. 323–30.
- [32] Tan C, Zhou X, Zhang P, Wang Z, Wang D, Guo W, et al. Predicting grain protein content of field-grown winter wheat with satellite images and partial least square algorithm. *PLoS ONE* 2020;15(3):e0228500.
- [33] Zhao C, Liu L, Wang J, Huang W, Song X, Li C. Predicting grain protein content of winter wheat using remote sensing data based on nitrogen status and water stress. *Int J Appl Earth Obs Geoinf* 2005;7(1):1–9.
- [34] Li CJ, Wang JH, Wang Q, Wang DC, Song XY, Wang Y, et al. Estimating Wheat Grain Protein Content Using Multi-Temporal Remote Sensing Data Based on Partial Least Squares Regression. *J Integr Agric* 2012;11(9):1445–52.
- [35] Song X, Wang J, Yang G, Feng H. Winter Wheat Cropland Grain Protein Content Evaluation through Remote Sensing. *Intell Autom Soft Comput* 2014;20(4):599–609.
- [36] Wang L, Tian Y, Yao X, Zhu Y, Cao W. Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. *F Crop Res* 2014;164:178–88.
- [37] Segarra J, Buchaillet ML, Araus JL, Kefauver SC. Remote Sensing for Precision Agriculture : Sentinel-2 Improved Features and Applications. *Agronomy* 2020;10(5):1–18.
- [38] Zhao H, Song X, Yang G, Li Z, Zhang D, Feng H. Monitoring of nitrogen and grain protein content in winter wheat based on Sentinel-2A data. *Remote Sens* 2019;11(14):1–25.
- [39] Reyniers M, Vrindts E. Measuring wheat nitrogen status from space and ground-based platform. *Int J Remote Sens* 2006;27(3):549–67.
- [40] Du M, Noguchi N, Ito A, Shibuya Y. Correlation analysis of vegetation indices based on multi-temporal satellite images and unmanned aerial vehicle images with wheat protein contents. *Eng Agric Environ Food* 2019;10:1–12.
- [41] Farquhar GD, Ehleringer JR, Hubick KT. Carbon isotope discrimination and photosynthesis. *Annu Rev Plant Biol* 1989;40(1):503–37.
- [42] Rezzouk FZ, Gracia-Romero A, Kefauver SC, Gutiérrez NA, Aranjuelo I, Serret MD, et al. Remote sensing techniques and stable isotopes as phenotyping tools to assess wheat yield performance: Effects of growing temperature and vernalization. *Plant Sci* 2020;295:110281–97.
- [43] Crain J, Ortiz-Monasterio I, Raun B. Evaluation of a reduced cost active NDVI sensor for crop nutrient management. *J Sensors* 2012;2012:1–10.
- [44] Barmeier G, Schmidhalter U. High-throughput phenotyping of wheat and barley plants grown in single or few rows in small plots using active and passive spectral proximal sensing. *Sensors* 2016;16(11):1–14.
- [45] Bendig J, Bolten A, Bennertz S, Broscheit J, Eichfuss S, Bareth G. Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. *Remote Sens* 2014;6(11):10395–412.
- [46] Gracia-Romero A, Kefauver SC, Fernandez-Gallego JA, Vergara-Díaz O, Nieto-Taladriz MT, Araus JL. UAV and ground image-based phenotyping: A proof of concept with durum wheat. *Remote Sens* 2019;11(10):1244–69.
- [47] Casadesús J, Kaya Y, Bort J, Nachit MM, Araus JL, Amor S, et al. Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. *Ann Appl Biol* 2007;150(2):227–36.
- [48] Gracia-Romero A, Kefauver SC, Vergara-Díaz O, Zaman-Allah MA, Prasanna BM, Cairns JE, et al. Comparative performance of ground vs. Aerially assessed rgb and multispectral indices for early-growth evaluation of maize performance under phosphorus fertilization. *Front Plant Sci* 2017;8:1–13.
- [49] Hunt ER, Doraiswamy PC, McMurtrey JE, Daughtry CST, Perry EM, Akhmedov B. A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. *Int J Appl Earth Obs Geoinf* 2013;21:103–12.
- [50] Hunt ER, Cavigelli M, Daughtry CST, McMurtrey JE, Walthall CL. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis Agric* 2005;6(4):359–78.
- [51] Zaman-Allah M, Vergara O, Araus JL, Tarekegne A, Magorokosho C, Zarco-Tejada PJ, et al. Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* 2015;11(1).
- [52] Vergara-Díaz O, Kefauver SC, Elazab A, Nieto-Taladriz MT, Araus JL. Grain yield losses in yellow-rusted durum wheat estimated using digital and conventional parameters under field conditions. *Crop J* 2015;3(3):200–10.
- [53] Barnes EM, Clarke TR, Richards SE, Colaizzi PD, Haberland J, Kostrzewski M, et al. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In: Proc fifth int conf precis agric Bloomington, USA; 2000. p. 1–15.
- [54] Gamon JA, Huemmrich KF, Wong CYS, Ensminger I, Garrity S, Hollinger DY, et al. A remotely sensed pigment index reveals



- photosynthetic phenology in evergreen conifers. *PNAS* 2016;113(46):13087–92.
- [55] Gitelson AA, Viña A, Arkebauer TJ, Rundquist DC, Keydan G, Leavitt B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys Res Lett* 2003;30(5):1248–54.
- [56] Vincini M, Frazzi E, D'Alessio P. A broad-band leaf chlorophyll vegetation index at the canopy scale. *Precis Agric* 2008;9(5):303–19.
- [57] Huete A, Didan K, Miura T, Rodriguez EP, Gao X, Ferreira LG. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens Environ* 2002;83(1–2):195–213.
- [58] Haboudane D, Miller JR, Tremblay N, Zarco-Tejada PJ, Dextraze L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens Environ* 2002;81(2–3):416–26.
- [59] Rouse Jr. JW, Haas R, Schell J, Deering D. Monitoring vegetation systems in the great plains with erts. In: *Third earth res tech satellite-1 symposium NASA*. Washington D.C., USA; 1974. p. 309–17.
- [60] Baret F, Jacquemoud S, Hanocq JF. The soil line concept in remote sensing. *Remote Sens Rev* 1993;7(1):65–82.
- [61] Roujean JL, Breon FM. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens Environ* 1995;51(3):375–84.
- [62] Huete A. A soil-adjusted vegetation index (SAVI). *Remote Sens Environ* 1988;25(3):295–309.
- [63] Hunt Jr E, Daughtry CS, Eitel JU, Long DS. Remote sensing leaf chlorophyll content using a visible band index. *Agron J* 2011;103(4):1090–9.
- [64] R Core Team. A language and environment for statistical computing: <https://www.r-project.org/index.html>. 2020.
- [65] Harrell FE. Regression modeling strategies. USA: Springer; 2001. p. 78–9.
- [66] Meloun M, Militký J, Hill M, Brereton RG. Crucial problems in regression modelling and their solutions. *Analyst* 2002;127(4):433–50.
- [67] Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography (Cop)* 2013;36(1):27–46.
- [68] Arnold C. The determination and significance of the base temperature in a linear heat unit system. *Proc Am Soc Hortic Sci* 1959;74(1):430–45.
- [69] Segarra J, González-Torralba J, Aranjuelo I, Araus JL, Kefauver SC. Estimating wheat grain yield using Sentinel-2 imagery and exploring topographic features and rainfall effects on wheat performance in Navarre. Spain *Remote Sens* 2020;12(14):1–24.
- [70] Goñi J, Caballero A Nuevas variedades de cereal. *Navarra Agrar* 2017;242:11–21.
- [71] Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate Data Analysis. USA: Macmillan; 1995. p. 133–43.
- [72] Hansen PM, Schjoerring JK. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens Environ* 2003;86(4):542–53.
- [73] Buchaillot ML, Gracia-Romero A, Vergara-Díaz O, Zaman-Allah MA, Tarekegne A, Cairns JE, et al. Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques. *Sensors* 2019;19(8):1815.
- [74] Gracia-Romero A, Vergara-Díaz O, Thierfelder C, Cairns JE, Kefauver SC, Araus JL. Phenotyping conservation agriculture management effects on ground and aerial remote sensing assessments of maize hybrids performance in Zimbabwe. *Remote Sens* 2018;10(2):1–21.
- [75] Ali AM, Ibrahim SM, Bijay-Singh. Wheat grain yield and nitrogen uptake prediction using atLeaf and GreenSeeker portable optical sensors at jointing growth stage. *Inf Process Agric* 2020;7(3):375–83.
- [76] Deng L, Mao Z, Li X, Hu Z, Duan F, Yan Y. UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras. *ISPRS J Photogramm Remote Sens* 2018;146:124–36.
- [77] Guo B-B, Qi S-L, Heng Y-R, Duan J-Z, Zhang H-Y, Wu Y-P, et al. Remotely assessing leaf N uptake in winter wheat based on canopy hyperspectral red-edge absorption. *Eur J Agron* 2017;82:113–24.
- [78] Prey L, Schmidhalter U. Sensitivity of vegetation indices for estimating vegetative N status in winter wheat. *Sensors* 2019;19(17):1–17.
- [79] Zheng H, Li W, Jiang J, Liu Y, Cheng T, Tian Y, et al. A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. *Remote Sens* 2018;10(12):2026.
- [80] Huete AR, Didan K, Van Leeuwen W. MODIS Vegetation Index (MOD13) algorithm theoretical basis document. USA: University of Arizona; 1999.
- [81] Segarra J, Buchaillot ML, Stefani U, Araus JL, Kefauver SC. Sentinel-2 responsiveness to fertilization gradients in wheat at field level in Córdoba Province, Argentina. In: *Mediterr middle-east geosci remote sens symp M2GARSS 2020*. Tunis, Tunisia; 2020. p. 322–5.
- [82] Pettersson CG, Eckersten H. Prediction of grain protein in spring malting barley grown in northern Europe. *Eur J Agron* 2007;27(2–4):205–14.
- [83] Dupont FM, Altenbach SB. Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. *J Cereal Sci* 2003;38(2):133–46.
- [84] Van Sanford DA, MacKown CT. Variation in nitrogen use efficiency among soft red winter wheat genotypes. *Theor Appl Genet* 1986;72(2):158–63.
- [85] Palta JA, Fillery IR. N application enhances remobilisation and reduces losses of pre-anthesis N in wheat grown on an Duplex soil. *Aust J Agric Res* 1995;46(3):519–31.
- [86] Papakosta DK, Gagianas AA. Nitrogen and Dry Matter Accumulation, Remobilization, and Losses for Mediterranean Wheat during Grain Filling. *Agron J* 1991;83(5):864–70.
- [87] Kong L, Xie Y, Hu L, Feng B, Li S. Remobilization of vegetative nitrogen to developing grain in wheat (*Triticum aestivum* L.). *F Crop Res* 2016;196:134–44.
- [88] Revill A, Florence A, MacArthur A, Hoad SP, Rees RM, Williams M. The value of Sentinel-2 spectral bands for the assessment of winter wheat growth and development. *Remote Sens* 2019;11(17):1–18.
- [89] Revill A, Florence A, MacArthur A, Hoad S, Rees R, Williams M. Quantifying uncertainty and bridging the scaling gap in the retrieval of leaf area index by coupling sentinel-2 and UAV observations. *Remote Sens* 2020;12(11):1843–61.
- [90] Daughtry CS, Walthall CL, Kim MS, Brown de Colstoun E, McMurtrey III JE. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment* 2000(74):229–39.





## **Capítol 5. L'agricultura i l'observació de la Terra: estimació de la variació del rendiment del blat a nivell de parcel·la**

### **Chapter 5. Farming and Earth Observation: Sentinel-2 data to estimate within-field wheat grain yield**

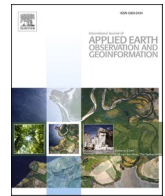
Joel Segarra <sup>a,b</sup>, Jose Luis Araus <sup>a,b,\*</sup>, Shawn C. Kefauver <sup>a,b,\*</sup>

<sup>a</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain

<sup>b</sup> AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198 Lleida, Catalonia, Spain

\*Correspondance

Publicat a la revista International Journal of Applied Earth Observation and Geoinformation



# Farming and Earth Observation: Sentinel-2 data to estimate within-field wheat grain yield

Joel Segarra<sup>a,b</sup>, Jose Luis Araus<sup>a,b,\*</sup>, Shawn C. Kefauver<sup>a,b,\*</sup>

<sup>a</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain

<sup>b</sup> AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198 Lleida, Catalonia, Spain

## ARTICLE INFO

### Keywords:

Grain yield  
Machine learning  
Precision farming  
Remote sensing  
Sentinel-2  
Wheat

## ABSTRACT

Wheat grain yield (GY) is a crop feature of central importance affecting agricultural, environmental, and socioeconomic sustainability worldwide. Hence, the estimation of within-field variability of GY is pivotal for the agricultural management, especially in the current global change context. In this sense, Earth Observation Systems (EOS) are key technologies that use satellite data to monitor crop yield, which can guide the application of precision farming. Yet, novel research is required to improve the multiplatform integration of data, including data processing, and the application of this discipline in agricultural management. This article provides a novel methodological analysis and assessment of its applications in precision farming. It presents an integration of wheat GY, Global Positioning Systems (GPS), combine harvester data, and EOS Sentinel-2 multispectral bands. Moreover, it compares several indices and machine learning (ML) approaches to map within-field wheat GY. It also analyses the importance of multi-date remote sensing imagery and explores its potential applications in precision agriculture. The study was conducted in Spain, a major European wheat producer. Within-field GY data was obtained from a GPS combine harvester machine for 8 fields over three seasons (2017–2019) and consecutively processed to match Sentinel-2 10 m pixel size. Seven vegetation indices (NDVI, GNDVI, EVI, RVI, TGI, CVI and NGRDI) as well as the biophysical parameter LAI (leaf area index) retrieved with radiative transfer models (RTM) were calculated from Sentinel-2 bands. Sentinel-2 10 m resolution bands alone were also used as variables. Random forest, support vector machine and boosted regressions were used as modelling approaches, and multilinear regression was calculated as baseline. Different combinations of dates of measurement were tested to find the most suitable model feeding data. LAI retrieved from RTM had a slightly improved performance in estimating within-field GY in comparison with vegetation indices or Sentinel-2 bands alone. At validation, the use of multi-date Sentinel-2 data was found to be the most suitable in comparison with single date images. Thus, the model developed with random forest regression (e.g.  $R^2 = 0.89$ , and  $RSME = 0.74$  t/ha when using LAI) outperformed support vector machine ( $R^2 = 0.84$  and  $RSME = 0.92$  t/ha), boosting regression ( $R^2 = 0.85$  and  $RSME = 0.88$  t/ha) and multilinear regression ( $R^2 = 0.69$  and  $RSME = 1.29$  t/ha). However, single date images at specific phenological stages (e.g.  $R^2 = 0.84$ , and  $RSME = 0.88$  t/ha using random forest at stem elongation) also posed relatively high  $R^2$  and low RMSE, with potential for precision farming management before harvest.

## 1. Introduction

Crop yield gathers maximum attention in agricultural research and public policies due to its importance for guiding economic decisions and ensuring farmers' incomes and food security (Rosegrant and Cline, 2003; van Ittersum, 2016). In the current global context, crop management faces several challenges which affect negatively both the production and the environment. On the one hand, crop yield is highly

dependent on external inputs (Sutton et al., 2011), which, when mismanaged, can lead to environmental pollution and a decline in the performance. On the other hand, crop yields are highly susceptible to external pressures such as climatic conditions or pests (Savary et al., 2019), as well as of spatial variability in soil fertility which jeopardize the performance of crops. Therefore, the high-resolution prediction of crop yields before harvest plays a central role in addressing these challenges and guiding management decisions affecting agricultural,

\* Corresponding authors at: Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain.

E-mail addresses: [jaraus@ub.edu](mailto:jaraus@ub.edu) (J.L. Araus), [sckefauver@ub.edu](mailto:sckefauver@ub.edu) (S.C. Kefauver).

<https://doi.org/10.1016/j.jag.2022.102697>

Received 25 October 2021; Received in revised form 21 January 2022; Accepted 23 January 2022

Available online 5 February 2022

1569-8432/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

environmental, and socioeconomic sustainability. In this sense, Earth Observation Systems (EOS) are key technologies that use satellite data and can be used to monitor crop yield and potentially guide the application of management strategies within precision agriculture. In the last years, EOS have improved their resolutions and consequently the potential to apply precision agriculture in crop management has increased (Migdall et al., 2018), as well as the capacities to monitor environmental features affecting crops (Hunt et al., 2019b).

Among the current openly accessible EOS, the system that has gathered more attention in recent years in precision agriculture is Sentinel-2. Since its full operationality was reached in 2018, the European-launched satellite has a revisit time of 5 days, in addition to an improved spatial and spectral resolution in comparison with other equivalent satellites (Segarra et al., 2020a) used in precision farming. The open access nature of Sentinel-2 data is also a central factor for its use, as many high to middle resolution satellites have paywalls. In addition to EOS, other technological advancements such as Global Positioning System (GPS)-supported agricultural machinery have been developed in the last years and allows for a multiplatform development of precision agriculture (Bach and Mauser, 2018). These technological advancements can coordinate to monitor key agricultural features such as yield in a more efficient fashion. Besides the technological advancements, modelling approaches such as machine learning (ML) have shown considerable promise in agricultural remote sensing applications (Chlingaryan et al., 2018). These ML computer algorithms are particularly useful for studying complex biological systems, as they can capture complex interactions among variables and find generalizable predictive patterns (Bzdok et al., 2018). The use of multiplatform data to develop models that allow estimating grain yield before harvest is central to guiding agricultural management and both secure crop yield and optimize the use of resources (Foley et al., 2011).

Sentinel-2 data in combination with GPS combine harvesters can contribute to precisely generate grain yield (GY) estimation models. Nonetheless, the potential for estimating within-field GY variability has yet to be fully explored and calls for a further integration of multiplatform data and streamline data processing. In the case of Sentinel-2, only a few articles have dealt with remote sensing and GPS combine harvesters, as for example the case of cereals maize (Kayad et al., 2019) and wheat (Cavalaris et al., 2021; Hunt et al., 2019b). For maize, ML random forest (RF) and support vector machine (SVM) regressions were tested to match Sentinel-2 vegetation indices with GPS GY points at a single field. The results were modest ( $R^2 = 0.48$ ) but showed potential in determining the best phenological stage when estimating within-field GY. In the case of wheat, Cavalaris et al. (2021) demonstrated the possibility to match Sentinel-2 derived vegetation indices (VIs) with within-field GY using simple regressions at different phenological stages ( $R^2 = 0.63$ ). Also in wheat, Hunt et al. (2019) used RF, VIs and Sentinel-2 bands alone to estimate wheat within-field GY in a single season. Their results suggested that Sentinel-2 bands alone and VIs have similar results ( $R^2 = 0.89$ ).

So far, however, biophysical variables, such as LAI (leaf area index), derived from radiative transfer models (RTM) have not been used in estimating within-field wheat GY. Overall, VIs have limitations in the radiometric information they can exploit and are not as robust as RTM (Maes and Steppe, 2019; Weiss et al., 2000). In this sense, and compared with VIs, LAI has shown improvements for grain yield estimation for several crops (Duan et al., 2021; Gilardelli et al., 2019; Lambert et al., 2018; Mokhtari et al., 2018; Zhou et al., 2017). Moreover, in the case of wheat, so far ML approaches have not been matched with phenological stages, which could be relevant to precision agriculture management needs for crop-specific models. Furthermore, in the case of wheat within-field grain yield, ML learning approaches have not been fully explored, as only simple regressions and RF have been tested. In this sense, support vector machine (SVM) was also tested for within-field maize grain yield estimation (Kayad et al., 2019), but the study focused exclusively on the use of vegetation indices. Similarly, boosted

regression (BR) has shown improved performance for yield estimation applied to winter wheat compared to SVM by several authors (Heremans et al., 2015; Stas et al., 2016), but again both studies used only vegetation indices and lower-resolution satellite data.

We aim to make the most of current technological advancements to define the most suitable variables for estimating within-field wheat GY. On the one hand, we use GPS technological advancements in combine harvesters, which have provided large (within fields) geolocated grain yield datasets. On the other hand, we match this data with remotely sensed Sentinel-2 satellite spectral information and RTM to feed ML GY estimation models. Over three seasons (2017–2019) eight different fields in the Province of Burgos (Spain) were monitored with Sentinel-2 time series, biophysical parameter LAI derived from RTM and different VIs (NDVI, GNDVI, EVI, RVI, TGI, CVI and NGRDI) were calculated for each field over the various image dates obtained throughout the seasons. The Sentinel-2 bands alone, the biophysical parameter and VIs were independently matched to the GPS combine harvester dataset. The ML approaches RF, SVM, generalized boosting regression (BR), together with a multilinear regression as baseline, were used. The study was structured around three questions.

- 1) What processing or combination of Sentinel-2 derived spectral information is more suitable to estimate within-field wheat grain yield?
- 2) Can wheat GY models matching phenological stages be accurate and have potential to be applied in precision agriculture?
- 3) Which ML approach is more suitable for high resolution crop-specific wheat grain yield modelling?

This study concludes by discussing the potential of the findings presented here to be applied in field-level management for precision farming. Moreover, it suggests some guidelines to advance towards a generalized use of these technologies among farmers.

## 2. Materials and methods

### 2.1. Study site

The study was conducted during three seasons (2017–2019) in the Province of Burgos. The region where this study takes place is the Northern inner plateaus (*Meseta Norte*) of Spain. This region is in the Duero Basin and is characterized by agricultural fields growing mainly winter cereals (barley and wheat). Regarding wheat types, Spain is among the top durum wheat producers in Europe (Ranieri, 2015), with the Duero Basin region (Castilla y León) being among the largest producers at the country level after Andalucía in Southern Spain. Data from 8 fields in the Province of Burgos in Spain (Fig. 1) were used. Fields were conventionally managed and the variety Athoris of durum wheat (*Triticum turgidum* L. subsp. *durum* (Desf) Husn.), which is widely grown in the region, was sowed. The location of the fields is shown in Fig. 2.

*Meseta Norte* has an inner-Mediterranean/continental climate with an average temperature of 10–14 °C, annual rainfall under 600 mm and average height above sea level of 800–850 m (Font Tullot, 2000). The fields were sown between mid-October and first week of November for the three seasons. The fields were fertilized with chemical fertilizers averaging between 35 and 40 kg/ha of N, 80–130 kg/ha of P and 60–90 kg/ha of K in the case of basal fertilization. Regarding top-dressing fertilization, between 85 and 110 kg/ha N was applied throughout the three seasons in the different fields between tillering and stem elongation, corresponding to Z30–Z32 phenological stages (Zadoks et al., 1974). The average grain yield over the three seasons was 4.81 t/ha.

Regarding the description of the fields shown in Fig. 2, number 1 has a surface of 21.30 ha and a slope of less than 1%, number 2 has a surface of 25.80 ha and a slope of less than 1%, number 3 has a surface of 5.80 ha and a slope of 3%, number 4 has surface of 55.64 ha and a slope of 7%, number 5 has a surface of 18.50 ha and a slope of 2%, and number 6 has



**Fig. 1.** Map of the Iberian Peninsula with the province of Burgos, where the fields are located, highlighted in green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

a surface of 9.70 ha and a slope of 2%. The area is characterized by loam/clay soils with a homogenous distribution (<https://suelos.itacyl.es/mapas>).

## 2.2. Sentinel-2 data and indices calculation

The Sentinel-2 multispectral bands (Table 1) were downloaded without cloud cover from Copernicus Open Access Hub (<https://scihub.copernicus.eu/>) as a 2A product (Bottom of Atmosphere reflectance images) for the following dates: 2017 (05–22), 2018 (04–17, 05–26), 2019 (04–12, 05–12). In addition, the imagery captured on 04–14-2017 was downloaded as a L1C product (Top of Atmosphere reflectance images) and was subsequently corrected to level 2A using the Sen2Cor tool on SNAP (Sentinel Application Platform), obtaining Bottom-Of-Atmosphere (BOA) and cirrus corrected reflectance images. The sensing dates were matched to the phenological stages *in situ* evaluated following the Zadoks scale (Zadoks et al., 1974): stem elongation (30–39), heading (41–59) and anthesis (61–69).

Seven widely used VIs (Table 2) were calculated on ArcGIS Pro 2.3.0 as shown in Table 2. Moreover, the biophysical parameter, LAI (leaf area index) was calculated at 10 m of spatial resolution following neural network algorithms trained with PROSAIL radiative transfer models (Weiss and Baret, 2016) on SNAP (Sentinel Application Platform) (Table 2).

## 2.3. Matching of sentinel-2 and GPS combine harvester grain yield data

The fields were harvested between mid-July to the beginning of August throughout the three seasons. A combine harvester (John Deere T660i + 625R) equipped with a GPS was used and relative GY results (t/ha) were reported. As the obtained results were relative to the surface and not absolute weights, we considered that no calibration between actual weight and harvested weight was necessary. All GPS GY points obtained with the combine harvester were uploaded as a shapefile in

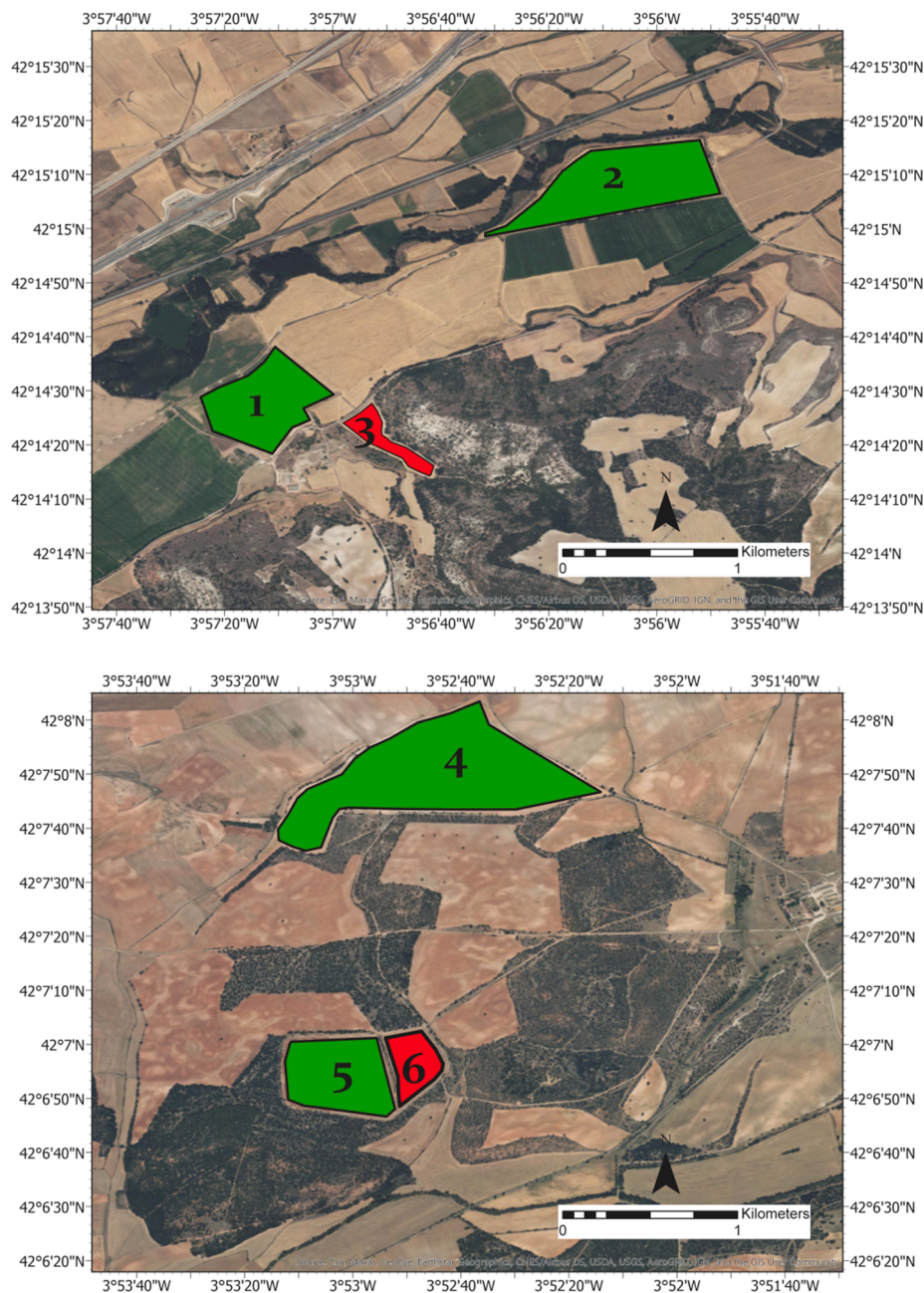
ArcGIS Pro 2.3.0, and subsequently all points contained within each pixel were averaged using the Spatial Analysis Menu Tool Point to Raster. As a result, each single pixel value of the calculated LAI, NDVI, GNDVI, EVI, RVI, TGI, CVI and NGRDI layers and the original 10 m resolution Sentinel-2 bands were associated to one GY value. This processing protocol was carried out at each phenological stage. An overview of few indices calculated is shown in Fig. 3. A total number of 20,124 pixels with their corresponding unique GY values and Sentinel-2 derived data was obtained as data set. The raw data was manually buffered to avoid edge effects of mix pixels as shown in Fig. 4. All the data was matched in a point vector file, it was easier to process and use as data is organized in attribute tables. The processing was carried out on ArcGIS Pro 2.3.0 and a schematic illustration describing the process is shown in Fig. 4. The cut width of the combine harvester was 7.60 m and therefore adequate to 10 m pixels size of Sentinel-2 imagery. Notwithstanding its suitability for Sentinel-2 pixel size, the GY spatial resolution of the combine harvester, as it is a function of the monitoring equipment, the cutting head and the software has some lack of precision, as reported by other authors (Lyle et al., 2014).

All the GPS GY points with values lower to 0.5 t/ha were discarded. Moreover, a portion of the raw data was trimmed using the “tclust” package in R (R Core Team, 2021) based on GY and Sentinel-2 bands inaccuracies. Trimming allows the removal of the most outlying data which can strongly influence the results; this approach has been introduced by several authors (García-Escudero et al., 2008; Woodruff and Reiners, 2004). The number of clusters ( $k$ ) and portion of data trimmed ( $\alpha$ ) was optimized and subsequently applied to remove outliers ( $k = 5$ ,  $\alpha = 0.05$ ). The general process is shown in Fig. 5.

## 2.4. Machine learning approaches

One of the most widely used machine learning (ML) approaches is random forest (RF) regression. It has been successfully used for estimating wheat grain yield (Hunt et al., 2019a; Jeong et al., 2016; Wang





**Fig. 2.** Locations of the fields used in the study during season 2017–2019. Six fields are shown, two of them were used in alternate seasons, and the reference number for each field is indicated. Fields used independently for validation of the GY estimation model are highlighted in red, those used for training are in green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

et al., 2016). The RF algorithm is a ML method developed by (Breinman, 2001), where the set of predictors is randomly distributed at each split, while the variables randomly selected at each split is a user-defined parameter as well as the number of trees. The optimal parameters regarding the number of trees and the number of variables randomly sampled as candidates at each split was optimized to 300 and 6 respectively. In this case the package “RandomForest” was used in R (R Core Team, 2021).

Support vector machine (SVM) used in regressions (Vapnik, 2013) is a ML approach that has also been used in agriculture for grain yield estimation (Oguntunde et al., 2018; Saruta et al., 2013). In R (R Core Team, 2021) the package “e1071” was used for the SVM in eps-regression, and the degree parameter needed for kernel was set at 3. Generalized boosted regression (BR) are combinations of two techniques: decision tree algorithms and boosting methods (Freund et al.,

1999; Friedman, 2001). These approaches have been successfully used in grain yield estimation (Arumugam et al., 2021; Zhang et al., 2019). In R (R Core Team, 2021) package “gbm” was used, the number of trees was optimized at 5000 and the interaction depth at 5.

For all ML approaches, a 70% of the dataset (13,685 pixels) was used to train the models and a 30% of the dataset (6439 pixels) containing two fields not used in the training were used for validation. Fig. 6 shows the distribution of the data set, as density due to the different number of pixels in each set. We assume the distribution of data is equivalent in train and validation sets. Both training and validation datasets followed a similar distribution of data (Fig. 6). The accuracy assessment was carried out by calculating  $R^2$ , RMSE (eq. (1)) and actual GY MEAN (eq. (2)) to derive the % RMSE (eq. (3)) of the validation dataset.

**Table 1**  
Sentinel-2 multispectral instrument information.

Band	Spatial Resolution (m)	Central Wavelength (nm)
B1: Coastal Aerosol	60	443
B2: Blue	10	490
B3: Green	10	560
B4: Red	10	665
B5: Red-Edge	20	705
B6: Red-Edge	20	740
B7: Red-Edge	20	783
B8: NIR	10	842
B8A: Vegetation RE	20	865
B9: Water Vapour	60	945
B10: SWIR Cirrus	60	1375
B11: SWIR	20	1610
B12: SWIR	20	2190

**Table 2**  
The calculated vegetation indices and biophysical parameters.

Abbreviation	Parameter	Calculation	Reference
NDVI	Vegetation index	$\frac{B8 - B4}{B8 + B4}$	(Rouse Jr., 1974)
GNDVI	Vegetation index	$\frac{B8 - B3}{B8 + B3}$	(Gitelson et al., 1996)
EVI	Vegetation index	$2.5 \cdot \frac{B8 - B4}{(B8 - 6 \cdot B4 - 7.5 \cdot B2) + 1}$	(Haboudane et al., 2002)
RVI	Vegetation index	$\frac{B8}{B4}$	(Tucker, 1979)
TGI	Vegetation index	$-0.5 \cdot (190 \cdot (B4 - B3) - 120 \cdot (B4 - B2))$	(Hunt et al., 2013)
NGRDI	Vegetation index	$\frac{(B3 - B4)}{(B3 + B4)}$	(Hunt et al., 2005)
CVI	Vegetation index	$\frac{B8 \cdot B3}{B3 \cdot B4}$	(Vincini et al., 2008)
LAI	Biophysical parameter	RTM on SNAP	(Weiss and Baret, 2016)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\text{predicted}_i - \text{actual}_i)^2}{n}} \quad (1)$$

$$MEAN = \frac{\sum_{i=1}^n \text{actual}_i}{n} \quad (2)$$

$$\%RMSE = \frac{RMSE}{MEAN} \cdot 100 \quad (3)$$

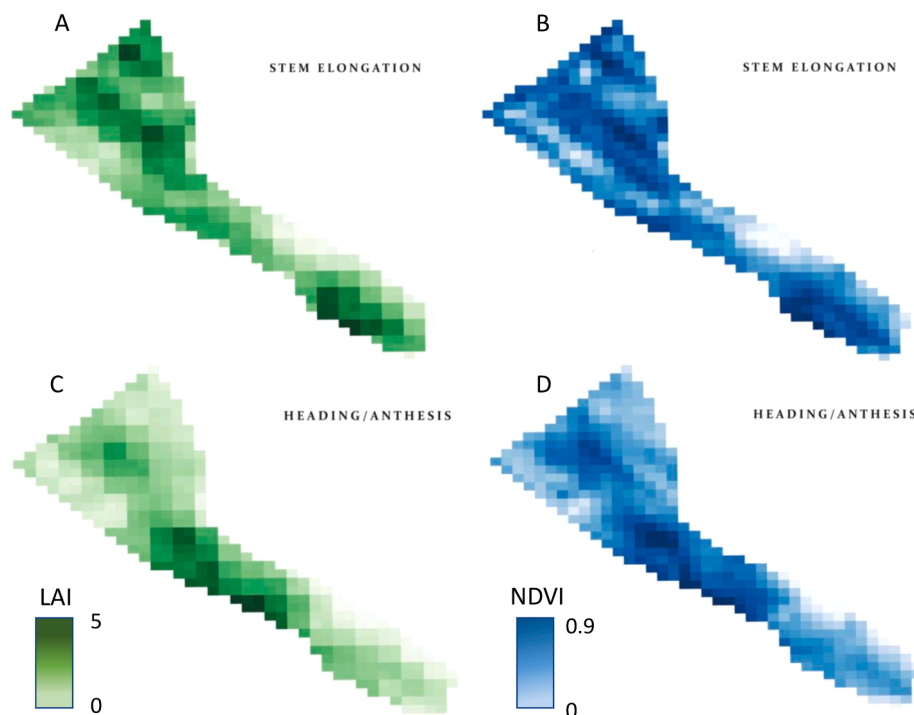
### 3. Results

#### 3.1. Machine learning approaches

The ML approach that produced higher  $R^2$  and lower %RMSE is RF regression (Fig. 7), followed by BR and SVM. All machine learning approaches had better results in comparison with the baseline, multilinear regression (Table 3). For all phenological stages and indices, RF regression averaged between a highest  $R^2 = 0.89$  and %RMSE = 15.4 and a lowest of  $R^2 = 0.83$  and %RMSE = 19.1; meanwhile BR averaged between a highest  $R^2 = 0.85$  and %RMSE = 18.3 and a lowest of  $R^2 = 0.81$  and %RMSE = 20.7; SVM averaged between a highest  $R^2 = 0.84$  and %RMSE = 19.1 and a lowest of  $R^2 = 0.74$  and %RMSE = 23.7. The multilinear regression as baseline averaged between a highest  $R^2 = 0.69$  and %RMSE = 26.8 and a lowest of  $R^2 = 0.37$  and %RMSE = 35.7.

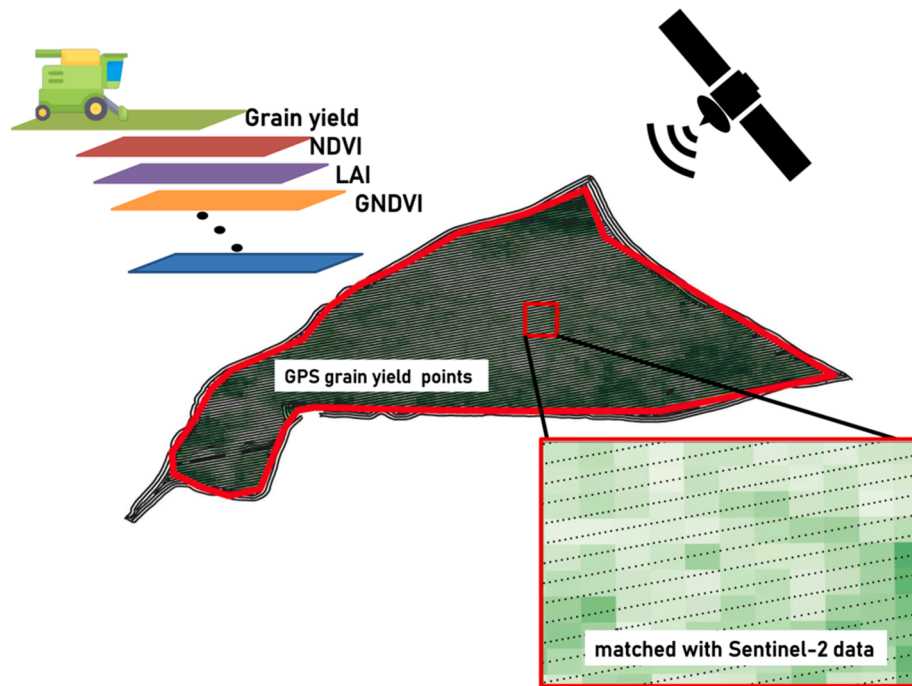
#### 3.2. Phenology and date

Multidate Sentinel-2 imagery improved all data feeding instances here studied. There is a slightly improved result when using Sentinel-2 timeseries, rather than stem elongation and heading/anthesis phenological stages alone. In general terms, with single date images, heading/anthesis phenological stage performed slightly better in comparison with stem elongation (Table 3). In the case of RF regression, for instance, at stem elongation  $R^2$  ranged between 0.83 and 0.84 and %RMSE



**Fig. 3.** Overview of LAI (in green A and C) and NDVI (in blue B and D) calculated from the Sentinel-2 images, with Field 3 as an example of the model validation with the two phenological stages studied, stem elongation at the top of the Figure (A and B) and heading/anthesis at the bottom (C and D). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)





**Fig. 4.** Explanation of the distribution of yield points obtained with the combine harvester. Fields were manually buffered to avoid edge effects. The interpolation of GPS yield points to Sentinel-2 derived data is also shown.

18.3–19.1. Meanwhile at heading/anthesis,  $R^2$  ranged between 0.86 and 0.88 and %RMSE 16.8–17.0. Finally, when using the whole timeseries,  $R^2$  ranged between 0.87 and 0.89 and %RMSE 15.4–16.2. Hence, the estimation of within-field grain yield has a slightly improve when combining sensing dates.

### 3.3. Vegetation indices and biophysical parameters

In Fig. 7 a linear regression between the observed and estimated values at validation using multivariate LAI data is shown. A higher density of points is observed around 3 to 6 t/ha as most sampled points were around these values. The linearity of the results is evident, yet the density of points is slightly uneven due to the nature of the data obtained in this area. The results suggest that with RF regression LAI using stem elongation and heading/anthesis dates together has the best performance ( $R^2 = 0.89$  and %RMSE = 15.4). However, for the same specific situation VIs ( $R^2 = 0.87$  and %RMSE = 16.2) and 10 m resolution Sentinel-2 bands ( $R^2 = 0.88$  and %RMSE = 15.8) performed relatively well too. This trend was also observed for the rest of machine learning approaches here studied. Therefore, within-field wheat GY variability can be estimated relatively accurately with all the data feeding approaches here studied (vegetation indices, LAI and 10 m Sentinel-2 bands alone) being LAI the most accurate.

Regarding the vegetation indices alone (Table 4), we observed that RF regression yielded the best results. Support vector machine had more limited results and linear regression showed unacceptable %RMSE. With RF regression and multivariate data, RVI, GNDVI and NDVI had the lowest %RMSE (16.2 in all the cases) as well as the highest  $R^2$  (0.87 in all the cases). Greenness sensitive indices such as TGI or CVI had lower  $R^2$  (0.85) and higher %RMSE (17.4) in contrast with the previous biomass sensitive indices.

RF regression is the most promising ML approach when using LAI derived from Sentinel-2 data processing. The use of all available dates also improved the results. Within-field relative accuracy can be observed in Fig. 8, where a comparison between observed and estimated wheat GY using LAI data is shown. The pixels shown in the two fields are validation pixels, two independent fields of the 30% not used for

training the model; therefore, these present an independent visual validation of the accuracy of the developed model. Moreover, they reflect comparable general patterns of wheat GY within-field variability.

## 4. Discussion

This study is structured around three questions that cover three relevant factors for within-field GY variability mapping, with a central relevance in remote sensing for precision farming. First, we determined what Sentinel-2 derived data is more accurate to estimate within-field GY. Second, we analysed how relevant is the temporal variable of the information that is sensed, namely the potential of phenological stages assessment. And third, we inferred what modelling approach, weather simple linear regressions or more complex ML approaches, work best. Finally, we discuss the scope of the findings presented here in regard to precision farming.

### 4.1. Sentinel-2 data

Coming back to question one (What processing or combination of Sentinel-2 derived spectral information is more suitable to estimate within-field wheat GY?) we observed that the biophysical parameter LAI had a slightly improved performance in contrast with Sentinel-2 bands and VIs. The lower %RMSE (15.4) in the model derived from LAI in comparison with the one developed with VIs (16.2 %RMSE) is in line with other studies, which showed improved performance of LAI for estimating cereal (e.g. wheat) grain yield (Lambert et al., 2018; Zhou et al., 2020b). However, VIs constitute a simple approach to extract information from remotely sensed data and have been linked to crop yield features in the last years (Gracia-Romero et al., 2017; Liu et al., 2006; Segarra et al., 2020b). VIs are easy to use, yet the spectral information RTM can capture is more robust as it captures reflectance changes in crop stages and changing architectures (Viña et al., 2011), in this sense, LAI was calculated using RTM, which are able to capture improved information of crop canopies (Wolanin et al., 2019). Moreover, most indices experience saturation at certain crop stages (Bannari et al., 1995), although a combination of trait specific indices might

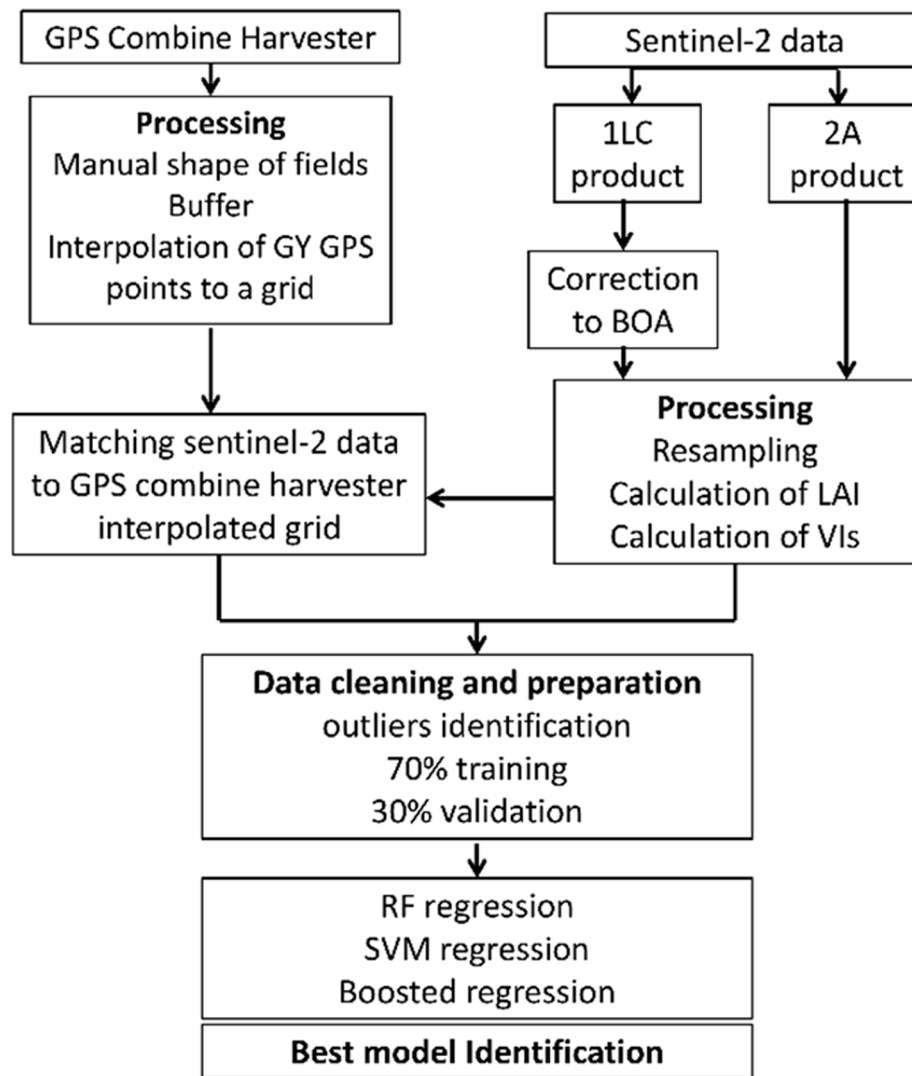


Fig. 5. Method overview of the process to estimate wheat GY using several data and ML approaches.

overcome this issue. Regarding Sentinel-2 bands alone, there is still a slightly improvement of GY models when using LAI (%RMSE 15.4) than bands alone (%RMSE 15.8). Other authors (Hunt et al., 2019a) argued that ML algorithms extract the most from the spectral information due to their capacity to capture complex interactions among variables, and therefore are successfully linked with within-field GY. Nonetheless, on top of this, we can argue that hybrid models that use advanced RTMs to describe biophysical parameters such as LAI and subsequently are matched to GY to feed ML models can even improve and add robustness to the models. These approaches have mainly been used in phenotyping (Jin et al., 2021), but as observed here can be almost operational for use in within-field wheat GY estimation.

#### 4.2. Time series and phenology

Regarding question two (Can wheat GY models matching phenological stages be accurate and have potential to be applied in precision agriculture?) we observed that multi-date data improves the capacity to estimate within-field grain yield variability. In all three data feeding instances (Sentinel-2 bands, LAI and VIs),  $R^2$  was higher and RMSE lower when using multi-date Sentinel-2 data (Table 3). Yet, the full temporal resolution of Sentinel-2 images could not be exploited as the availability of images for the regions studied was limited throughout the three seasons due to cloud cover. Despite of the limitations, several

images throughout the season could be analysed following the phenological stages stem elongation and heading/anthesis. Interestingly, LAI assessed at stem elongation and heading/anthesis predicted better GY (lower %RMSE) than VIs and Sentinel-2 bands. This is coherent with the fact that LAI is a good indicator of potential canopy photosynthesis, with these reproductive stages being crucial in defining yield (Miralles et al., 2000; Villegas, 2001). In the case of the VIs and eventually of the single bands they may reflect later (during the crop cycle); particularly in terms of stay green and the onset of crop senescence during grain filling (Aparicio et al., 2000). Decrease in Stay Green, as response for example of water stress or lack of mineral nutrients effects may affect GY through an accelerated reduction of photosynthetic assimilates during the grain filling (Christopher et al., 2016), which might be too late for a reasoned management (e.g. additional irrigation or top-dressing fertilization). Moreover, LAI and VIs in wheat are frequently weakly correlated, which means their performance predicting yield are not necessarily comparable (Serrano et al., 2000). In this sense, for single phenological stages, the efficiency of the models was relatively high. Hence the estimation of GY before the whole season is completed, around March to May can help farmers to guide management decisions and ensure a better harvest. Several authors have matched Sentinel-2 images with phenology to estimate wheat GY at various resolutions (Fieuzal et al., 2020; Segarra et al., 2020b; Toscano et al., 2019) that allows an improved opportunity for precision farming management. The technical progress of Sentinel-2

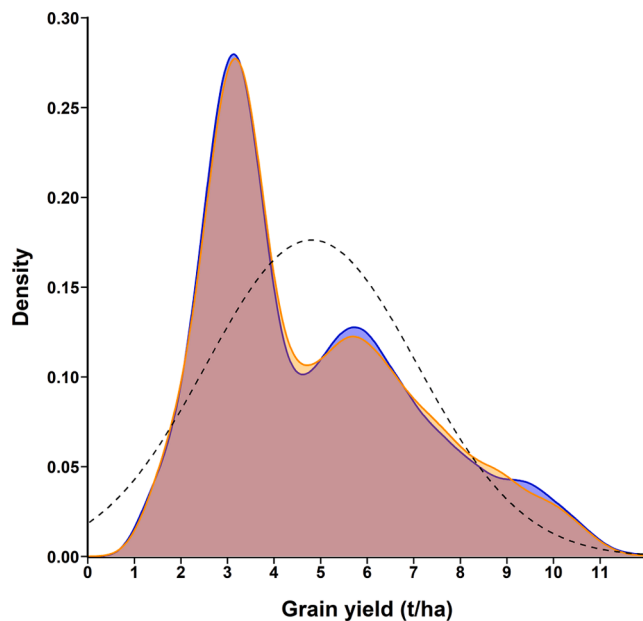


Fig. 6. Datasets distribution, orange shading indicates data forming the test set, blue shading the data forming the training set. The dotted line indicates a theoretical perfect normal distribution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and the increasing development of modelling approached for GY estimation now makes precision farming affordable and cost-effective and thus, almost operational (Weiss et al., 2020).

#### 4.3. Modelling approaches

Regarding the third question (Which ML approach is more suitable for high resolution crop-specific wheat GY modelling?) we observed that RF clearly outperformed BR, SVM and multilinear regression (Table 4). In contrast with this, other authors (Heremans et al., 2015) observed an improved performance of BR in comparison with RF regression, when using VIs. Nonetheless, we can argue that the findings of this article

show RF as an improved modelling approach when using LAI to estimate within field grain yield, as discussed previously. In contrast with other authors who did not find improvements between simple linear regressions and ML (Uno et al., 2005; Zhou et al., 2020a), we found that the relationship between crop yield and reflectance is complex enough for ML approaches and results in improvements in within-field yield predictions. RF is less prone to outliers, and hence one would expect yield estimation performance to increase, as we have corroborated here. Besides, RF algorithm is powerful in handling both linear and non-linear relationship as wheat yield and spectral information might have some non-linearity (Kumhálová and Matějková, 2017; Tesfaye et al., 2021). In general, we observed that the modelling approach is central for an effective GY estimation. We can argue that the modelling is the most important factor, followed by the sensing date (timeseries availability importantly linked to phenology) and the processing of the spectral data (LAI, for instance). In an equivalent study developed in the UK, RF was also used and compared to simple regression but no other machine learning approaches were presented (Hunt et al., 2019a). The results here obtained confirm the effectiveness of RF regression to estimate within-field GY variability using Sentinel-2 imagery for the case of Spain. Nonetheless, a further research and comparison of ML approaches could be developed aiming to standardize methodologies for specific regions and crop-specific cases for this almost operational precision agriculture technology. The ability of RF to cope with multivariate relationships between data of different types and resolutions is a key advantage over methods such as linear regression, which can only address univariate relationships.

#### 4.4. Precision farming

To our knowledge, hitherto advanced models using ML approaches to estimate within-field GY in cereals have focused on demonstrations (Hunt et al., 2019a; Kayad et al., 2019) of single machine learning models and VIs or simple regressions (Cavalaris et al., 2021). Overall, there are a handful of articles (only two focused on bread wheat and none on durum wheat) dealing with GPS combine harvester and Sentinel-2 images and none have assessed several ML approaches, spectral indices and biophysical parameters retrieved from RTM. Hence, the results here presented assessing several ML approaches and Sentinel-2 data processing approaches are a novelty for this growing research

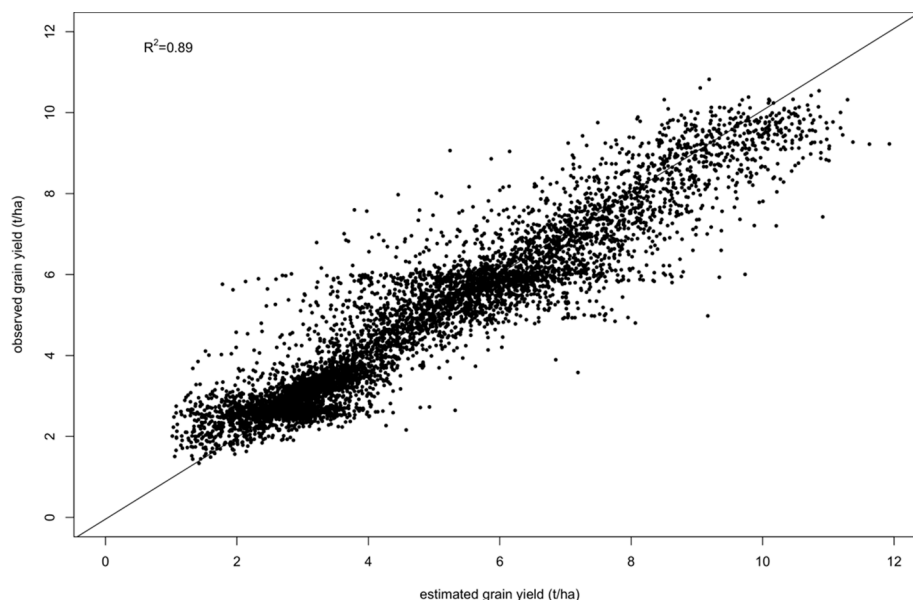


Fig. 7. Regression between observed and estimated-GY relative values at validation using RF regression, LAI data) and the whole timeseries combined. The black line indicates perfect correlation.

**Table 3**

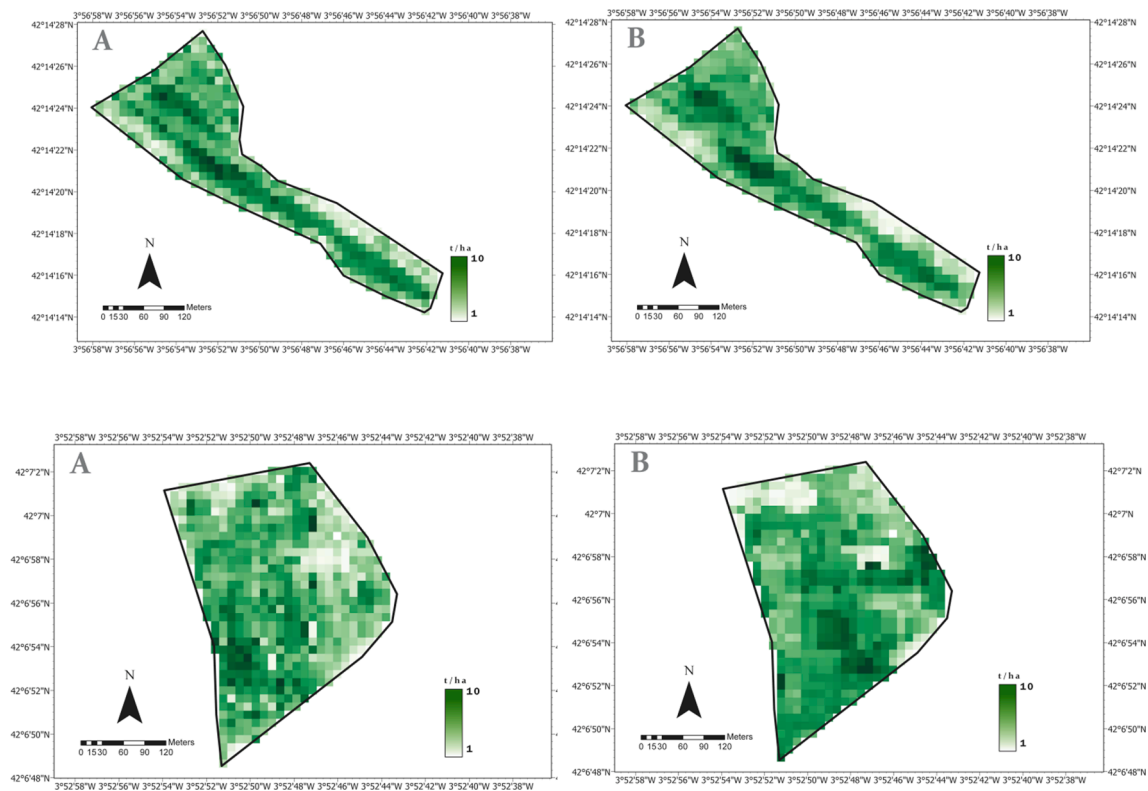
At validation, model accuracies achieved for the various sensing periods and indices with random forest regression, support vector machine, boosted regression and baseline multilinear regression. S2 refers to Sentinel-2, while RMSE stands for root mean square error.

S2 date	Phenological stages	data	Random Forest			Support Vector Machine			Boosted Regression			Multilinear regression		
			R <sup>2</sup>	RMSE (t/ha)	% RMSE	R <sup>2</sup>	RMSE (t/ha)	% RMSE	R <sup>2</sup>	RMSE (t/ha)	% RMSE	R <sup>2</sup>	RMSE (t/ha)	% RMSE
March/April	Stem elongation	VIs	0.83	0.92	19.1	0.74	1.11	23.0	0.82	0.96	19.9	0.61	1.33	27.6
		LAI	0.84	0.84	17.4	0.72	1.09	22.6	0.84	0.82	17.1	0.47	1.73	35.9
		10 m S2	0.84	0.88	18.3	0.76	1.14	23.7	0.81	1.00	20.7	0.53	1.57	32.6
May	Heading/Anthesis	VIs	0.86	0.82	17.0	0.80	0.99	20.5	0.84	0.92	19.1	0.54	1.41	29.3
		LAI	0.88	0.76	15.8	0.81	1.01	20.9	0.82	1.01	21.0	0.38	1.86	38.6
		10 m S2	0.86	0.81	16.8	0.81	1.09	22.6	0.82	0.97	20.1	0.37	1.72	35.7
March/April + May/June	Stem elongation + Heading/Anthesis	VIs	0.87	0.78	16.2	0.82	0.89	18.5	0.84	0.91	18.9	0.69	1.29	26.8
		LAI	0.89	0.74	15.4	0.84	0.92	19.1	0.85	0.88	18.3	0.48	1.74	36.1
		10 m S2	0.88	0.76	15.8	0.83	0.91	18.9	0.83	0.94	19.5	0.58	1.49	30.9

**Table 4**

Testing of single vegetation indices for all the analysed dates together.

Vegetation index	R <sup>2</sup>	Random Forest		Support Vector Machine			Boosted Regression			Linear regression		
		RMSE (t/ha)	%RMSE	R <sup>2</sup>	RMSE (t/ha)	%RMSE	R <sup>2</sup>	RMSE (t/ha)	%RMSE	R <sup>2</sup>	RMSE (t/ha)	%RMSE
GNDVI	0.87	0.78	16.2	0.78	1.00	20.7	0.83	0.79	16.4	0.47	1.60	33.2
NDVI	0.87	0.78	16.2	0.79	0.98	20.3	0.83	0.79	16.4	0.53	1.56	32.4
EVI	0.84	0.81	16.8	0.77	1.02	21.2	0.81	0.81	16.8	0.56	1.44	29.9
RVI	0.87	0.78	16.2	0.78	1.03	21.4	0.83	0.79	16.4	0.51	1.56	32.4
TGI	0.85	0.84	17.4	0.74	1.20	24.9	0.83	0.82	17.0	0.37	1.75	36.3
NGRDI	0.85	0.79	16.4	0.75	1.05	21.8	0.84	0.79	16.4	0.55	1.55	32.2
CVI	0.85	0.84	17.4	0.74	1.24	25.7	0.83	0.81	16.8	0.37	1.77	36.7



**Fig. 8.** Comparison between actual GY pixels (A) and estimated GY with the independent validation dataset (B) in the field 3 (top) and 6 (bottom) using random forest regression and LAI data derived from RTM and Sentinel-2 data.



field. In this sense, we believe that the methodology here developed, and the results obtained can contribute to precision farming. First, the ML approaches analysed present RF as an efficient model to monitor within-field wheat GY. The use of ML to retrieve GY has been considered one of the most important areas develop associated with remote sensing and agriculture (Weiss et al., 2020). On top of that, based on the results, we can argue that retrieving biophysical parameters with RTM that can be linked to crop traits and subsequently used for ML models is a step forward in precision agriculture. The usage of RTMs with ML regression algorithms is opening up a powerful and promising field of vegetation properties retrieval from EOS data (Berger et al., 2020; Estévez et al., 2020), but has yet to be fully explored in precision agriculture. LAI, leaf structure parameters, leaf chlorophyll content, leaf carotenoid content, leaf anthocyanin content or leaf equivalent water thickness are, among others, traits of agricultural crops that can be retrieved from RTM (Danner et al., 2021). The availability of biophysical processors in software such as SNAP of the European Space Agency used here opens a door to the applicability of these parameters. In this sense, with increased computing power available in the next years, RTM inversion can provide an elegant alternative to estimate fertilizer requirements in precision agriculture (Maes and Steppe, 2019). Furthermore, RTM have shown the potential of Sentinel-2 data to map disease-incidence dynamics in agriculture plots (Hornero et al., 2020). Regarding irrigation, cropland canopy water content thematic layers have been developed with RTM and Sentinel-2 imagery (Boren and Boschetti, 2020) using several parameters such as water thickness or dry matter content, among others.

In addition to this, the results obtained in this study with single image acquisition and its matching with phenological stage can help to locate low-yielding spots in croplands and subsequently apply precise farming decisions. In this sense, the performance of LAI and other biophysical parameters of interest for precision agriculture retrieved from RTM could contribute to find physiological or agronomic explanations and apply reasoned managements. Our study also has some limitations to overcome such as its applicability in other crops and in different agroclimates.

We believe that a way to implement the adoption of precision farming for wheat in the inner Mediterranean/continental regions of Spain following these findings would be:

- 1) Further expanding this study by increasing the geolocated dataset of GPS combine harvesters across different environmental and crop type conditions in Spain. This machinery is more accessible nowadays and is being applied more frequently now. Moreover, Spain, where this study takes places, meets a wide range of climatic conditions and crop types with relevance to European agricultural systems.
- 2) Based on our results, RF regression is the most suitable modelling approach to estimate within-field GY at different growing stage and can serve as a starting point to further develop streamlined and standardized models
- 3) Biophysical parameters retrievable with RTM such as LAI, leaf structure parameters, leaf chlorophyll content, leaf carotenoid content, leaf anthocyanin content or leaf equivalent water thickness, among others, could be used as explanatory variables to understand limitations in specific field spots and guide the application of adequate management strategies (smart fertilization, regeneration of degraded soils, etc.). These parameters can be relevant to physiologically and agronomic-related crop performances at each phenological stages.
- 4) The public agricultural institutions of the country together with farmers could lead the creation of an openly accessible platform to estimate within field wheat GY and guide specific management decision for farmers with this information. Following the example of Belgium and the platform WatchItGrow (Curnel, 2017) to monitor potato yield with Sentinel-2 data.

## 5. Conclusions

This article studied the performance of Sentinel-2 derived spectral information, EOS Sentinel-2 multi-date images, and machine learning approaches to define the most effective parameters for within-field GY variability mapping. To our knowledge, no study has focused on the use of biophysical parameters (LAI) retrieved from RTM and a complete set of VIs as training data for high resolution wheat GY estimation as they have in general terms focused on few VIs or spectral bands alone. As most studies focus on field to regional scale GY estimation, only a scarce number of studies have focused on within-field wheat GY. Hence, this study contributes with novel methodological advancements in within-field wheat GY estimation (and to the best of our knowledge none in durum wheat) with implications in precision farming. When comparing the individual performances of different categories of parameters (LAI versus VIs / single bands) assessed at a given phenological stage we found that each category of traits best predicted GY at a different phenological stage, which may have also practical applications for precision agriculture. We found that the modelling approach (specifically RF in our study) is the most relevant factor for GY estimation, followed by the sensing date (single date or multi-date time series) and finally the Sentinel-2 derived information (biophysical parameters, vegetation indices or band alone). At the best performing scenario RF regression with the whole phenological images available the model reached and  $R^2$  of 0.89 and a % RMSE of 15.4, which we can consider a relatively high performing model. Understanding wheat GY variability within a field and applying precision farming management practices is central for environmental, agricultural, and socioeconomic sustainability.

## CRedit authorship contribution statement

**Joel Segarra:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization. **Jose Luis Araus:** Funding acquisition, Resources, Supervision, Writing – original draft, Writing – review & editing. **Shawn C. Kefauver:** Conceptualization, Project administration, Resources, Supervision.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We acknowledge the support of the project PID2019-106650RB-C21 from the Ministerio de Ciencia e Innovación, Spain. J.S. is a recipient of a FPI doctoral fellowship from the same institution (grant: PRE2020-091907). J.L.A. acknowledges support from the Institució Catalana de Recerca i Estudis Avançats (ICREA), Generalitat de Catalunya, Spain). S. C.K. is supported by the Ramon y Cajal RYC-2019-027818-I research fellowship from the Ministerio de Ciencia e Innovación, Spain. We acknowledge the support of Cerealto Siro Group, together with Cristina de Diego and Javier Velasco, technical staff from the company, by providing the wheat yield data. This research was also supported by the COST Action CA17134 SENSECO (Optical synergies for spatiotemporal sensing of scalable ecophysiological traits) funded by COST (European Cooperation in Science and Technology, [www.cost.eu](http://www.cost.eu)).

## References

- Aparicio, N., Villegas, D., Casadesus, J., Araus, J.L., Royo, C., 2000. Spectral vegetation indices as nondestructive tools for determining durum wheat yield. *Agron. J.* 92 (1), 83–91. <https://doi.org/10.2134/agronj2000.92183x>.

- Arumugam, P., Chemura, A., Schaubberger, B., Gornott, C., 2021. Remote sensing based yield estimation of rice (*Oryza sativa* L.) using gradient boosted regression in India. *Remote Sens.* 13, 1–18. <https://doi.org/10.3390/rs13122379>.
- Bach, H., Mauser, W., 2018. Sustainable agriculture and smart farming. In: *Earth Observation Open Science and Innovation*. Springer, Cham, pp. 261–269. [https://doi.org/10.1007/978-3-319-65633-5\\_12](https://doi.org/10.1007/978-3-319-65633-5_12).
- Bannari, A., Morin, D., Bonn, F., Huete, A.R., 1995. A review of vegetation indices. *Remote Sens. Rev.* 13 (1–2), 95–120. <https://doi.org/10.1080/02757529509532298>.
- Berger, K., Verrelst, J., Féret, J.-B., Hank, T., Woche, M., Mauser, W., Camps-Valls, G., 2020. Retrieval of aboveground crop nitrogen content with a hybrid machine learning method. *Int. J. Appl. Earth Obs. Geoinf.* 92, 102174. <https://doi.org/10.1016/j.jag.2020.102174>.
- Boren, E.J., Boschetti, L., 2020. Landsat-8 and sentinel-2 canopy water content estimation in croplands through radiative transfer model inversion. *Remote Sens.* 12, 2803. <https://doi.org/10.3390/rs12122803>.
- Breiman, L., 2001. Random forest. *Mach. Learn.* <https://doi.org/10.1023/A:1010933404324>.
- Bzdok, K., Altman, N., Krzywinski, M., 2018. Points of significance: statistics versus machine learning. *Nat. Methods* 15 (4), 233–234. <https://doi.org/10.1038/nmeth.4642>.
- Cavalaris, C., Megoudi, S., Maxouri, M., Anatolitis, K., Sifakis, M., Levizou, E., Kyparissis, A., 2021. Modeling of durum wheat yield based on sentinel-2 imagery. *Agronomy* 11 (8), 1486. <https://doi.org/10.3390/agronomy11081486>.
- Chlingaryan, A., Sukkarieh, S., Whelan, B., 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and electronics in agriculture* 151, 61–69. <https://doi.org/10.1016/j.compag.2018.05.012>.
- Christopher, J.T., Christopher, M.J., Borrell, A.K., Fletcher, S., Chenu, K., 2016. Stay-green traits to improve wheat adaptation in well-watered and water-limited environments. *J. Exp. Bot.* 67 (17), 5159–5172. <https://doi.org/10.1093/jxb/erw276>.
- Curnel, Y., 2017. Watch It Grow, an innovative platform for a sustainable growth of the Belgian potato production, FACCE MACSUR Reports.
- Danner, M., Berger, K., Woche, M., Mauser, W., Hank, T., 2021. Efficient RTM-based training of machine learning regression algorithms to quantify biophysical & biochemical traits of agricultural crops. *ISPRS J. Photogramm. Remote Sens.* 173, 278–296. <https://doi.org/10.1016/j.isprsjprs.2021.01.017>.
- Duan, B.o., Fang, S., Gong, Y., Peng, Y.i., Wu, X., Zhu, R., 2021. Remote estimation of grain yield based on UAV data in different rice cultivars under contrasting climatic zone. *F. Crop. Res.* 267, 108148. <https://doi.org/10.1016/j.fcr.2021.108148>.
- Estévez, J., Vicent, J., Rivera-Caicedo, J.P., Morcillo-Pallarés, P., Vuolo, F., Sabater, N., Camps-Valls, G., Moreno, J., Verrelst, J., 2020. Gaussian processes retrieval of LAI from Sentinel-2 top-of-atmosphere radiance data. *ISPRS J. Photogramm. Remote Sens.* 167, 289–304. <https://doi.org/10.1016/j.isprsjprs.2020.07.004>.
- Fieuzal, R., Bustillo, V., Collado, D., Dedieu, G., 2020. Combined use of multi-temporal Landsat-8 and sentinel-2 images for wheat yield estimates at the intra-plot spatial scale. *Agronomy* 10 (3), 327. <https://doi.org/10.3390/agronomy10030327>.
- Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P.M., 2011. Solutions for a cultivated planet. *Nature* 478 (7369), 337–342. <https://doi.org/10.1038/nature10452>.
- Font Tullot, I., 2000. *Climatología de España y Portugal*. Universidad de Salamanca, Salamanca.
- Freund, Y., Schapire, R., Abe, N., 1999. A short introduction to boosting. *J.-Japanese Soc. Artif. Intell.* 1612, 771–780.
- Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 1189–1232. <https://doi.org/10.1214/aos/1013203451>.
- García-Escudero, L.A., Gordaliza, A., Matrán, C., Mayo-Iscar, A., 2008. A general trimming approach to robust cluster analysis. *Ann. Stat.* 36, 1324–1345. <https://doi.org/10.1214/07-AOS515>.
- Gilardelli, C., Stella, T., Confalonieri, R., Ranghetti, L., Campos-Taberner, M., García-Haro, F.J., Boschetti, M., 2019. Downscaling rice yield simulation at sub-field scale using remotely sensed LAI data. *Eur. J. Agron.* 103, 108–116. <https://doi.org/10.1016/j.eja.2018.12.003>.
- Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58 (3), 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7).
- Gracia-Romero, A., Kefauver, S.C., Vergara-Díaz, O., Zaman-Allah, M.A., Prasanna, B.M., Cairns, J.E., Araus, J.L., 2017. Comparative performance of ground vs. Aerially assessed rgb and multispectral indices for early-growth evaluation of maize performance under phosphorus fertilization. *Front. Plant Sci.* 8, 1–13. <https://doi.org/10.3389/fpls.2017.02004>.
- Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81 (2–3), 416–426. [https://doi.org/10.1016/S0034-4257\(02\)00018-4](https://doi.org/10.1016/S0034-4257(02)00018-4).
- Heremans, S., Dong, Q., Zhang, B., Bydekerke, L., Van Orshoven, J., 2015. Potential of ensemble tree methods for early-season prediction of winter wheat yield from short time series of remotely sensed normalized difference vegetation index and in situ meteorological data. *J. Appl. Remote Sens.* 9 (1), 097095. <https://doi.org/10.1117/1.JRS.9.097095>.
- Hornero, A., Hernández-Clemente, R., North, P.R.J., Beck, P.S.A., Boscía, D., Navas-Cortés, J.A., Zarco-Tejada, P.J., 2020. Monitoring the incidence of *Xylella fastidiosa* infection in olive orchards using ground-based evaluations, airborne imaging spectroscopy and Sentinel-2 time series through 3-D radiative transfer modelling. *Remote Sens. Environ.* 236, 111480. <https://doi.org/10.1016/j.rse.2019.111480>.
- Hunt, E.R., Cavigelli, M., Daughtry, C.S.T., McMurtrey, J.E., Walthall, C.L., 2005. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis. Agric.* 6 (4), 359–378. <https://doi.org/10.1007/s11119-005-2324-5>.
- Hunt, E.R., Doraiswamy, P.C., McMurtrey, J.E., Daughtry, C.S.T., Perry, E.M., Akhmedov, B., 2013. A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. *Int. J. Appl. Earth Obs. Geoinf.* 21, 103–112. <https://doi.org/10.1016/j.jag.2012.07.020>.
- Hunt, M.L., Blackburn, G.A., Carrasco, L., Redhead, J.W., Rowland, C.S., 2019. High resolution wheat yield mapping using Sentinel-2. *Remote Sens. Environ.* 233, 111410. <https://doi.org/10.1016/j.rse.2019.111410>.
- Hunt, M.L., Blackburn, G.A., Rowland, C.S., 2019b. Monitoring the sustainable intensification of arable agriculture: the potential role of earth observation. *Int. J. Appl. Earth Obs. Geoinf.* 81, 125–136. <https://doi.org/10.1016/j.jag.2019.05.013>.
- Jeong, J.H., Resop, J.P., Mueller, N.D., Fleisher, D.H., Yun, K., Butler, E.E., Timlin, D.J., Shim, K.M., Gerber, J.S., Reddy, V.R., Kim, S.H., 2016. Random forests for global and regional crop yield predictions. *PLoS One* 11, 1–15. <https://doi.org/10.1371/journal.pone.0156571>.
- Jin, X., Zarco-Tejada, P.J., Schmidhalter, U., Reynolds, M.P., Hawkesford, M.J., Varshney, R.K., Yang, T., Nie, C., Li, Z., Ming, B., Xiao, Y., Xie, Y., Li, S., 2021. High-throughput estimation of crop traits: a review of ground and aerial phenotyping platforms. *IEEE Geosci. Remote Sens. Mag.* 9, 200–231. <https://doi.org/10.1109/MGRS.2020.2998816>.
- Kayad, A., Sozzi, M., Gatto, S., Marinello, F., Pirotti, F., 2019. Monitoring within-field variability of corn yield using sentinel-2 and machine learning techniques. *Remote Sens.* 11. <https://doi.org/10.3390/rs11232873>.
- Kumhálová, J., Matějková, Š., 2017. Yield variability prediction by remote sensing sensors with different spatial resolution. *Int. Agrophys.* 31, 195–202. <https://doi.org/10.1515/intag-2016-0046>.
- Lambert, M.J., Traoré, P.C.S., Blaes, X., Baret, P., Defourny, P., 2018. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens. Environ.* 216, 647–657. <https://doi.org/10.1016/j.rse.2018.06.036>.
- Liu, L., Wang, J., Bao, Y., Huang, W., Ma, Z., Zhao, C., 2006. Predicting winter wheat condition, grain yield and protein content using multi-temporal EnviSat-ASAR and Landsat TM satellite images. *Int. J. Remote Sens.* 27, 737–753. <https://doi.org/10.1080/01431160500296867>.
- Lyle, G., Bryan, B.A., Ostendorf, B., 2014. Post-processing methods to eliminate erroneous grain yield measurements: review and directions for future development. *Precis. Agric.* 15, 377–402. <https://doi.org/10.1007/s11119-013-9336-3>.
- Maes, W.H., Steppe, K., 2019. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci.* 24, 152–164. <https://doi.org/10.1016/j.tplants.2018.11.007>.
- Migdall, S., Brüggemann, L., Bach, H., 2018. Earth observation in agriculture. In: *Satellite-Based Earth Observation*. Springer International Publishing, Cham, pp. 85–93. [https://doi.org/10.1007/978-3-319-74805-4\\_9](https://doi.org/10.1007/978-3-319-74805-4_9).
- Miralles, D.J., Richards, R.A., Slafer, G.A., 2000. Duration of the stem elongation period influences the number of fertile florets in wheat and barley. *Funct. Plant Biol.* 27, 931. <https://doi.org/10.1071/PP00021>.
- Mokhtari, A., Noory, H., Vazifedoust, M., 2018. Improving crop yield estimation by assimilating LAI and inputting satellite-based surface incoming solar radiation into SWAP model. *Agric. For. Meteorol.* 250–251, 159–170. <https://doi.org/10.1016/j.agrformet.2017.12.250>.
- Oguntunde, P.G., Lischheid, G., Dietrich, O., 2018. Relationship between rice yield and climate variables in southwest Nigeria using multiple linear regression and support vector machine analysis. *Int. J. Biometeorol.* 62, 459–469. <https://doi.org/10.1007/s00484-017-1454-6>.
- R Core Team, 2021. A language and environment for statistical computing.
- Ranieri, R., 2015. Geography of the durum wheat crop. *Pastaria Int* (6), 24–36.
- Rosegrant, M.W., Cline, S.A., 2003. Global food security: challenges and policies. *Science* (80-) 302, 1917–1919. <https://doi.org/10.1126/science.1092958>.
- Rouse Jr., J.W., Haas, R., Schell, J., Deering, D., 1974. Monitoring vegetation systems in the great plains with erts. *NASA Spec. Publ.* 351.
- Saruta, K., Hirai, Y., Tanaka, K., Inoue, E., Okayasu, T., Mitsuka, M., 2013. Predictive models for yield and protein content of brown rice using support vector machine. *Comput. Electron. Agric.* 99, 93–100. <https://doi.org/10.1016/j.compag.2013.09.003>.
- Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., Nelson, A., 2019. The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* 3, 430–439. <https://doi.org/10.1038/s41559-018-0793-y>.
- Segarra, J., Buchailot, M.L., Araus, J.L., Kefauver, S.C., 2020a. Remote sensing for precision agriculture: sentinel-2 improved features and applications. *Agronomy* 1–18. <https://doi.org/10.3390/agronomy10050641>.
- Segarra, J., González-Torralba, J., Aranjuelo, I., Araus, J.L., Kefauver, S.C., 2020b. Estimating wheat grain yield using Sentinel-2 imagery and exploring topographic features and rainfall effects on wheat performance in Navarre. Spain. *Remote Sens.* 12, 1–24. <https://doi.org/10.3390/rs12142278>.
- Serrano, L., Filella, I., Peñuelas, J., 2000. Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. *Crop Sci.* 40, 723–731. <https://doi.org/10.2135/cropsci2000.403723x>.
- Stas, M., Van Orshoven, J., Dong, Q., Heremans, S., Zhang, B., 2016. A comparison of machine learning algorithms for regional wheat yield prediction using NDVI time series of SPOT-VGT. In: *2016 5th Int. Conf. Agro-Geoinformatics, Agro-Geoinformatics*, pp. 16–20. doi: 10.1109/Agro-Geoinformatics.2016.7577625.

- Sutton, M.A., Howard, C.M., Erisman, J.W., Billen, G., Bleeker, A., Grennfelt, P., Van Grinsven, H., Grizzetti, B., 2011. The European nitrogen assessment: sources, effects and policy perspectives. Cambridge University Press. <https://doi.org/10.1002/met.1290>.
- Tesfaye, A.A., Osgood, D., Aweke, B.G., 2021. Combining machine learning, space-time cloud restoration and phenology for farm-level wheat yield prediction. *Artif. Intell. Agric.* 5, 208–222. <https://doi.org/10.1016/j.aiia.2021.10.002>.
- Toscano, P., Castrignanò, A., Di Gennaro, S.F., Vonella, A.V., Ventrella, D., Matese, A., 2019. A precision agriculture approach for durum wheat yield assessment using remote sensing data and yield mapping. *Agronomy* 9. <https://doi.org/10.3390/agronomy9080437>.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0).
- Uno, Y., Prasher, S.O., Lacroix, R., Goel, P.K., Karimi, Y., Viau, A., Patel, R.M., 2005. Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data. *Comput. Electron. Agric.* 47, 149–161. <https://doi.org/10.1016/j.compag.2004.11.014>.
- van Ittersum, M.K., 2016. Crop yields and global food security. Will yield increase continue to feed the world? *Eur. Rev. Agric. Econ.* 43, 191–192. <https://doi.org/10.1093/erae/jbv034>.
- Vapnik, V., 2013. The Nature of Statistical Learning Theory. Springer Science & Business Media. [https://doi.org/10.1007/978-1-4757-3264-1\\_1](https://doi.org/10.1007/978-1-4757-3264-1_1).
- Villegas, D., 2001. Biomass accumulation and main stem elongation of durum wheat grown under mediterranean conditions. *Ann. Bot.* 88, 617–627. <https://doi.org/10.1006/anbo.2001.1512>.
- Viña, A., Gitelson, A.A., Nguy-Robertson, A.L., Peng, Y., 2011. Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sens. Environ.* 115, 3468–3478. <https://doi.org/10.1016/j.rse.2011.08.010>.
- Vincini, M., Frazzi, E., D'Alessio, P., 2008. A broad-band leaf chlorophyll vegetation index at the canopy scale. *Precis. Agric.* 9, 303–319. <https://doi.org/10.1007/s11119-008-9075-z>.
- Wang, L., Zhou, X., Zhu, X., Dong, Z., Guo, W., 2016. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop J.* 4, 212–219. <https://doi.org/10.1016/j.cj.2016.01.008>.
- Weiss, M., Baret, F., 2016. S2ToolBox Level 2 products: LAI, FAPAR, FCOVER - Version 1.1. Sentin. ToolBox Level2 Prod. 53.
- Weiss, M., Baret, F., Myneni, R.B., Pragnère, A., Knyazikhin, Y., 2000. Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. *Agronomie* 20, 3–22. <https://doi.org/10.1051/agro:2000105>.
- Weiss, M., Jacob, F., Duveiller, G., 2020. Remote sensing for agricultural applications: a meta-review. *Remote Sens. Environ.* 236, 111402. <https://doi.org/10.1016/j.rse.2019.111402>.
- Wolanin, A., Camps-Valls, G., Gómez-Chova, L., Mateo-García, G., van der Tol, C., Zhang, Y., Guanter, L., 2019. Estimating crop primary productivity with Sentinel-2 and Landsat 8 using machine learning methods trained with radiative transfer simulations. *Remote Sens. Environ.* 225, 441–457. <https://doi.org/10.1016/j.rse.2019.03.002>.
- Woodruff, D.L., Reiners, T., 2004. Experiments with, and on, algorithms for maximum likelihood clustering. *Comput. Stat. Data Anal.* 47, 237–253. <https://doi.org/10.1016/j.csda.2003.11.002>.
- Zadoks, J., Chang, T., Konzak, C., 1974. A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. <https://doi.org/10.1111/j.1365-3180.1974.tb01084.x>.
- Zhang, L., Traore, S., Ge, J., Li, Y., Wang, S., Zhu, G., Cui, Y., Fipps, G., 2019. Using boosted tree regression and artificial neural networks to forecast upland rice yield under climate change in Sahel. *Comput. Electron. Agric.* 166, 105031. <https://doi.org/10.1016/j.compag.2019.105031>.
- Zhou, X., Kono, Y., Win, A., Matsui, T., Tanaka, T.S.T., 2020a. Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches. *Plant Prod. Sci.* 00, 1–15. <https://doi.org/10.1080/1343943X.2020.1819165>.
- Zhou, X., Wang, P., Tansey, K., Zhang, S., Li, H., Tian, H., 2020b. Reconstruction of time series leaf area index for improving wheat yield estimates at field scales by fusion of Sentinel-2, -3 and MODIS imagery. *Comput. Electron. Agric.* 177, 105692. <https://doi.org/10.1016/j.compag.2020.105692>.
- Zhou, X., Zheng, H.B., Xu, X.Q., He, J.Y., Ge, X.K., Yao, X., Cheng, T., Zhu, Y., Cao, W.X., Tian, Y.C., 2017. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS J. Photogramm. Remote Sens.* 130, 246–255. <https://doi.org/10.1016/j.isprsjprs.2017.05.003>.





## Discussió general

### 1. Estimació de rendiment del blat amb teledetecció

Un dels majors reptes dels models empírics d'estimació de rendiment és provar l'aplicabilitat dels models en diferents anys i camps de cultiu, és a dir provar la seva robustesa. Això permet avaluar la fiabilitat i reproductibilitat d'aquests models a través del temps, o com a mínim il·lustrar el funcionament dels índexs de vegetació o variables biofísiques (en el seu sentit fisiològic i agronòmic) en relació amb el rendiment. En aquest cas en els capítols 3 i 5 he vist que tant en anys diferents com en diferents camps de blat els models han pogut estimar de manera correcta el rendiment. Pel que fa al capítol 3, el model multilineal dividit entre el conjunt de dades d'entrenament, corresponent a la temporada 2018, i el conjunt de dades de validació, corresponent a la temporada 2019, va revelar l'eficiència d'aquest enfocament durant almenys un període de dos anys a Navarra. Això és poc comú perquè generalment aquests models empírics es desenvolupen i validen amb dades del mateix any (Lambert et al., 2018; Schwalbert et al., 2018), cosa que succeeix per la dificultat d'obtenir dades de diverses temporades, i, per tant, dona robustesa a les tècniques desenvolupades en aquesta tesi.

D'ençà que l'any 2018 el ple desplegament de la constel·lació de satèl·lits Sentinel-2 va ser efectiu, s'ha pogut assolir un millor monitoratge dels estadis fenològics del blat gràcies a la freqüència de cinc dies del Sentinel-2. Aquesta característica ha permès millorar els models d'estimació com mostren els resultats presentats en els capítols 3 i 5. Amb altres satèl·lits era poc probable que la predicció del rendiment de blat es pogués basar en diferents estadis fenològics perquè la freqüència era més baixa, com en el cas del satèl·lit Landsat 8 (imatges cada dues setmanes).

Els resultats suggereixen que l'espigat és l'etapa fenològica més adequada per estudiar les relacions empíriques entre els índexs de vegetació o paràmetres biofísics amb el rendiment del blat (Capítol 3 i 5). Això és coherent amb altres mesures espectrals per estimar el rendiment de blat en estadis fenològics similars, com els resultats obtinguts per Fernandez-Gallego et al. (2019). A més, aquesta etapa també s'ha descrit com a òptima per a la predicció del rendiment amb teledetecció de drons i condicions de cultiu equivalents (Hassan et al., 2019). Així doncs centrar-se en la fenologia per a l'estimació del rendiment és possible fent servir imatges obtingudes amb Sentinel-2 i no només amb dron o avió. Tanmateix, tot sovint, per poder aplicar manejos adequats abans de la maduresa i la recol·lecció del gra, l'estimació del rendiment de blat en estadis primerencs és molt rellevant. Una estimació adequada en estadis primerencs permet tenir més marge per aplicar un maneig específic i poder millorar el rendiment final.

Un altre factor rellevant per poder estimar el rendiment és quin ús es fa del valor dels píxels detectats durant el creixement. En el capítol 3 he vist que les mitjanes de valor màxim i mínim de píxels és la combinació estadística més adequada per estimar el rendiment en aquest cas. Això s'explica perquè la mateixa heterogeneïtat dels camps fa que diferents zones tinguin diferències lleugeres en el desenvolupament, així doncs poder capturar el moment màxim de desenvolupament, per exemple quant a activitat fotosintètica o quantitat de biomassa, permet posteriorment relacionar-ho més finament amb el rendiment. Això millora altres aproximacions estadístiques com per exemple la mitjana de píxel per camp en la imatge presa. Així doncs, per estudiar el rendiment, explorar l'evolució del cultiu a través del temps pot proporcionar un millor resultat en comparació amb les dades mitjanes dels píxels d'una sola imatge o índex de vegetació.

Malgrat la importància d'estimar el rendiment amb antelació, centrar-se en el període fenològic d'espigat proporciona correlacions relativament més altes amb el rendiment, tot explicant més del 80% de la variabilitat (capítols 3 i 5). És a dir, segregar períodes fenològics és un enfocament útil per calcular estimacions empíriques de rendiment a diferència d'usar valors de píxels sense considerar explícitament l'estadi fenològic. Estimar l'estadi fenològic amb l'acumulació de temperatura (growing degree days) i centrar-se en el període específic, espigat per exemple, facilita el processament i dona resultats precisos. La resolució espacial del Sentinel-2 permet una avaluació fina dels resums estadístics dels píxels (formulats en diferents índexs de vegetació o com a bandes soles) i, en conseqüència, permet fer estimacions més precises. Una altra manera de poder explorar al màxim la relació entre les dades de teledetecció i rendiment és poder tenir la variabilitat de rendiment dins de la parcel·la i relacionar-ho amb el rendiment, com en el cas de les dades de recol·lectores georeferenciades (capítol 5).

En el cas de bases de dades de gran envergadura (centenars d'entrades) vaig observar, capítol 5, que el paràmetre biofísic LAI (índex de l'àrea de la fulla) tenia una estimació del rendiment lleugerament millorada en comparació amb les bandes Sentinel-2 soles i els índexs de vegetació. En aquest cas emprant models de ML, concretament *random forest* com el més efectiu. La reducció del percentatge d'error (15 %RMSE) en el model derivat del LAI en comparació amb el desenvolupat amb índexs de vegetació (16,2 %RMSE) concorda amb altres estudis que van mostrar una millora amb l'ús del LAI per estimar el rendiment en cereals i d'altres cultius (Zhou et al., 2020). No obstant això, els índexs de vegetació constitueix un enfocament senzill per extreure informació a partir de dades de teledetecció i s'han relacionat efectivament amb el rendiment dels cultius (Gracia-Romero et al., 2017; Liu et al., 2006). Els índexs de vegetació són

fàcils d'utilitzar, però les dades obtingudes a partir de RTM són més robustes perquè capturen els canvis de reflectància en les etapes de creixement del cultiu i les arquitectures canviants de les plantes (Viña et al., 2011). En aquest sentit, el LAI es va calcular fent servir models pre-establerts de RTM. Aquests són capaços de capturar informació millorada dels canvis dels cultius (Wolanin et al., 2019). A més, a diferència de la majoria dels índexs de vegetació, que se saturen en determinades etapes de cultiu (Bannari et al., 1995), el LAI no presenta una elevada saturació.

Pel que fa a les bandes Sentinel-2, hi ha una lleugera millora dels models d'estimació de rendiment quan s'utilitza LAI (%RMSE 15.4) que les bandes soles (%RMSE 15.8). Altres autors (Hunt et al., 2019) han argumentat que els algorismes de ML són els que extreuen més informació espectral a causa de la seva capacitat de capturar interaccions complexes entre variables, i per tant es poden vincular exitosament amb el rendiment. No obstant això, podem argumentar que els models híbrids que utilitzen RTM avançats per descriure paràmetres biofísics com el LAI i posteriorment es relacionen amb el rendiment per alimentar models de ML poden fins i tot millorar i afegir robustesa als models. Aquests enfocaments s'han utilitzat principalment en el fenotipatge (Jin et al., 2021), però com s'observa aquí també poden ser operatius pel seu ús en l'estimació, dins de la parcel·la, de les diferències de rendiment en blat.

Les dades obtingudes en diverses dates milloren la capacitat d'estimar la variabilitat del rendiment dins del camp. En els tres casos estudiats (bandes soles de Sentinel-2, LAI i índex de vegetació), l' $R^2$  era més alt i RMSE més baix quan s'utilitzaven dades de Sentinel-2 de diferents dates (Taula 3, capítol 5). No obstant això, sovint la resolució temporal completa de les imatges de Sentinel-2 no està disponible completament perquè la disponibilitat d'imatges per a les regions estudiades depèn de si el cel està ennuvolat. Malgrat les limitacions, es van poder analitzar i seguir les etapes fenològiques en diverses imatges durant la temporada de creixement del blat. El LAI va ser la variable que millor es va ajustar al rendiment (un error % RMSE més baix) en comparació amb les bandes de Sentinel-2 i els índexs de vegetació sols. Això és coherent amb el fet que el LAI és un bon indicador de la fotosíntesi potencial i que els estadis fenològics estudiats corresponen a les etapes reproductives crucials per definir el rendiment en blat (Miralles et al., 2000; Villegas, 2001).

En el cas dels índexs de vegetació i de les bandes soles mostren resultats menys definits. Al fer servir centenars de dades les mateixes bandes soles de Sentinel-2 es poden combinar per dotar d'un millor entrenament a l'algoritme en comptes de calcular només un índex concret amb unes bandes limitades. Tanmateix, l'ús de les bandes soles pel model, tot i que equivalent, o fins i tot superior que els índexs de vegetació, no permet descriure un sentit fisiològic ni guiar exactament

quin maneig pot ser més adequat, com sí que es pot extrapolar de les dades derivades d'índex de vegetació. És una de les limitacions del *black box* del ML.

En aquest sentit, l'estimació del rendiment abans que es completi tota la temporada del cultiu, al voltant de març a maig, pot ajudar els agricultors a orientar les decisions de gestió agrícola i assegurar una millor collita. Diversos autors han relacionat imatges de Sentinel-2 amb la fenologia per estimar el rendiment de blat en diverses resolucions (Fieuzal et al., 2020; Toscano et al., 2019) i permetre una gestió millorada dins del marc de l'agricultura de precisió. Les característiques del Sentinel-2 i el desenvolupament creixent de la modelització enfocada a l'estimació del rendiment fa que cada cop més l'agricultura de precisió sigui assequible i rendible i, per tant, gairebé operativa (Weiss et al., 2020). És doncs cabdal definir el seu potencial ús en el marc de la sostenibilitat agroecològica a través de la infraestructura de dades de teledetecció satel·lital pel cas mediterrani.

## 2. Estimació de qualitat del blat amb teledetecció

Per estimar el percentatge de nitrogen en els grans de blat, els models desenvolupats en aquesta tesi utilitzant sensors multiespectrals en drons han resultat més eficients que els models en el visible, RGB (a escala de terra i dron), així com els models desenvolupats amb un sensor multiespectral àmpliament utilitzat (GreenSeeker) a nivell de terra. En validar els models d'estimació del contingut de nitrogen en el gra de blat amb dades multiespectrals a escala de dron, els millors resultats van mostrar (Figura 5, capítol 4) un  $R^2$  de 0.42 i RMSE de 0.18% a l'estadi fenològic d'espigat, que concorda amb la investigació de Zhao et al. (2019). No es van observar, tanmateix, correlacions significatives en els models construïts amb dades RGB a nivell de terra ni dron. S'ha demostrat que una alta resolució espectral a la regió del visible fotosintèticament actiu és capaç d'avaluar el contingut de nitrogen dels cultius (Hansen and Schjoerring, 2003), no obstant això, hem vist que el contingut de nitrogen en el gra no és avaluable quan s'utilitzen mesures de reflectància de banda ampla proporcionades per càmeres RGB comercials, independentment de la seva alta resolució espacial (Buchaillot et al., 2019).

Per tant, podem argumentar que el model construït amb dades multiespectrals a escala de dron ha permès estimar el contingut de nitrogen en el gra de blat fruit de la seva major resolució espectral. L'instrument multiespectral utilitzat a nivell de terra només mesura la reflectància vermella i infraroja propera (computada com NDVI) i no va mostrar correlacions significatives. Això podria estar relacionat, primer, amb la manca de relació entre un índex sensible a la biomassa com el NDVI i el contingut de nitrogen en el gra que està relacionat amb el contingut

de nitrogen a la planta. A més, la saturació del NDVI, que succeeix en cobertes d'alta densitat de vegetació, com en l'estadi d'espigat, no contribueix a generar un indicador de la planta relacionat amb el contingut de nitrogen en el gra. El GreenSeeker també té limitacions de resolució espacial en comparació amb els sensors emprats en els drons (Taula 2, capítol 4).

Les dades multiespectrals detectades remotament amb dron tenen diversos avantatges sobre les dades RGB. Observem que entre els tres índexs de vegetació que millor s'han relacionat amb el % de nitrogen en el gra (TCARI, Clred-edge i EVI) (Taula 5, capítol 4), EVI utilitza bandes RGB i NIR, mentre que TCARI i Clred-edge, a més d'aquestes, utilitzen una banda de l'espectre a la vora del vermell (700 nm). Els avantatges dels índexs que usen bandes a la vora del vermell per determinar el nitrogen en els grans s'han observat en diversos estudis (Guo et al., 2017; Prey and Schmidhalter, 2019), com també es corrobora en aquesta tesi. L'EVI és un índex sensible a la biomassa, que, no obstant això, s'ha utilitzat amb èxit en estimacions de concentració de nitrogen de les plantes (Zheng et al., 2018), també en l'estimació del contingut de proteïnes en el gra de blat (Tan et al., 2020). L'EVI té unes característiques millorades per evitar influències atmosfèriques i del sol, cosa que el fa més sensible a les variacions de la vegetació en comparació amb altres índexs (Huete et al., 1999).

En aquest sentit, els resultats suggereixen que els sensors multiespectrals, sobretot en la vora del vermell, augmenten l'eficàcia de l'estimació del contingut de nitrogen en el gra en comparació amb les dades en el visible de tres bandes, vermell-verd-blau. La part de l'espectre a la vora del vermell és altament sensible als canvis de clorofil·la i, per tant, es pot utilitzar com a indicador del contingut de nitrogen (Guerif et al., 2007; Segarra et al., 2020). L'índex TCARI també s'ha utilitzat amb èxit per estimar el contingut de proteïnes en el gra de cereals (Pettersson and Eckersten, 2007), així com el Clred-edge a escala de dron amb un  $R^2$  reportat de 0,50 (Zhou et al., 2021). Podem argumentar que la resolució espectral (bandes en la vora del vermell) és especialment rellevant quan l'objectiu de detecció, el % de nitrogen en els grans, es detecta indirectament i és altament dependent de la informació espectral relacionada amb la clorofil·la de la planta i la remobilització del nitrogen cap als grans en relació a l'estadi fenològic.

En aquest sentit, he observat que la fenologia és central per a l'estimació del contingut de nitrogen i, per tant, de proteïna en el gra (Fig. 3, Fig. 5, capítol 4). A l'espigat, el model d'estimació multiespectral va arribar a un  $R^2$  de 0,42, mentre que a l'ompliment de gra va disminuir a 0,29. Això concorda amb la conceptualització que l'acumulació de nitrogen del blat en els grans comença a produir-se en omplir el gra en el moment en el qual el nitrogen es remobilitza. És a dir, la concentració de nitrogen de les plantes comença a descendir en l'etapa d'ompliment del

gra perquè es concentra en forma de proteïnes als grans. Així doncs, els resultats suggereixen que els índexs de vegetació sensibles a la concentració de nitrogen a l'espigat poden ser un indicador del contingut de nitrogen en el gra de blat, com també han observat Dupont i Altenbach (2003). En blat, entre el 60 i el 95% del nitrogen es remobilitza de les fulles i tiges als grans, sent les fonts més importants del contingut de nitrogen en el gra (Palta i Fillery, 1995; Van Sanford i MacKown, 1986) i per tant de proteïna. Els resultats de l'estudi mostren que és possible monitoritzar el contingut de nitrogen en el gra a través d'aquests índexs de vegetació sensibles a la clorofil·la si es detecta el contingut de nitrogen vegetal abans de l'ompliment del gra. Aquesta etapa fenològica correspon a l'acumulació màxima de nitrogen (Kong et al., 2016; Papakosta and Gagianas, 1991) en les plantes de blat abans de remobilitzar el nitrogen en grans.

En general, el contingut de nitrogen en el gra és més alt a les zones mitjana i sud de Navarra, que corresponen a zones mediterrànies més seques, que no pas a la zona nord de Navarra que té menys estrès hídric. Això està probablement relacionat amb una mida de gra més petita, causada per l'estrès de la sequera durant l'ompliment de gra, a les zones més al sud. En aquest sentit, el  $\delta^{13}\text{C}$  i el contingut en nitrogen en el gra estaven correlacionats positivament (Fig. 4, capítol 4) i servien per explicar les característiques agroclimàtiques vinculades a la fisiologia del blat i el contingut de nitrogen en el gra.

Aquest estudi demostra que el model a escala de dron pot ser escalat a satèl·lit, en el cas de Sentinel-2, tot emprant bandes equivalents i estimacions del contingut de nitrogen en el gra a escala de camp de cultiu. Per tant, he vist que el model aplicat al llarg de dues temporades de cultiu va assolir un coeficient  $R^2$  de 0,40 i RMSE de 0,29% (Fig. 7, capítol 4), que es troba en el mateix rang que l'obtingut amb la validació a les parcel·les experimentals amb les dades multiespectrals a escala de dron ( $R^2$  de 0,42 i RMSE de 0,18 %, Fig. 6, capítol 4). El RMSE és lleugerament més alt quan el model s'aplica a escala de camp; tanmateix, aquests resultats s'alineen amb altres estudis que van utilitzar sensors muntats en drons per obtenir dades espectrals, construir models i aplicar-los a dades de Sentinel-2 equivalents (Revill et al., 2020, 2019). Pel cas d'estimació del contingut de nitrogen en el gra de blat (Zhao et al., 2019) diversos models han mostrat un rang de RMSE de 0,19–0,53%. Els models amb dades reals a escala de camp requereixen un treball de camp entretingut i cal cobrir grans extensions. Alternativament, la construcció de models en parcel·les experimentals més petites facilita el procés de modelatge i permet extrapolar-los a escala de satèl·lit. Els resultats d'aquesta tesi confirmen que l'ús de sensors multiespectrals en drons amb bandes espectrals equivalents de Sentinel-2 permet l'estimació del contingut de nitrogen en el gra de blat també a escala de camp de cultiu.



Així doncs, els models desenvolupats en el capítol 4 han estat capaços d'estimar el contingut de nitrogen en el gra de blat amb aplicacions tant en parcel·les experimentals com en camps agrícoles. La comparativa innovadora entre diferents plataformes de detecció i estadis fenològics estableix les bases per a una millora del fenotipat orientat a la qualitat i, alhora, del monitoratge dels cultius. Sovint, els trets de qualitat han estat poc considerats en els programes de millora (Pronin et al., 2020; Zhu et al., 2001) i el desenvolupament de models de teledetecció pot ajudar a identificar les línies amb millors característiques (pel que fa a qualitat del gra) en poblacions genèticament diverses de blat. En aquest sentit, la variació utilitzada per a aquest estudi va ser del 2,07 al 2,74 % del contingut de nitrogen en el gra de blat, tot i això, podria ser més àmplia (1,2-3,7%) si s'afegís una major diversitat de genotips, des de varietats modernes fins a ecotips.

Per tant, el model és específic per un rang de contingut de nitrogen en el gra de blat. Tot i que el  $R^2$  assolit és inferior a 0,5, la importància dels models i els resultats a la validació demostren la seva eficàcia en el fenotipatge de parcel·les experimentals i el monitoratge amb satèl·lit del contingut de nitrogen en el gra. Actualment, aquests models no poden determinar el contingut de nitrogen en el gra de blat amb una precisió molt alta. Per a això encara es necessitarien mesures destructives o sensors híperespectrals. No obstant això, els models obtinguts poden orientar els agricultors sobre la qualitat del gra després d'un vol de dron abans de la collita i també podrien ser potencialment utilitzats en l'agricultura de precisió amb satèl·lit per detectar àrees de concentració baixes i altes de contingut de nitrogen en el gra de blat en un camp.

Això és especialment rellevant en la producció agroecològica perquè la disponibilitat de fertilitzant és limitada i generalment el % de nitrogen en gra menor, un millor monitoratge i gestió dels inputs en el marc agroecològic ha de permetre mantenir un nivell suficient de contingut de nitrogen en el gra de blat. Aquesta informació podria contribuir a optimitzar l'ús dels recursos prioritzant una major fertilització d'alta qualitat en zones pobrament fertilitzades i ajudar així a avançar cap a la sostenibilitat agrícola. A més, a escala d'agroecosistema, amb l'ús d'imatges Sentinel-2 la fiabilitat actual dels models podria proporcionar suport per desenvolupar mapes regionals d'estimació del contingut de nitrogen en el gra de blat. Aquests mapes poden contribuir a comprendre les característiques ambientals i del sòl que afecten el contingut de nitrogen en el gra de blat, així com ajudar els agricultors i les institucions a gestionar les terres de cultiu de manera sostenible. La resolució espacial del Sentinel-2 fa que la seva aplicació directa com a plataforma de fenotipatge en programes de millora sigui inadequada, però per a l'aplicació a l'escala de camp de cultiu el model desenvolupat en aquest estudi podria ser-hi implementat.

### 3. Efectes ambientals i del paisatge en el rendiment del blat

Les característiques topogràfiques i les precipitacions representen conjuntament entre l'11% i el 20% de la variació espacial observada en el rendiment de blat, capítol 3. Aquest rang percentual és coherent amb altres estudis que tracten aquests factors. Pel que fa als efectes topogràfics sobre el blat, a escala de camp de cultiu s'ha estimat que el pendent i l'altura representen entre el 5% i el 42,5% de variació en el rendiment de blat en zones centreeuropees (Kumhálová i Moudrý, 2014), mentre que en un altre estudi als EUA s'ha vist que representa entre el 39% i el 65% de la variació en el rendiment (Green et al., 2007). A més, Basso et al. van demostrar que en el context mediterrani l'orientació, l'altura del camp i el pendent estan directament relacionats amb el rendiment del blat (Basso et al., 2009) i que la distribució de la pluja és un factor cabdal (Basso et al., 2012).

En aquest sentit, s'ha observat que les característiques topogràfiques afecten de manera més pronunciada la producció de blat en anys secs o en zones amb precipitacions naturalment irregulars (és a dir, terres mediterrànies de cultiu de secà) (Ferrara et al., 2010). Això explica el rang de variabilitat relativament ampli del rendiment de blat (capítol 3) conseqüència de les precipitacions i els efectes topogràfics. Es podria esperar doncs un major efecte de les característiques topogràfiques en els agrosistemes de blat ibèrics, també tenint en compte les seves característiques geomorfològiques i climàtiques heterogènies com el cas descrit en el capítol 2 a Castella.

En un escenari de canvi climàtic, singularment pronunciat en la Mediterrània, aquestes característiques mereixen una especial atenció. En aquest sentit, les relacions entre la topografia i el rendiment de blat són innovadores perquè han estat poc estudiades (Kravchenko and Bullock, 2000), segurament a causa de la dificultat d'obtenir prediccions precises de rendiment per a múltiples camps. Això, però, la teledetecció ho facilita. Un altre factor és també la tendència de configurar paisatges agrícoles homogenis des de l'agricultura convencional. Aquesta tendència ha prioritzat establir produccions intensives en zones més fàcils de llaurar i, per tant, aquells paisatges agrícoles més complexos no han centrat l'atenció de l'acadèmia i han esdevingut marginals. Pel que fa a centrar-se en dades topogràfiques i climàtiques d'escala regional, estudis previs han definit zones d'idoneïtat de cultiu (Fekadu and Negese, 2020; Motuma et al., 2016; Pilevar et al., 2020); però sovint no hi han afegit dades de rendiment o la mateixa estimació dels efectes d'aquests factors en el rendiment del cultiu.

En l'estudi descrit en el capítol 3, es fan servir dades topogràfiques i climàtiques de Navarra per explorar geogràficament els seus efectes en el rendiment mitjançant imatges de Sentinel-2. Estudar els impactes climàtics i topogràfics sobre el rendiment a una escala més gran pot proporcionar noves idees per millorar la gestió de les terres de cultiu i alhora avançar cap a una millor avaluació agroecològica del paisatge. Tanmateix, en el cas del capítol 3, factors com el tipus de sòl, l'evapotranspiració, la radiació i la fertilització o les pràctiques agrícoles, entre d'altres, no varen poder ser avaluats amb les tècniques de teledetecció ni les bases de dades emprades. Aquests factors podrien ajudar a explicar la variabilitat de rendiment en major mesura, tenint en compte també l'important efecte del genotip en la producció. En el capítol 2 sí que s'han afegit alguns d'aquests factors per poder descriure millor els efectes ambientals i del paisatge.

#### 4. Mapatge de cultius amb teledetecció i sistemes d'informació geogràfica

La constel·lació de satèl·lits Sentinel-2 és una bona eina per monitoritzar l'activitat agrícola (Segarra et al., 2020a), com ja he descrit, però també per fer mapes dels tipus de cultius. Això té una gran incidència en l'avaluació de les polítiques agrícoles de la Unió Europea (Defourny et al., 2019). A més, aquests avenços tecnològics poden guiar polítiques per promoure pràctiques agrícoles sostenibles com rotacions de cultius i monitoritzar la seva eficàcia a diferents escales: des del camp fins al paisatge. Les pràctiques agrícoles produeixen el paisatge i tenen efectes a llarg termini amb conseqüències considerables per a la integritat mediambiental (Rounsevell et al., 2003). Per tant, els canvis en la gestió agrícola tenen un ampli impacte en l'ecosistema i uns clars efectes socioeconòmics. L'actual infraestructura de dades satel·lital permet controlar aquests aspectes i monitoritzar pràctiques sostenibles.

La classificació d'imatges de satèl·lit amb algorismes de ML permet detectar el tipus de cultius i tenir un mapa del paisatge agrícola força precís. Per tal d'obtenir mapes de cultius, diverses estratègies poden ser útils. En el cas de la detecció de parcel·les de blat a Navarra (Capítol 3) he aprofitat que l'ordi, com a cereal similar al blat i àmpliament cultivat, té una senescència més primerenca que no pas el blat. Així doncs, emprant imatges tardanes en les que l'ordi hagi entrat en senescència però el blat no es possible diferenciar-los millor. Aquesta aproximació ha demostrat ser eficient per classificar parcel·les de blat (capítol 3). Novament, aquest mapatge aprofita la freqüència temporal del Sentinel-2 per diferenciar cultius similars i altres usos del sòl. Els resultats de les superfícies agrícoles obtingudes en el capítol 3 (taula 9) són coherents amb les estadístiques oficials del govern regional de Navarra.

La resolució espacial de 10 m del Sentinel-2 va permetre un mapatge precís i la possibilitat d'aplicar en els camps de blat classificats els models d'estimació de rendiment (capítol 3). El model d'estimació del rendiment del gra es va aplicar a les parcel·les classificades sense considerar diferències genotípiques dins de la regió (tant pel que fa a la classificació com a les qüestions de modelització del rendiment) a causa del fet que per a la temporada 2017-2018 el 80% dels camps de tot Navarra van utilitzar Camargo o Marcopolo (Goñi and Caballero, 2017). Per tant, se suposava una homogeneïtat genotípica general del blat en tota la regió. Això assumeix la plasticitat dels genotips elit a adaptar-se a un rang ambiental ampli, tanmateix la heterogeneïtat dels paisatges i les característiques geomorfològiques diferenciades a Navarra requeririen un increment de la diversitat de genotips emprats. Sobretot, com descrit en el capítol 1, per als ambients marginals, és a dir zones complexes no altament productives. Els marges de rendiment de les varietats desenvolupades en ambients més productius no tenen un gran marge

de millora en les zones marginals. Així doncs, adaptar els recursos genètics a la complexitat del paisatge per poder tenir millors rendiments i manejos més sostenibles és cabdal.

En general la classificació del blat amb imatges del Sentinel-2 presentada en el capítol 3 és precisa. Tanmateix la precisió de la classificació del blat a la zona sud de Navarra és lleugerament més baixa. Això podria ser a causa de la superfície limitada dedicada a aquest tipus de cultiu en comparació amb altres zones de Navarra. En la zona sud es va observar una major dificultat per classificar amb precisió els camps de blat amb 13 classificacions errònies (Taula 8, Capítol 3). També podria ser conseqüència de la menor grandària dels camps d'aquesta zona sud. Les classificacions errònies van ocórrer majoritàriament entre píxels de blat i guaret en totes les tres zones estudiades.

Pel que fa a les pràctiques sostenibles, en dues temporades de cultiu a Castella he observat, mitjançant mapes de cultius obtinguts a partir de dades de Sentinel-2, una rotació escassa de cultius a la regió (Capítol 2). La majoria dels camps que conreen blat conreen posteriorment o bé blat de nou (amb el 18,5% dels casos) o un altre cereal com l'ordi, en el 34% dels casos. Això indica que més del 50% dels camps conreen el mateix cultiu o un cultiu relativament equivalent en dues temporades consecutives. Això suggereix una alta dependència dels pesticides i herbicides, ja que en escenaris de cultiu continuus, l'aparició de plagues i altres malures als cultius són comunes, especialment en sistemes d'alts input (Jalli et al., 2021). Però també s'ha demostrat que les rotacions de cultius milloren el rendiment del blat, els sòls i l'entorn a les zones mediterrànies (López-Bellido et al., 2020). En aquest sentit, múltiples serveis ecosistèmics poden beneficiar-se de rotacions de cultius per mantenir l'estructura del sòl i els microorganismes, millorar l'eficiència de l'ús d'aigua i nutrients i regular les plagues (Barbieri et al., 2017).

En aquest sentit, la producció agroecològica del blat millora l'estructura del sòl i evita l'erosió. A més, augmenta la biodiversitat salvatge de flora i fauna i té una petjada de carboni més petita que l'agricultura convencionals. Cal destacar que els mètodes agrícoles industrials, com el llaurat i l'ús de fertilitzants químics, contribueixen substancialment a l'erosió de les terres de cultiu, exercint així un impacte en un component crucial dels sistemes agrícoles. L'erosió del sòl i la gestió inadequada del sòl són especialment freqüents a les regions mediterrànies (Vanwalleghem et al., 2011). En el context de la producció mediterrània de blat de secà a la península Ibèrica, s'ha comprovat que la implementació de pràctiques de gestió de baixos inputs tenen efectes positius sobre les propietats del sòl (Kosmas et al., 1997; Martínez-Mena et al., 2020).

En aquest sentit, els sistemes agroecològics, que abasten diversos cultius com el blat, tenen entre un 13% i un 9% més d'eficàcia en l'emmagatzematge de carboni del sòl i nitrogen del sòl, respectivament, que el cultiu convencional de blat (Migliorini et al., 2014). A més, cap pràctica de cultiu augmenta el contingut de matèria orgànica del sòl (3,3%) i mostra una biomassa microbiana del sòl més elevada (71%) (Sapkota et al., 2012). Deixar els rostolls dels cultius en el camp, tal com se segueix en les pràctiques de conservació, és millor a l'hora de fer del sòl un embornal de carboni i representa una millora de l'estructura del sòl a través d'un enfocament agroecològic (Di Bene et al., 2016). La dimensió espacial i la presència d'elements naturals són factors clau per determinar la sostenibilitat de les interaccions biòtiques i abiòtiques dins dels sòls.

Definir zones on poder desenvolupar aquetes pràctiques agrícoles pel cas del blat sense comprometre els rendiments de manera desproporcionada és de gran importància. En aquest sentit, els resultats obtinguts en el capítol 2 a partir de dades georeferenciades edàfiques, climàtiques i topogràfiques permeten descriure espacialment les zones marginals a Castella i poden orientar on desenvolupar pràctiques agroecològiques que contribueixin a generar paisatges agrícoles sostenibles i productius. En aquest cas el 25% de zones de producció de blat marginals encaixen amb l'objectiu de tenir un 25% de la superfície agrícola en producció ecològica o si més no alternativa a la producció industrial convencional d'alts inputs. La classificació de la idoneïtat de blat ha estat precisa en més d'un 80% en comparació amb altres classificacions d'idoneïtat.

L'ús de conjunts de dades geolocalitzades que estiguin relacionades amb el creixement i el desenvolupament dels cultius permet definir espacialment els ambients agrícoles i la seva adequació. Això permet guiar la millora genètica o gestió agrícola. A més, permet descriure diverses condicions a una escala regional. En aquest sentit, altres estudis han tractat, fins a cert punt, aquesta qüestió i han utilitzat imatges de satèl·lit per extrapolar dades a escala del paisatge i estimar el rendiment (Lambert et al., 2018). Així com els trets de qualitat, com el descrit en el capítol 4, combinant diversos instruments de teledetecció i mesures fisiològiques i agronòmiques. El flux de dades actual permet definir ambients i zones marginals de conreu (que ocupen àrees importants), com en el cas del centre nord peninsular descrit en el capítol 2.

## 5. Un maneig sostenible del blat

En el Capítol 2, en condicions de secà, he observat que la producció de blat en ecològic ha produït aproximadament la meitat del rendiment obtingut en comparació amb la producció convencional. La diferència de rendiment observada en aquest estudi és més gran que la descrita per De Ponti et al. (2012), qui, després d'abordar una extensa revisió, va observar un 19% de diferència de rendiment entre l'ecològic i el convencional en el cas del blat. Per a la regió estudiada, l'avaluació oficial de rendiment màxim de blat per a l'agricultura ecològica és de 3354 kg/ha, mentre que en l'agricultura convencional és de 5504 kg/ha (<https://genvce.org/productos-genvce/informes/>), indicant una diferència aproximada del 40%. S'han suggerit que aquestes diferències de rendiment es redueixen dràsticament si el blat es cultiva en ambients marginals. Una altra qüestió és la varietat de blat triada. És habitual que els agricultors ecològics utilitzin varietats tradicionals (Ortman et al., 2023). Aquests genotips tenen habitualment un rendiment més baix que les varietats moderns.

Per a la Unió Europea, s'ha vist una manca general de varietats de cultiu específicament adaptades a les condicions ecològiques (Padel et al., 2021). O bé adaptades a agroambients específics (no només a manejos com l'ecològic), és a dir al mateix paisatge. Així doncs, caldria una millora genètica específica per a condicions marginals i de baixos inputs. En aquest sentit, en els resultats descrits en el Capítol 2, diversos factors afecten la menor producció de blat en ecològic. Per exemple: la ubicació dels camps ecològics es trobaven en les zones de producció marginal. En canvi, la ubicació dels camps de cultiu de blat convencional es trobaven en zones d'alt rendiment (Figura 5, Capítol 2). Així doncs, a part del maneig, el genotip i l'entorn en què es conrea el blat afecten significativament la producció agrícola. Això podria haver jugat un paper en l'augment de la diferència de rendiment en el cas del capítol 2 en comparació amb altres estudis i les estadístiques oficials.

Pel que fa als trets de qualitat, en el capítol 2 no he observat diferències clares pel que fa al contingut de proteïnes en el gra, el que suggereix que el maneig (règim d'aigua i fertilització) juntament amb l'elecció del genotip no va afectar aquest tret, excepte en els camps de regadiu i els fertilitzats químicament. En aquest cas, els valors més alts de proteïna en gra poden ser conseqüència de la fertilització intensiva de nitrogen associada a les condicions altament productives d'un cultiu de regadiu (Ibrahim et al., 2018). Per altra banda, els isòtops  $\delta^{15}\text{N}$  indiquen el cultiu de lleguminoses en la rotació de cultius prèvia al blat (Peoples et al., 2015) per als camps de secà en ecològic.



El blat en regadiu té un rendiment més alt en comparació amb qualsevol altre maneig, principalment perquè l'estat hídric en el cas mediterrani és el factor més limitant de la producció. Això concorda amb el menor (és a dir, més negatiu)  $\delta^{13}\text{C}$  de cereals en regadiu en comparació amb les condicions de secà, com també s'ha corroborat en altres estudis (Rezzouk et al., 2020). A més, recolza l'ús del  $\delta^{13}\text{C}$  com a indicador del total d'aigua (pluja més reg) rebut pel cultiu (Araus et al., 2003). No obstant això, hi ha una manca general de reg a tota la península Ibèrica (Berbel et al., 2019) a causa de la insuficient modernització del regadiu i les sequeres estacionals, tot sovint extremes en algunes zones. L'aigua disponible per al reg es fa servir normalment per altres cultius econòmicament més rendibles com l'alfals, el blat de moro pel bestiar o fins i tot la remolatxa sucrera. El canvi climàtic, tanmateix, pot obligar a revertir aquest escenari en relació al cultiu de blat. Així doncs, l'augment de les temperatures i la disminució de les precipitacions a la regió com a conseqüència del canvi climàtic (Calvin et al., 2023) empitjorarà les condicions per al blat de secà, fent més necessari el regadiu així com altres estratègies d'adaptació.

Un punt rellevant descrit al capítol 2 és que la fertilització amb fonts orgàniques, en aquest cas de digestat, permeten obtenir rendiments equivalents en comparació amb les parcel·les de blat fertilitzades químicament. Això concorda amb altres estudis en blat (Yan i Gong, 2010; Yang et al., 2020) que no van trobar diferències significatives entre les fonts de fertilització orgànica i química. Aquest resultat mostra un camí cap a pràctiques més sostenibles sense comprometre els rendiments ja que la fertilització química s'ha relacionat amb grans pèrdues de nutrients en el medi (eutrofització dels aqüífers i rius, emissions d'òxid nitrós i d'amoníac) (Tuomisto et al., 2012). A més, té efectes negatius sobre els sòls, cosa que afecta la sostenibilitat futura dels rendiments agrícoles i el medi ambient.

En aquest sentit, la Unió Europea té previst aconseguir el 25% del total de terres de cultiu en producció ecològica per a la propera dècada (2030), tal com s'afirma en la declaració de la Comissió Europea titulada "Pla d'acció per al desenvolupament de la producció ecològica a la UE" (2021) i he descrit prèviament. Per tant, trobar terres de cultiu aptes per aquest maneig pot ajudar a definir àrees amb diferències de rendiment reduïdes en comparar la producció ecològica amb la convencional. Les terres de cultiu marginals es troben sovint en entorns amb condicions agroecològiques que limiten la seva productivitat (Ceccarelli, 1994). Des de la implantació de l'agricultura industrial, aquestes terres de cultiu han estat abandonades progressivament en diverses regions mediterrànies (Jiménez-Olivencia et al., 2021; Segarra et al., 2023), tal com s'observa en el capítol 1 la superfície agrícola recentment abandonada a totes les ecoregions euromediterrànies és molt extensa. Per tant, mapejar la idoneïtat del blat com s'ha descrit en el

capítol 2, així com el monitoratge amb dades de teledetecció de les pràctiques agrícoles i els seus efectes en el cultiu (capítols 3,4, i 5) poden contribuir a l'aplicació de pràctiques agrícoles adequades a paisatges específics.

Observem, per exemple, una reducció mitjana del 12% en el rendiment del gra en utilitzar varietats antigues en contrast amb les modernes (Figura 5, capítol 1) sota en un ambient amb baixos inputs (De Vita et al., 2006; Dinelli et al., 2013; Iannucci et al., 2018; Stagnari et al., 2013). No obstant això, en condicions extremadament marginals, Carranza-Gallego et al. (2018) van observar millors rendiments de varietats antigues en condicions agroecològiques, especialment a causa de les estacions relativament més humides i, per tant, una major infestació de males herbes i una millora de la viabilitat de genotips més alts (és a dir, més antics). En termes generals, però, les varietats modernes superen les varietats més antigues, fins i tot en terres de cultiu extremadament marginals com les descrites per Wang et al. (2021). És a dir que cal una millora per ambients específics i manejos sostenibles si és vol mantenir uns marges de producció suficients en el marc agroecològic.

En general, en la producció convencional, la millora de producció de les varietats modernes en contrast amb les antigues és del 38% (Figura 5), com a mitjana descrita en la literatura científica analitzada en el capítol 1 (Carranza-Gallego et al., 2019; Migliorini et al., 2016, 2014; Stagnari et al., 2013). Això il·lustra que, sota la producció convencional, les varietats modernes tenen una diferència de rendiment més gran amb les varietats antigues, però aquest es redueix quan es cultiven en ambients marginals. A més, sota una gestió agroecològica en condicions ambientals mediterrànies relativament marginals, les diferències de producció entre agroecològiques i convencionals no existeixen amb varietats equivalents. Com descriu Pardo et al. (2011), a les regions semiàrides mediterrànies, no es va observar cap diferència de rendiment entre producció ecològica i convencional amb l'aplicació o no d'herbicides o la fertilització orgànica vs. química que també va ser observat per Dinelli et al. (2013).

Pel que fa a la viabilitat econòmica de la producció agroecològica del blat a l'Euromediterrani, diversos autors han demostrat que cal superar importants limitacions econòmiques. A diferència de l'agricultura convencional, Bux et al. (2022) van observar un benefici major d'un 55% en la producció de blat convencional que en ecològic a Itàlia, mentre que Pardo et al. (2009) va observar que a l'Espanya mediterrània semiàrida, ni les produccions de blat convencional ni l'ecològica són sostenibles econòmicament i necessiten ser subvencionades. No obstant això, un estudi dut a terme per Varia et al. (2021) ha demostrat que els sistemes ecològics que conreen les varietats tradicionals de blat dur a Sicília presenten un benefici econòmic notable en

comparació amb l'ús de varietats estàndard perquè es poden vendre a preus més alts i son preferides pels consumidors. Una alternativa és incrementar la rotació i diversificar els cultius en els paisatges mediterranis (Zimmerer et al., 2024).

## 6. La importància geogràfica

Les terres de cultiu abandonades en totes les ecoregions estudiades ocupen àrees extenses (Taula 2, capítol 1). En algunes d'aquestes zones hi podria tenir lloc una re-cultivació del blat amb enfocaments agroecològics alhora que es deixessin les terres més productives a enfocaments més productius. Al meu parer seria un punt entremig entre els postulats maximalistes de *landsparing* i *landsharing*, en el cas euromediterrani concretament els paisatges estan profundament intervinguts i la seva diversitat i sostenibilitat no és contradictòria amb l'acció de l'home (Blondel, 2006). Per tant, crec que els resultats obtinguts en el capítol 1 i 2 mostren prou evidència per argumentar que les terres de cultiu euromediterrànies poden sostenir la producció agroecològica de blat en almenys un 25% de les terres de cultiu.

Per tant, la identificació de terres de cultiu marginals pot ajudar a delimitar àrees agrícoles amb uns manejos específics. El principal factor que ha contribuït a la reducció de les terres de cultiu en les darreres dècades és la seva conversió a bosc com a resultat de l'abandonament. L'abandonament de les activitats agrícoles no millora significativament la conservació de la biodiversitat, augmenta el risc d'incendis i utilitza excessivament els recursos hídrics de les conques mediterrànies. Per tant, és rellevant planificar la recuperació d'algunes d'aquestes terres de cultiu per afrontar aquests reptes. L'extensió de l'agricultura del blat, per exemple, en aquestes zones hauria d'adherir-se a principis agroecològics dins d'un paisatge divers, certament també amb l'ús d'eines tecnològiques innovadores com la teledetecció satel·lital. L'imperatiu de preservar, incrementar i rehabilitar els paisatges agrícoles de cereals a la regió euromediterrània és evident sota l'escenari actual de canvi climàtic, la inestabilitat econòmica i el debat de la política agrària de la Unió Europea.

Mapejar la idoneïtat del cultiu de blat combinant dades de factors ambientals i d'ús del sòl pot contribuir a fer front a la reducció de la collita (Capítol 2). En un escenari de reducció del rendiment del blat, totes les terres de cultiu disponibles, ja siguin altament productives o marginals, hauran de ser conreades per a afrontar els reptes ambientals i socioeconòmics que planteja el canvi climàtic. Per tant, centrar-se en les condicions específiques de les zones de cultiu de blat requerirà cada vegada més atenció per aconseguir una producció suficient d'aliments i mantenir els ecosistemes dins de les seves respectives capacitats de càrrega.

La recerca ha tingut diversos enfocaments per abordar la producció de blat, des d'estudis genotípics fins a estudis fenotípics. En aquest sentit, el maneig dels cultius i la millora genètica han donat lloc a un valuós corpus de literatura científica i exemples pràctics sobre la selecció i obtenció de trets específics i pràctiques de gestió de cultius adaptades a la sequera, la salinitat o les plagues, per esmentar-ne alguns (Araus et al., 2021; Barmeier i Schmidhalter, 2016; Rezzouk et al., 2020). No obstant això, molts d'aquests estudis han estat generalment limitats geogràficament, també pel que fa al número estudiat d'ambients i estacions experimentals (Resende et al., 2021). Per superar algunes d'aquestes limitacions i utilitzar la majoria de dades disponibles, Xu (2016) va proposar entendre l'ambient com un nou element a detectar en relació a la producció del cultiu, un nou enfocament tècnic per ajudar a desxifrar els impactes ambientals en els cultius i superar algunes d'aquestes limitacions espacials. Actualment el medi agrícola es produeix a partir de sensors geolocalitzats i dades interpolades espacialment. Això incrementa les possibilitats de modelatge per ajudar a trobar condicions agrícoles òptimes i definir el maneig de les terres de cultiu.

Els estudis amb una perspectiva agroecològica que es duen a terme en el territori mediterrani generen un coneixement que permet produir l'espai d'un manera sostenible. L'absència d'aquest coneixement suposa un biaix i una disminució de les possibilitats de produir paisatges agrícoles sostenibles. En aquest sentit les zones a l'euromediterrània amb més mancances d'estudis en el marc agroecològic són la Grècia continental (0 i 0,04 estudis per 10.000 ha de cultiu de blat) i el centre peninsular ibèric (0,06 i 0,02 estudis per 10.000 ha de cultiu de blat) (Taula 1, capítol 1). Això és rellevant perquè el centre d'Espanya té la major zona de producció de blat comú a l'Euromediterrani. A més, té una superfície de cultiu de blat a les zones marginals extensa. Amb un 25% dels camps de cultiu en aquestes condicions marginals en algunes zones (com s'ha descrit en el capítol 2) s'hi podria desenvolupar una producció agroecològica.

En particular, el centre peninsular està constituït per una gran estepa cerealística que es beneficiaria d'una major comprensió geogràfica de la seva sostenibilitat i una millora de les pràctiques de maneig agrícola. Grècia, per l'altra banda, té una producció significativa de blat dur i una rica agrobiodiversitat de varietats de blat (Pecetti et al., 1992) que afavoriria la gestió agroecològica. Encara que l'agricultura tradicional està molt estesa a Grècia i podria semblar-se en alguns aspectes a l'agroecologia, l'adaptació al canvi climàtic requereix el desenvolupament de nous enfocaments. El sud d'Itàlia ha rebut més atenció acadèmica que el sud d'Espanya i Portugal. Mentre que el centre-nord d'Itàlia i el nord-est d'Espanya/sud-est de França són les àrees amb més estudis sobre l'agroecologia del blat (Capítol 1). En aquest sentit, la producció

agroecològica del blat varia molt a les regions mediterrànies (de 1,4 a 4,9 t/ha, (David et al., 2005)) i depèn de molts factors. Per tant, avaluar geogràficament els múltiples factors que s'han discutit aquí podria contribuir a millorar espacialment la producció sostenible del blat i destacar les àrees amb més manca de recerca de manera que es puguin dedicar més recursos a investigar-les.

## 7. Línies a seguir per establir agrosistemes de blat sostenibles

Produir aliments i salvaguardar els ecosistemes dins de les seves respectives capacitats de càrrega constitueix el pilar fonamental de les societats contemporànies. En aquest sentit, s'ha produït un debat ben viu entre la comunitat científica i els responsables polítics en els últims anys (Canfield et al., 2021; Giraldo and Rosset, 2018). Des d'aquesta perspectiva, alguns estudiosos sostenen que l'agroecologia té potencial com a forma alternativa per aconseguir una producció sostenible que alhora salvaguardi el medi ambient (Altieri and Nicholls, 2020; Espluga-Trenc et al., 2021) i d'altres consideren que no és així (Trewavas, 2001).

L'agroecologia implica una disminució dels inputs i la utilització dels serveis ambientals i de biodiversitat existents de manera natural, com la gestió de plagues agrícoles utilitzant enemics naturals en lloc de pesticides químics (Tilman et al., 2002). Però com es pot estendre i implementar aquest marc quan aquesta mateixa concepció dels recursos i el metabolisme terrestre no és comprès de forma suficient? Si la naturalesa és l'home prenent consciència de si mateix (Reclus, 1885), com formulà el geògraf Elisee Reclus fa més d'un segle, actualment l'ús de la tecnologia en la nostra circumstància ha d'anar cap a aquesta consciència. També en el cas de l'agricultura, desenvolupat en aquesta tesi, per produir paisatges agroecològics sostenibles.

Les aplicacions de dades de satèl·lit en l'agricultura de precisió s'han desenvolupat principalment en ambients agrícoles estandarditzats i relativament homogenis, com els que són habituals en l'agricultura industrial. No obstant això, la majoria de les activitats agrícoles a la Terra es duen a terme en paisatges agrícoles relativament heterogenis, amb policultius, arbres i cultius herbacis cultivats simultàniament dins del camp, i en camps relativament petits (Altieri and Nicholls, 2017). D'una manera especial això succeeix a la Mediterrània com a zona amb una composició de paisatges molt diversa (Zimmerer et al., 2024). Tanmateix la infraestructura de dades satel·lital així com les dades obtingudes amb sensors i altres instruments de recerca no s'ajusten a aquesta

diversitat de paisatges i esdevenen territoris també abandonats per la comprensió geogràfica de la tècnica actual.

En aquests paisatges agrícoles complexos, en comparació amb els satèl·lits de resolució mitjana, els satèl·lits d'alta resolució poden capturar millor la variabilitat dins dels camps i donar una visió de l'estat dels cultius per guiar-ne la gestió. És cert, però, que les imatges de Sentinel-2 s'han utilitzat per avaluar paisatges agrícoles heterogenis i diversos, com en el cas mediterrani desenvolupat en aquesta tesi amb resultats prometedors. No obstant això, les imatges de satèl·lits d'alta resolució com WorldView poden millorar la comprensió de l'agroecosistema, com ara el mapeig d'arbres dins del camp (Lambert et al., 2018) i una millor avaluació del cultiu principal. Així doncs, és cabdal entendre l'orografia juganera mediterrània, configurada per microclimes i amb una producció de paisatges bioculturals antiquíssima, per poder produir aquestes geografies de manera sostenible amb les actuals tecnologies digitals.

La disponibilitat actual de dades per satèl·lit, que té restriccions significatives per les imatges d'alta resolució, és una limitació important per als paisatges agrícoles heterogenis. A més, en moltes zones, la complexitat de les interaccions en la producció agrícola dificulta el seguiment de moltes variables. Per exemple, en el cas de controlar l'estrès biòtic a nivell regional i guiar enfocaments específics de gestió a escala de camp, Buchaillet et al. (2022) va observar que les imatges d'alta resolució de PlanetScope oferien millors resultats en comparació amb les imatges de Sentinel-2. No obstant això, també hi ha dificultats per descriure detalladament els agroecosistemes a partir d'imatges d'alta resolució, especialment a causa dels policultius i la gestió heterogènia que hi ha en els camps que es troben, tot sovint, en un mosaic agrícola heterogeni.

Per tant, en paisatges agrícoles heterogenis, l'agricultura de precisió té diverses limitacions per detectar el desenvolupament dels cultius. Aquesta limitació prové principalment de les mides reduïdes dels camps, així com els tipus de maneig i els recursos que tenen els agricultors per abordar-los. En aquest sentit, l'enfocament en l'agricultura digital desenvolupat en aquesta tesi, amb la comprensió de l'agroecologia del blat mitjançant dades de teledetecció, pot contribuir en aplicar l'agricultura de precisió en el marc agroecològic i planificar paisatges sostenibles.





## Conclusions

- Més d'un 80% de la variabilitat del rendiment en blat es pot estimar amb imatges de Sentinel-2 de forma precisa.
- La resolució temporal de Sentinel-2, freqüència de visita, permet extreure dades en estadis fenològics específics i millorar els models d'estimació de rendiment i qualitat.
- La resolució espacial de Sentinel-2 permet mapejar de forma precisa els tipus de cultius en ambients mediterranis.
- Més enllà dels índexs de vegetació, l'ús de tota la resolució espectral de Sentinel-2 així com paràmetres biofísics com el LAI obtinguts a partir de models invertits de RTM milloren l'estimació de rendiment en blat.
- La informació espectral en la vora del vermell (700 nm) millora l'estimació del contingut de nitrogen en gra de blat.
- Més d'un 40% de la variabilitat del contingut de nitrogen en gra de blat es pot estimar a partir d'imatges de Sentinel-2 en l'estadi d'espigat abans de remobilitzar el nitrogen en la planta.
- La superfície del blat a l'euromediterrània ocupa, aproximadament, un terç de les terres de cultiu i es pot mapejar amb imatges de Sentinel-2 i d'altres satèl·lits de manera precisa espacialment i temporal.
- La producció convencional en blat produeix entre un 19 i un 50% més que la producció ecològica sobretot pels ambients on es produeix, les varietats emprades i les pèrdues derivades de manques de maneig efectiu (fertilització, control de plagues, etc).
- Des de la perspectiva de les dades i programes de lliure accés, la integració de la infraestructura de dades satel·lital i l'agricultura de precisió en la producció agroecològica pot ajudar a reduir la diferència de producció amb l'agricultura convencional.
- L'orografia i la distribució de pluges determinen fins a un 20% el rendiment del blat en condicions mediterrànies.
- Un 25% de la superfície de cultiu en la zona nord de Castella té unes condicions marginals que permetrien desenvolupar-hi manejos agroecològics pel cas del blat.

-En ambients agrícoles marginals la diferència de producció entre varietats modernes i tradicionals és redueix substancialment fins a un màxim de 12%, caldria doncs fer una millora genètica de varietats de blat per els ambients marginals mediterranis.

-El modelatge del rendiment del blat a partir d'algoritmes de ML (aprenentatge automàtic) permet obtenir mapes de variabilitat de rendiment i entendre la heterogeneïtat de la parcel·la.

-L'ús de dades multispectrals obtingudes amb drons amb bandes equivalent al Sentinel-2 permet escalar les estimacions de parcel·la experimental al nivell de camp.

-L'ús de la infraestructura de dades satel·lital ha d'ajustar-se a la heterogeneïtat dels paisatges mediterranis per produir-los de manera sostenible en l'actual interacció digital de la societat amb el medi.



## Bibliografia

Adams JB, Gillespie AR. 2006. Remote sensing of landscapes with spectral images: A physical modeling approach. *Remote Sensing of Landscapes with Spectral Images: A Physical Modeling Approach* 1–362. DOI: 10.1017/CBO9780511617195

Alvarez Leon L. 2022. An emerging satellite ecosystem and the changing political economy of remote sensing. In: Goldstein J and Nost E (eds) *The Nature of Data: Infrastructures, Environments, Politics*. University of Nebraska Press, 37–60

Araus JL, Cabrera-Bosquet L, Serret MD, Bort J, Nieto-Taladriz MT. 2013. Comparative performance of  $\delta^{13}\text{C}$ ,  $\delta^{18}\text{O}$  and  $\delta^{15}\text{N}$  for phenotyping durum wheat adaptation to a dryland environment. *Functional Plant Biology* 40: 595. DOI: 10.1071/FP12254

Araus JL, Villegas D, Aparicio N, del Moral LFG, El Hani S, Rharrabti Y, Ferrio JP, Royo C. 2003. Environmental Factors Determining Carbon Isotope Discrimination and Yield in Durum Wheat under Mediterranean Conditions. *Crop Science* 43: 170. DOI: 10.2135/cropsci2003.1700

Avery TE, Berlin GL. 1992. *Fundamentals of Remote Sensing and Airphoto Interpretation*. Macmillan

Bakker K, Ritts M. 2018. Smart Earth: A meta-review and implications for environmental governance. *Global Environmental Change* 52: 201–211. DOI: 10.1016/j.gloenvcha.2018.07.011

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF. 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data* 5: 180214. DOI: 10.1038/sdata.2018.214

Bento VA, Ribeiro AFS, Russo A, Gouveia CM, Cardoso RM, Soares PMM. 2021. The impact of climate change in wheat and barley yields in the Iberian Peninsula. *Scientific Reports* 11: 1–12. DOI: 10.1038/s41598-021-95014-6

Blas A, Garrido A, Unver O, Willaarts B. 2019. A comparison of the Mediterranean diet and current food consumption patterns in Spain from a nutritional and water perspective. *Science of The Total Environment* 664: 1020–1029. DOI: 10.1016/j.scitotenv.2019.02.111

Blazquez CH, Edwards GJ. 1986. Spectral reflectance of healthy and diseased watermelon leaves. *Annals of Applied Biology* 108: 243–249. DOI: 10.1111/j.1744-7348.1986.tb07646.x

- Blondel J. 2006. The “design” of Mediterranean landscapes: A millennial story of humans and ecological systems during the historic period. *Human Ecology* 34: 713–729. DOI: 10.1007/s10745-006-9030-4
- Burke M, Lobell DB. 2017. Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proceedings of the National Academy of Sciences of the United States of America* 114: 2189–2194. DOI: 10.1073/pnas.1616919114
- Bzdok D, Altman N, Krzywinski M. 2018. Points of Significance: Statistics versus machine learning. *Nature Methods* 15: 233–234. DOI: 10.1038/nmeth.4642
- Cameron KC, Di HJ, Moir JL. 2013. Nitrogen losses from the soil/plant system: a review. *Annals of Applied Biology* 162: 145–173. DOI: 10.1111/aab.12014
- Cavalaris C, Megoudi S, Maxouri M, Anatolitis K, Sifakis M, Levizou E, Kyparissis A. 2021. Modeling of durum wheat yield based on sentinel-2 imagery. *Agronomy* 11. DOI: 10.3390/agronomy11081486
- Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM. 2015. Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances* 1. DOI: 10.1126/sciadv.1400253
- Ceccarelli S. 1994. Specific adaptation and breeding for marginal conditions. *Breeding Fodder Crops for Marginal Conditions*, 101–127. DOI: 10.1007/978-94-011-0966-6\_15
- Chairi F, Vergara-Diaz O, Vatter T, Aparicio N, Nieto-Taladriz MT, Kefauver SC, Bort J, Serret MD, Araus JL. 2018. Post-green revolution genetic advance in durum wheat: The case of Spain. *Field Crops Research* 228: 158–169. DOI: 10.1016/j.fcr.2018.09.003
- Chen P, Haboudane D, Tremblay N, Wang J, Vigneault P, Li B. 2010. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sensing of Environment* 114: 1987–1997. DOI: 10.1016/j.rse.2010.04.006
- Chlingaryan A, Sukkarieh S, Whelan B. 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture* 151: 61–69. DOI: 10.1016/j.compag.2018.05.012
- Cos J, Doblas-Reyes F, Jury M, Marcos R, Bretonnière P-A, Samsó M. 2022. The Mediterranean climate change hotspot in the CMIP5 and CMIP6 projections. *Earth System Dynamics* 13: 321–340. DOI: 10.5194/esd-13-321-2022

d'Andrimont R, Verhegghen A, Lemoine G, Kempeneers P, Meroni M, van der Velde M. 2021. From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. *Remote Sensing of Environment* 266. DOI: 10.1016/j.rse.2021.112708

Daniel C, Triboí E. 2002. Changes in wheat protein aggregation during grain development: Effects of temperatures and water stress. *European Journal of Agronomy* 16: 1–12. DOI: 10.1016/S1161-0301(01)00114-9

Defourny P, Bontemps S, Bellemans N, Cara C, Dedieu G, Guzzonato E, Hagolle O, Inglada J, Nicola L, Rabaut T, Savinaud M, Udrou C, Valero S, Bégué A, Dejoux JF, El Harti A, Ezzahar J, Kussul N, Labbassi K, Lebourgeois V, Miao Z, Newby T, Nyamugama A, Salh N, Shelestov A, Simonneaux V, Traore PS, Traore SS, Koetz B. 2019. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sensing of Environment* 221: 551–568. DOI: 10.1016/j.rse.2018.11.007

Disney M, Lewis P, Saich P. 2006. 3D modelling of forest canopy structure for remote sensing simulations in the optical and microwave domains. *Remote Sensing of Environment* 100: 114–132. DOI: 10.1016/j.rse.2005.10.003

Doxaran D, Froidefond J-M, Lavender S, Castaing P. 2002. Spectral signature of highly turbid waters. *Remote Sensing of Environment* 81: 149–161. DOI: 10.1016/S0034-4257(01)00341-8

Evenson RE, Gollin D. 2003. Assessing the Impact of the Green Revolution, 1960 to 2000. *Science* 300: 758–762. DOI: 10.1126/science.1078710

Foley JA, DeFries R, Asner GP, Barford C, Bonan G, Carpenter SR, Chapin FS, Coe MT, Daily GC, Gibbs HK, Helkowski JH, Holloway T, Howard EA, Kucharik CJ, Monfreda C, Patz JA, Prentice IC, Ramankutty N, Snyder PK. 2005. Global Consequences of Land Use. *Science* 309: 570–574. DOI: 10.1126/science.1111772

Foley JA, Ramankutty N, Brauman KA, Cassidy ES, Gerber JS, Johnston M, Mueller ND, O'Connell C, Ray DK, West PC, Balzer C, Bennett EM, Carpenter SR, Hill J, Monfreda C, Polasky S, Rockström J, Sheehan J, Siebert S, Tilman D, Zaks DPM. 2011. Solutions for a cultivated planet. *Nature* 478: 337–342. DOI: 10.1038/nature10452

- García-Martín M, Torralba M, Quintas-Soriano C, Kahl J, Plieninger T. 2021. Linking food systems and landscape sustainability in the Mediterranean region. *Landscape Ecology* 36: 2259–2275. DOI: 10.1007/s10980-020-01168-5
- Gitelson AA, Merzlyak MN. 1996. Signature analysis of leaf reflectance spectra: Algorithm development for remote sensing of chlorophyll. *Journal of Plant Physiology* 148: 494–500. DOI: 10.1016/S0176-1617(96)80284-7
- González-Fernández A, Segarra J, Sunny A, Couturier S. 2022. Forest cover loss in the Nevado de Toluca volcano protected area (Mexico) after the change to a less restrictive category in 2013. *Biodiversity and Conservation* 31. DOI: 10.1007/s10531-022-02368-y
- Graeber D. 2023. Pirate Enlightenment, or the real Libertalia.
- Jiménez-Olivencia Y, Ibáñez-Jiménez Á, Porcel-Rodríguez L, Zimmerer K. 2021. Land use change dynamics in Euro-mediterranean mountain regions: Driving forces and consequences for the landscape. *Land Use Policy* 109. DOI: 10.1016/j.landusepol.2021.105721
- Kokaly RF. 2001. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment* 75: 153–161. DOI: 10.1016/S0034-4257(00)00163-2
- Lang S, Hay GJ, Baraldi A, Tiede D, Blaschke T. 2019. GEOBIA achievements and spatial opportunities in the era of big earth observation data. *ISPRS International Journal of Geo-Information*. MDPI AG. DOI: 10.3390/ijgi8110474
- Lark TJ, Spawn SA, Bougie M, Gibbs HK. 2020. Cropland expansion in the United States produces marginal yields at high costs to wildlife. *Nature Communications* 11: 4295. DOI: 10.1038/s41467-020-18045-z
- Lasanta T, Vicente-Serrano SM. 2012. Complex land cover change processes in semiarid Mediterranean regions: An approach using Landsat images in northeast Spain. *Remote Sensing of Environment* 124: 1–14. DOI: 10.1016/j.rse.2012.04.023
- Le Gouis J, Oury FX, Charmet G. 2020. How changes in climate and agricultural practices influenced wheat production in Western Europe. *Journal of Cereal Science* 93: 102960. DOI: 10.1016/J.JCS.2020.102960
- Lechner AM, Foody GM, Boyd DS. 2020. Applications in Remote Sensing to Forest Ecology and Management. *One Earth* 2: 405–412. DOI: 10.1016/J.ONEEAR.2020.05.001



Liang S, Li X, Wang J. 2012. Advanced Remote Sensing. Academic Press

Lionello P, Malanotte-Rizzoli P, Boscolo R, Alpert P, Artale V, Li L, Luterbacher J, May W, Trigo R, Tsimplis M, Ulbrich U, Xoplaki E. 2006. The Mediterranean climate: An overview of the main characteristics and issues. , 1–26. DOI: 10.1016/S1571-9197(06)80003-0

Lloveras J, Lopez A, Ferran J, Espachs S, Solsona J. 2001. Bread-Making Wheat and Soil Nitrate as Affected by Nitrogen Fertilization in Irrigated Mediterranean Conditions. *Agronomy Journal* 93: 1183–1190

Matsushita B, Yang W, Chen J, Onda Y, Qiu G. 2007. Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to topographic effects: A case study in high-density cypress forest. *Sensors* 7: 2636–2651. DOI: 10.3390/s7112636

Mukherjee A, Panayotov G, Shon J. 2021. Eye in the sky: Private satellites and government macro data. *Journal of Financial Economics* 141: 234–254. DOI: 10.1016/j.jfineco.2021.03.002

Nagaraj A. 2022. The Private Impact of Public Data: Landsat Satellite Maps Increased Gold Discoveries and Encouraged Entry. *Management Science* 68: 564–582. DOI: 10.1287/mnsc.2020.3878

Nost E, Goldstein JE. 2022. A political ecology of data. *Environment and Planning E: Nature and Space* 5: 3–17. DOI: 10.1177/25148486211043503

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D’Amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR. 2001. Terrestrial ecoregions of the world: A new map of life on Earth. *BioScience* 51: 933–938. DOI: 10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Osorno-Covarrubias F, Couturier S, Piceno Hernández M. 2018. Measuring from Space the Efficiency of Local Forest Management: The Successful Case of the Indigenous Community of Cherán, Mexico. *IOP Conference Series: Earth and Environmental Science* 151: 012010. DOI: 10.1088/1755-1315/151/1/012010

Palta JA, Fillery IR. 1995. N application enhances remobilisation and reduces losses of pre-anthesis N in wheat grown on an Duplex soil. *Australian Journal of Agricultural Research* 46: 519–531

Panikkar R. 2021. Ecosofia. La saviesa de la Terra. Fragmenta

Papakosta DK, Gagianas AA. 1991. Nitrogen and Dry Matter Accumulation, Remobilization, and Losses for Mediterranean Wheat during Grain Filling. *Agronomy Journal* 83: 864–870. DOI: 10.2134/agronj1991.00021962008300050018x

Peña RJ. 2002. Wheat for bread and other foods. Bread wheat improvement and production. Rome

Ponti L, Gutierrez AP, Altieri MA. 2016. Preserving the Mediterranean Diet Through Holistic Strategies for the Conservation of Traditional Farming Systems. *Environmental History (Netherlands)* 5: 453–469. DOI: 10.1007/978-3-319-26315-1\_24

Rautiainen M, Stenberg P. 2005. Application of photon recollision probability in coniferous canopy reflectance simulations. *Remote Sensing of Environment* 96: 98–107. DOI: 10.1016/j.rse.2005.02.009

Rockström J, Steffen W, Noone K, Persson Å, Chapin FS, Lambin E, Lenton TM, Scheffer M, Folke C, Schellnhuber HJ, Nykvist B, de Wit CA, Hughes T, van der Leeuw S, Rodhe H, Sörlin S, Snyder PK, Costanza R, Svedin U, Falkenmark M, Karlberg L, Corell RW, Fabry VJ, Hansen J, Walker B, Liverman D, Richardson K, Crutzen P, Foley J. 2009. Planetary boundaries: Exploring the safe operating space for humanity. *Ecology and Society* 14. DOI: 10.5751/ES-03180-140232

Rosegrant MW, Ringler C, Zhu T. 2009. Water for Agriculture: Maintaining Food Security under Growing Scarcity. *Annual Review of Environment and Resources* 34: 205–222. DOI: 10.1146/annurev.environ.030308.090351

Rounsevell MDA, Annetts JE, Audsley E, Mayr T, Reginster I. 2003. Modelling the spatial distribution of agricultural land use at the regional scale. *Agriculture, Ecosystems and Environment* 95: 465–479. DOI: 10.1016/S0167-8809(02)00217-7

Royo C, Soriano JM, Alvaro F. 2017. Wheat: A Crop in the Bottom of the Mediterranean Diet Pyramid. *Mediterranean Identities - Environment, Society, Culture*. DOI: 10.5772/intechopen.69184

Santos M. 2000. La naturaleza del espacio. Ariel geografía

Segarra J. 2024. Satellite Imagery in Precision Agriculture. *Digital Agriculture*. Springer International Publishing: Cham, 325–340. DOI: 10.1007/978-3-031-43548-5\_10

Segarra J, Buchailot ML, Araus JL, Kefauver SC. 2020. Remote Sensing for Precision Agriculture : Sentinel-2 Improved Features and Applications. *Agronomy* 1–18

Segarra J, Fernàndez-Martínez J, Araus JL. 2023. Managing abandoned Mediterranean mountain landscapes: The effects of donkey grazing on biomass control and floral diversity in pastures. *Catena* 233. DOI: 10.1016/j.catena.2023.107503

Serrano L, Peñuelas J, Ustin SL. 2002. Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data. *Remote Sensing of Environment* 81: 355–364. DOI: 10.1016/s0034-4257(02)00011-1

Serret M., Ortiz-Monasterio I, Pardo A, Araus J. 2008. The effects of urea fertilisation and genotype on yield, nitrogen use efficiency,  $\delta^{15}\text{N}$  and  $\delta^{13}\text{C}$  in wheat. *Annals of Applied Biology* 080617165316730-??? DOI: 10.1111/j.1744-7348.2008.00259.x

Spitters CJ. 1989. Crop growth models: their usefulness and limitations. VI Symposium on the Timing of Field Production of Vegetables 267. DOI: <https://doi.org/10.17660/ActaHortic.1990.267.42>

Subirats E. 1981. La ilustración insuficiente. Taurus

Sudmanns M, Tiede D, Lang S, Bergstedt H, Trost G, Augustin H, Baraldi A, Blaschke T. 2020. Big Earth data: disruptive changes in Earth observation data management and analysis? *International Journal of Digital Earth* 13: 832–850. DOI: 10.1080/17538947.2019.1585976

Tubiello FN, Salvatore M, Ferrara AF, House J, Federici S, Rossi S, Biancalani R, Condor Golec RD, Jacobs H, Flammini A, Prosperi P, Cardenas-Galindo P, Schmidhuber J, Sanz Sanchez MJ, Srivastava N, Smith P. 2015. The Contribution of Agriculture, Forestry and other Land Use activities to Global Warming, 1990-2012. *Global Change Biology* 21: 2655–2660. DOI: 10.1111/gcb.12865

Tucker CJ. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8: 127–150. DOI: 10.1016/0034-4257(79)90013-0

Tuel A, Eltahir EAB. 2020. Why Is the Mediterranean a Climate Change Hot Spot? *Journal of Climate* 33: 5829–5843. DOI: 10.1175/JCLI-D-19-0910.1

Union of Concerned Scientists. 2023. UCS Satellite Database.

von Keyserlingk J, de Hoop M, Mayor AG, Dekker SC, Rietkerk M, Foerster S. 2021. Resilience of vegetation to drought: Studying the effect of grazing in a Mediterranean rangeland using satellite time series. *Remote Sensing of Environment* 255: 112270. DOI: 10.1016/j.rse.2020.112270

Weiss M, Jacob F, Duveiller G. 2020. Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment* 236: 111402. DOI: 10.1016/J.RSE.2019.111402

Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D, DeClerck F, Wood A, Jonell M, Clark M, Gordon LJ, Fanzo J, Hawkes C, Zurayk R, Rivera JA, De Vries W, Majele Sibanda L, Afshin A, Chaudhary A, Herrero M, Agustina R, Branca F, Lartey A, Fan S, Crona B, Fox E, Bignet V, Troell M, Lindahl T, Singh S, Cornell SE, Srinath Reddy K, Narain S, Nishtar S, Murray CJL. 2019. Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet* 393: 447–492. DOI: 10.1016/S0140-6736(18)31788-4

Winkler K, Fuchs R, Rounsevell M, Herold M. 2021. Global land use changes are four times greater than previously estimated. *Nature Communications* 12: 1–10. DOI: 10.1038/s41467-021-22702-2

Zampieri M, Toreti A, Ceglar A, Naumann G, Turco M, Tebaldi C. 2020. Climate resilience of the top ten wheat producers in the Mediterranean and the Middle East. *Regional Environmental Change* 20. DOI: 10.1007/s10113-020-01622-9

Zomeni M, Tzanopoulos J, Pantis JD. 2008. Historical analysis of landscape change using remote sensing techniques: An explanatory tool for agricultural transformation in Greek rural areas. *Landscape and Urban Planning* 86: 38–46. DOI: 10.1016/j.landurbplan.2007.12.006



## **Annex 1.** L'aplicació de la teledetecció en l'agricultura de precisió: les característiques millorades de Sentinel-2

### **Annex 1.** Remote sensing for precision agriculture: Sentinel-2 improved features and applications

Joel Segarra <sup>1,2</sup>, Maria L. Buchailot <sup>1,2</sup>, Jose Luis Araus <sup>1,2</sup> and Shawn C. Kefauver <sup>1,2,\*</sup>

<sup>1</sup>Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain

<sup>2</sup>AGROTECNIO (Center for Research in Agrotechnology), 25198 Lleida, Spain

\*Correspondence: [sckefauver@ub.edu](mailto:sckefauver@ub.edu)

Publicat a la revista Agronomy

Review

# Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications

Joel Segarra <sup>1,2</sup>, Maria Luisa Buchaillot <sup>1,2</sup> , Jose Luis Araus <sup>1,2</sup>  and Shawn C. Kefauver <sup>1,2,\*</sup>

<sup>1</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain; joel.segarra@ub.edu (J.S.); luisa.buchaillot@gmail.com (M.L.B.); jaraus@ub.edu (J.L.A.)

<sup>2</sup> Agrotecnio Centre for Research in Agrotechnology, Av. Rovira Roure 191, 25198 Lleida, Spain

\* Correspondence: sckefauver@ub.edu

Received: 6 April 2020; Accepted: 29 April 2020; Published: 1 May 2020



**Abstract:** The use of satellites to monitor crops and support their management is gathering increasing attention. The improved temporal, spatial, and spectral resolution of the European Space Agency (ESA) launched Sentinel-2 A + B twin platform is paving the way to their popularization in precision agriculture. Besides the Sentinel-2 A + B constellation technical features the open-access nature of the information they generate, and the available support software are a significant improvement for agricultural monitoring. This paper was motivated by the challenges faced by researchers and agrarian institutions entering this field; it aims to frame remote sensing principles and Sentinel-2 applications in agriculture. Thus, we reviewed the features and uses of Sentinel-2 in precision agriculture, including abiotic and biotic stress detection, and agricultural management. We also compared the panoply of satellites currently in use for land remote sensing that are relevant for agriculture to the Sentinel-2 A + B constellation features. Contrasted with previous satellite image systems, the Sentinel-2 A + B twin platform has dramatically increased the capabilities for agricultural monitoring and crop management worldwide. Regarding crop stress monitoring, Sentinel-2 capacities for abiotic and biotic stresses detection represent a great step forward in many ways though not without its limitations; therefore, combinations of field data and different remote sensing techniques may still be needed. We conclude that Sentinel-2 has a wide range of useful applications in agriculture, yet still with room for further improvements. Current and future ways that Sentinel-2 can be utilized are also discussed.

**Keywords:** Sentinel-2; remote sensing; precision agriculture; crop monitoring

## 1. Introduction

In current and future climate scenarios the resilience and productivity of agricultural systems will be increasingly jeopardized [1]. Furthermore, population growth trends, expected to reach 8.7 billion by 2030 and 9.7 billion by 2050 [2], will further strain food production systems worldwide, which so far have not been able to keep pace [3]. Within this changing context, crops face a threefold obstacle: management-derived challenges as well as increased pressure from abiotic and biotic stressors. Thus, the application of all available advanced technologies towards managing crop variability and maintaining or improving yields and reducing negative impacts on environmental quality, namely advancements in precision agriculture [4], is central to approaching these issues.

The possible technological solutions for these issues are multifold within precision agriculture's paradigm. On the one hand, plant breeding and the use of phenotyping technological advances to improve breeding programs [5] allow for obtaining genotypes adapted to various stressors. However, improved management scales are also a central topic to approach. The fundamental functioning



of agroecosystems in food production and their role in environmental conservation calls for the development of operational global scale methodologies. In this sense, within the scope of precision agriculture, remote sensing technologies are able to provide crucial contributions to agriculture monitoring [6] and improve our understanding of the impacts of environmental variations [7]. Many of the limitations in the application of remote sensing techniques for precision agriculture of previous years [8] have been overcome with the launches of Sentinel-2 A + B. The Sentinel-2 constellation, with improved spatial, spectral, and temporal resolution was designed specifically to meet the needs of the agricultural community, both farmers and academic researchers with a focus on international agricultural development.

Management-derived challenges as well as abiotic and biotic stressors may all be confronted and their negative impacts greatly ameliorated by employing remote sensing technologies to better guide adaptive agricultural management decisions, such as variable rate applications and treatments, which can be assessed and confronted at different levels of precision using different sources of remotely sensed data. Remote sensing data has long been made freely available to help to deploy precision agriculture in crop management to optimize the use of scarce resources but requires the necessary know-how for correct implementation. The combination of imaging and spectroscopy has allowed, for the last five decades since the launch of the NASA (National Aeronautics and Space Administration) Landsat-1 or Earth Resources Technology Satellite (ERTS) and its Multispectral Scanner (MSS) in 1972, the application of satellite-based remote sensing technologies. The use of this technology has played an important role, yet with some limitations, in agriculture monitoring all around the world and will increasingly be needed to assess and confront issues related to global climate change. Remote sensing provides coverage of large areas with high precision and can be a very efficient tool for improved management across scales. In this sense, remote sensing using multispectral images is a proxy for extensive manual crop monitoring operations in the field [9] and provides potential savings in precious time and resources. Precision agriculture has grown and established itself as a central tool in agricultural management to forecast yield, monitor plant stress, and optimize fertilization, irrigation, and soil tilling activities for a more sustainable management of farming practices and improvements in yield for farmers.

The launches of Sentinel-2 A + B twin platforms by the European Space Agency (ESA) has resulted in an improvement of precision agriculture possibilities. The ESA openly and freely provides Sentinel-2 with an improved number of multispectral bands, a more frequent revisit time, and a higher spatial resolution that should be of great interest to the agricultural community and to international agricultural development researchers worldwide.

This review article focuses on the field of precision agriculture and monitoring, with a particular emphasis on the recent advances provided by the ESA Sentinel-2 constellation, as part of the European Union (EU) Copernicus Program. Our goal is to frame remote sensing principles and Sentinel-2 A + B constellation features in the context of the state-of-the-art in precision agriculture. In order to better understand the improved features of Sentinel-2, its characteristics were compared to other satellites with potential use for agricultural monitoring. We also present an updated review of scientific literature on Sentinel-2 applications and advancements in precision agriculture. Furthermore, we explore its applications in biotic and abiotic stress monitoring, as well as at different management scales for agriculture.

## 2. Remote Sensing for Agriculture

### 2.1. Plant Spectral Reflectance Properties

Plants interact with sunlight—the full spectrum of sun-emitted electromagnetic radiation—differently depending on the wavelength observed. In this sense, incident solar radiation can follow three pathways: it can be transmitted, reflected, or absorbed. The electromagnetic radiation that is reflected by plants contains information about their biophysical composition and physiological status

and can be measured using satellite sensors, such as those placed in the ESA Sentinel-2. Focusing on the visible domain, the chlorophyll pigments present in green leaves strongly absorb in the visible (VIS) region of the spectrum (400–700 nm), especially in the blue and red wavelengths, where energy is captured for photosynthesis. Meanwhile in the near-infrared (NIR), from approximately 700 to 1300 nm, leaves exhibit high reflectance values and transmission, mostly related to leaf structural properties and biomass, and absorb less radiation in this spectral region [10,11]. Furthermore, plant canopy structure and leaf area are also fundamental traits connected to canopy reflectance patterns [12,13] and are key parameters for monitoring growth. Regarding the shortwave infrared region (SWIR), from approximately 1300 to 2500 nm, the absorption of radiation is largely dominated by water, followed by other biochemical components present in leaves. Concerning agriculture, in situations where crops interact with any given aspect of their environment (seasonal climatic variations, meteorological extreme events, pests, soil properties, etc.) or as crops grow and pass through different phenological stages, the interactions between plants and light reflectance translate into changes in plant signal patterns that can be interpreted using satellite data. The use of remote sensing instruments to monitor electromagnetic radiation reflectance changes in crops is well demonstrated in the scientific literature [13–24]. For example, chlorophyll content plays a direct role in photosynthesis and it is responsive to a range of stresses, and hence is often used as a general plant health indicator [25] and interpreted along with the signals related to biomass and water content. These are examples of general potential uses of plant spectral reflectance properties for vegetation monitoring. However, often it is the temporal and spatial properties of satellites that determine if their use is appropriate for agriculture. This is a consequence of the characteristics of crops, which grow more quickly and frequently in smaller and more closely managed parcels of land compared to natural vegetation.

## *2.2. Satellites and Remote Sensing Scales of Observation*

Over the last 50 years of digital evolution, numerous satellites with multispectral sensors have been launched to orbit around the Earth for the purpose of environmental and agricultural remote sensing. Currently, one of the most relevant Earth observation programs focused on detection and assessment of vegetation is Copernicus. The Copernicus Program, led by the ESA, is a program aimed to increase the satellite-based Earth observation capacities and within which Sentinel-2 twin platforms were launched. Besides the multispectral Sentinel-2 system, various satellites have been launched or are planned to be launched as part of this still on-going program. Regarding potential agricultural applications within the Copernicus program, it is worth mentioning Sentinel-1 (<https://sentinel.esa.int/web/sentinel/missions/sentinel-1>) that carries a synthetic-aperture radar on board as well as Sentinel-3 that provides thermal data at 1 km of spatial resolution (<https://sentinel.esa.int/web/sentinel/missions/sentinel-3>).

Framing the concepts regarding satellites features is especially relevant in order to grasp Sentinel-2 improvements compared to other orbiting satellites. In this sense, it is central to understand first satellite resolution and observation scales as they determine satellite systems' agricultural monitoring capacities.

Regarding resolution, four parameters are central to approach a better understanding of the aims and scope of these instruments. First, the spectral resolution, namely the number, width, and location on the electromagnetic spectrum of spectral bands in the sensor system. Second, the spatial resolution, which is the measure of the smallest object that can be resolved by the sensor, e.g., pixel size. Third, one must consider the temporal resolution which is a measure of the frequency with which a sensor revisits the same area of the Earth's surface; and finally, the radiometric resolution, which is determined by the number of bits into which the recorded radiation is divided [26] and determines the spectral precision of the measurements.

Following Homolová et al. [27], on the matching of ecological, agricultural, and remote sensing scales, we currently find that on the basis of individual plants, plant communities, and agroecosystems, different remote sensing technologies may provide coverage across canopy and landscape scales. Furthermore, it provides coverage across a range of plant physiological processes from evapotranspiration and photosynthesis to phenology dynamics, productivity, and even quality

traits such as those of a crop’s harvestable component nutritional value. The spatial, temporal, and spectral resolutions of any sensor-satellite system are central to determining the scale at which an instrument can monitor crops, and therefore the most adequate approach towards guiding agricultural monitoring and assessments and subsequently management actions.

### 2.3. Sentinel-2 Spatial and Temporal Resolutions Compared to Other Satellites

Sentinel-2 has had two launches, Sentinel-2A that was launched on 23 June 2015 and Sentinel-2B on 7 March 2017, forming, up from this date, the Sentinel-2 A + B constellation (<https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/sentinel-2>). Both have a 7.25-year lifetime, including a 3 months commissioning phase with adequate propellant for up to 12 years. Both Sentinel-2A and Sentinel-2B satellites have on board the exact same MultiSpectral Instrument (MSI) with bands from the visible to the shortwave infrared: four (4) bands at 10 m, the classical broadband visible blue (490 nm), green (560 nm), red (665 nm), and near infrared (842 nm); six bands at 20 m, four narrow bands in the vegetation red edge spectral domain (705, 740, 775, and 865 nm), and two longer SWIR bands (1610 and 2190 nm); and three (3) bands at 60 m dedicated to atmospheric correction (443 nm for aerosols and 940 nm for water vapor) and to cirrus detection (1380 nm) (Table 1). After the second launch the satellite provides higher temporal resolution with a 5-day global revisit frequency, and up to 2-day revisit in top northern and southern parts of the globe.

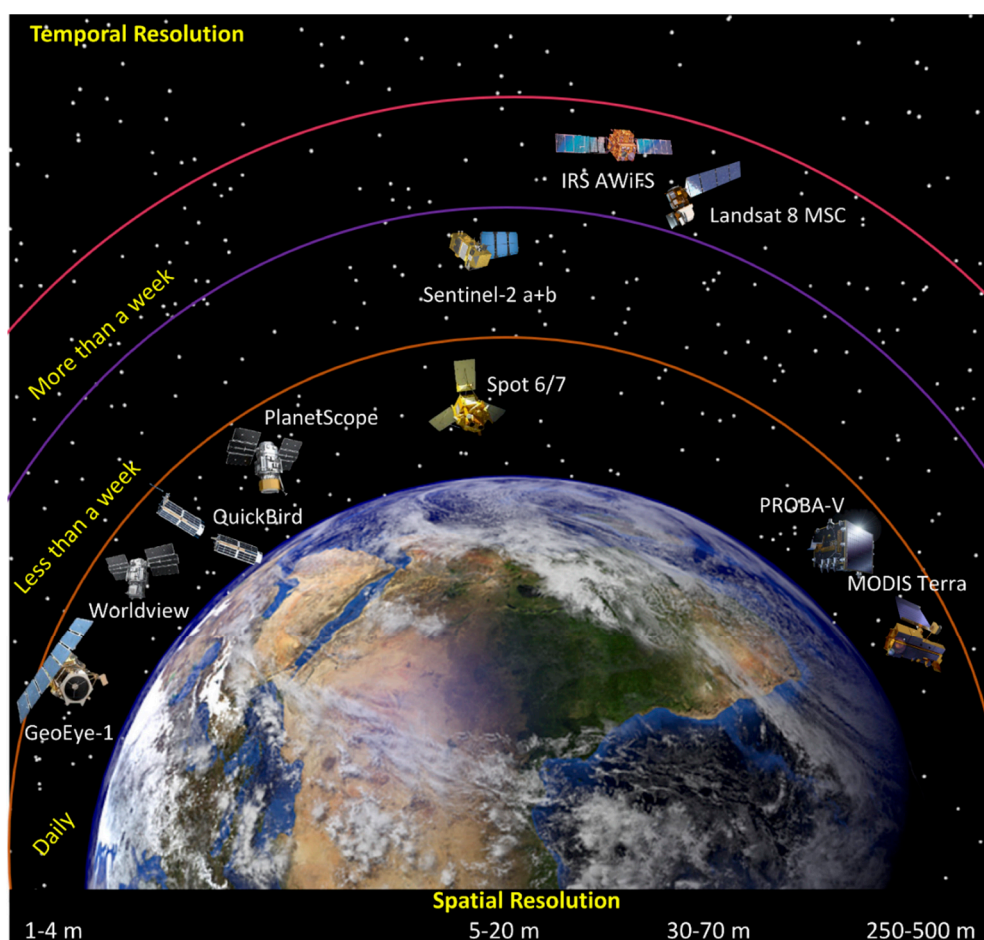
Satellite spatial and temporal resolutions depend on their type of orbit and their orbit altitude. Most of the satellites are two types: polar orbiting which are sun synchronous and have orbits that cross near the poles in low earth orbit (around 700 km); or are geostationary geosynchronous and orbit above the equator, located much farther distance from the earth (35,000 km). In this case Sentinel-2 is sun synchronous and has a systematic coverage of all land surfaces (including croplands) globally from  $-56^{\circ}$  (Southern America) to  $+83^{\circ}$  (Northern Greenland) latitude, with a swath width of 290 km [28] and an earth orbit of 786 km.

**Table 1.** Bands and resolutions of Sentinel-2 MultiSpectral Instrument (MSI) [28,29].

MSI Band	Spatial Resolution (m)	Central Wavelength (nm)	Bandwidth (nm)
B1: Coastal Aerosol	60	443	20
B2: Blue	10	490	65
B3: Green	10	560	35
B4: Red	10	665	30
B5: Red-Edge	20	705	15
B6: Red-Edge	20	740	15
B7: Red-Edge	20	783	20
B8: NIR	10	842	115
B8A: Vegetation RE	20	865	20
B9: Water Vapor	60	945	20
B10: SWIR Cirrus	60	1375	30
B11: SWIR	20	1610	90
B12: SWIR	20	2190	180

Focusing on spatial and temporal resolution, in Figure 1 various satellites are shown in comparison to the most recently launch satellite of interest to precision agriculture—Sentinel-2. For comparison, the MODIS (Moderate Resolution Imaging Spectroradiometer) and Proba-V, with spatial resolutions ranging from 150 to 500 m and both with a daily revisit time can monitor agricultural systems at the agroecosystem level. Meanwhile, the IRS AWiFS and Landsat 8 satellites have a spatial resolution of 30–100 m and a revisit time of 24 and 16 days, respectively. In this case, both agroecosystems and plant communities at local to regional scales, namely the crop field level, can be monitored. On the other hand, both the SPOT (Satellite Pour l’Observation de la Terre) 6/7 have similarities with Sentinel-2 A + B as both are constellations of two satellites and therefore the combined revisit time is reduced, which translates in providing global daily imagery in the case of SPOT and every 5 days in the case

of Sentinel-2. Additionally, the constellation combination also helps maintain relatively detailed spatial resolution, with SPOT spatial resolution at 6 m and the Sentinel-2 maximum resolution is 10 m. Specifically in the case of Sentinel-2 resolution ranges between 10 and 60 m, but the agricultural dedicated bands are either 10 or 20 m, depending on the band width, while the other bands at 60 m are for atmospheric aerosols and water vapor observations mainly for improved image calibration. Due to their shorter revisit time and more detailed spatial resolution (compared, for example, to LandSat at every 16 days and mostly 30 m pixel size), more precision in sub-field monitoring can be performed using these paired satellite constellations, hence covering from agroecosystem to field scales more precisely and thus gaining relevance for use in agricultural contexts and even more specifically smallholder farming systems in developing countries. Finally, GeoEye-1, Worldview-3, Planet, and Quickbird are each constellations of numerous commercial satellites allowing higher temporal resolution (as high as a 24 h revisit time) and have a more detailed spatial resolution with pixels smaller than 3 m, but in comparison to Sentinel-2 they do not provide free and open-access global coverage and may, in some cases, only capture data at controlled moments when contracted, which limits their potential use.

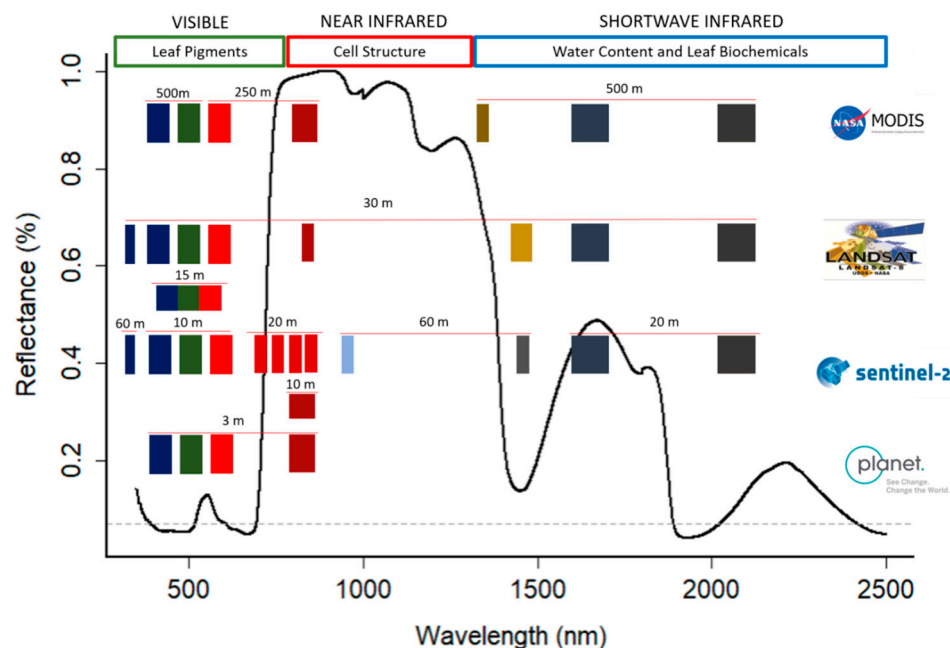


**Figure 1.** Current constellation of available satellites for land remote sensing that are highly relevant to crop monitoring and management, presented with their temporal and spatial resolution.

#### 2.4. Spectral Resolutions and Useful Vegetation Indices for Sentinel-2 Features

Besides temporal and spatial resolution, spectral resolution is a central feature of satellites as it determines the plant physiological properties that may be accurately quantified. In this sense, the number of spectral bands in several sensor systems is shown together with the canopy reflectance spectrum in Figure 2, taking the examples of landscape, canopy, and plant scale sensors placed in

satellites (MODIS, Landsat-8, Sentinel-2, and Planet, respectively). Comparing these satellites with Sentinel-2, we observe that all include sensors with visible, near infrared, and shortwave infrared bands; differently, Planet has only the three available bands in the visible (RGB, for red, green, and blue reflected light) and one in the near infrared (NIR). A different feature from Sentinel-2 is that it also captures four narrowband spectral regions of different wavelengths present in the Red Edge (RE) spectral region that are provided at 20 m resolution (Figure 2). The reflectance spectrum between 690 and 740 nm determines the point of maximum slope in the spectral signature of the vegetation [19], such that the RE marks the boundary between the processes of chlorophyll absorption in red wavelengths and within-leaf scattering in near-infrared wavelengths [20,21] (Figure 2), which provides unique capacities for separating crop photosynthetic activity from biomass, as indicated by some of the spectral indices that have been developed specifically for deployment by Sentinel-2 or other satellites that provide RE spectral band coverage. Moreover, the Sentinel-2 10 m spatial resolution that is achieved in the other broader spectral bands in the visible and near-infrared can provide improved details and allows for obtaining a large variety of spectral information to define crop topologies. In this sense, the big data volume of Sentinel-2 images provides novel opportunities though also challenges towards processing due to the increased complexity of interpreting additional spectral bands and the increased overall file size.



**Figure 2.** Crop reflectance spectral features, and representation of MODIS (Moderate Resolution Imaging Spectroradiometer), Landsat-8, Sentinel-2, and Planet multispectral bands and spatial resolution.

While many crop parameters may be estimated or indirectly inferred from their spectral light reflectance properties, some parameters, however, cannot be directly sensed but rather inferred through a combination of measurements of other plant properties. Chlorophyll spectral features are optically discernible within the visible and red-edge parts of the spectrum in a straightforward manner as these are classic plant pigments whose most basic function is the absorption of light for photosynthesis. On the other hand, plant nitrogen (N) content, for example, may not optically be discernible because N absorption features (1450–1940 nm) are obscured by water [22]. However, due to the close relationship between N concentration and chlorophyll content, chlorophyll-related indicators can be used as proxies of crop nitrogen concentration or status [23]. In this sense, the sensitivity of remote sensing data towards specific plant biophysical properties of interest can be estimated using vegetation indexes (VIs, Table 2) [24,30], which are usually fairly simple spectral wavelength (or spectral



bands to indicate both the width and the central wavelength of a spectral region as captured by satellite multispectral sensors) measurement combinations. More in depth reviews and extensive databases (e.g., <https://www.indexdatabase.de>) on the spectral regions most useful for agricultural monitoring are suggested for a more exhaustive coverage of spectral indexes in general [31,32]. Regarding the chosen VIs relevant to Sentinel-2 imagery presented in Table 2, for the visible spectrum the Normalized Green–Red Difference Index (NGRDI) compares the differences between the green and red visible bands for vegetation, together with the well-known Normalized Difference Vegetation Index (NDVI) have shown to be able to remotely sense total crop green biomass [33], which is often indicative of general crop vigor and may be closely correlated with crop nitrogen (N) content and may provide good predictions of yield at specific growth stages. Another index that uses the visible bands of the spectrum is the Triangular Greenness Index (TGI), which has been shown to estimate correctly chlorophyll concentration at leaf and canopy level [34].

In some cases, a combination of two VIs may allow for more specific measurements, given that the sensor has enough spectral bands for both calculations. Many of these more advanced calculations were taken into account in the specific design of Sentinel-2 as part of its development specifically for precision agriculture and environmental/vegetation monitoring in general. For example, the ratio of the Transformed Chlorophyll in Reflectance Index and the Optimized Soil-Adjusted Vegetation Index (TCARI/OSAVI) has been shown to estimate chlorophyll leaf content [35], basically by providing a canopy chlorophyll measurement (TCARI) and then correcting for biomass (OSAVI). In some instances, the spectral indexes were originally developed for a prior satellite with similar spectral but coarser spatial resolution that did not make them amenable to smaller-scale agricultural applications, such as in the case of the MERIS (Medium-spectral Resolution Imaging Spectrometer) Terrestrial Chlorophyll Index (MTCI) [36]. Based on the red-edge bands, many of which are newly featured at higher spatial resolutions with the Sentinel-2 satellite, MTCI has showed that Sentinel-2 derived indexes using red-edge bands perform very precisely in estimating chlorophyll and N content [36]. It is also the case of the chlorophyll index red edge (CI red edge) [37], effective for estimating leaf area index (LAI). The Chlorophyll Vegetation Index (CVI), a broad-band VI sensitive to leaf chlorophyll content at canopy level, has also shown good results for estimating chlorophyll content [38]. Another index of interest, in this case based on NIR and RE bands is the Inverted Red-Edge Chlorophyll Index (IRECI), which has shown good abilities to estimate chlorophyll content at the canopy level [39].

**Table 2.** An example of some spectral vegetation index (VIs) calculations of interest for use with European Space Agency (ESA) Sentinel-2 satellite data, where B is the spectral reflectance value of the band number of the Sentinel-2 band as detailed in Table 1.

VIs	Sentinel-2 Bands Used	Original Author
NGRDI	$(B3 - B4) / (B3 + B4)$	Hunt (2005)
TGI	$-0.5 \times [190 \times (B4 - B3) - 120 \times (B4 - B2)]$	Hunt (2012)
NDVI	$(B8 - B4) / (B8 + B4)$	Tucker (1979)
TCARI/OSAVI	$3 \times [(B5 - B4) - 0.2 \times (B5 - B3) \times (B5 / B4)] / [(1 + 0.16) \times (B7 - B4) \times (B7 + B4 + 0.16)]$	Haboudane (2002)
MTCI	$(B6 - B5) / (B5 - B4)$	Dash and Curran (2004)
CVI	$(B8 / B3) / (B3 / B4)$	Vincini (2008)
CI red-edge	$(B8 / B5) - 1$	Gitelson (2003)
IRECI	$(B7 - B4) / (B5 / B6)$	Frampton (2013)

## 2.5. Sentinel-2 Data Access, Secondary Products, and Support Tools

Regarding some of the highest spatial and temporal resolution satellites, such as Planet, the access to imagery is private or at least restricted, and therefore inaccessible for many institutions with limited funding, especially in developing countries. Besides that, NASA satellites, namely MODIS and LANDSAT-8 for instance, are publicly accessible and completely free, as well as ESA's entire Sentinel program, including Sentinel-2, along with the other sun-synchronous orbiting satellites that form part

of the greater ESA Copernicus Program. This makes the data more readily accessible and more openly free for a more globally coordinated remote sensing of agriculture in this case.

The accessibility to this data is central. In this sense, there are several ways to freely and openly download Sentinel-2 imagery; one of them is the direct download of the imagery from ESA's website, through Copernicus (<https://scihub.copernicus.eu/dhus/#/home>). Furthermore, third party tools for downloading the imagery are available. For instance, there is the United States Geological Survey (USGS) (<https://earthexplorer.usgs.gov/>) which allows comprehensive searching and downloading of full-resolution Sentinel-2 images. On the open-source software QGIS, there are various plug-ins that take advantage of the ESA's Application Programming Interface (APIs) at the Copernicus Open Access Hub (<https://scihub.copernicus.eu/apihub/>), such as the free and open source Semi-Automatic Classification Plugin (SCP), developed by L. Congedo [40], which provides tools to simply download Sentinel-2 imagery or even download and immediately calibrate and apply image analyses to the data. Furthermore, the API access portals allow for download only specific bands of select images, thus reducing the storage requirements of personal computers. Moreover, Google Earth Engine has daily updated copies of all the available Sentinel-2 data and provides both access to this data repository along with high processing capacity using their image processing servers. Many other similar tools and services exist on other software applications and web portals and are being developed continuously.

Another relevant point is data itself, its plan for continuity, and its access to both raw data and pre-calibrated and pre-processed satellite images that provide value added benefits to end-users. Sentinel-2 products consist of Level-1B, which provides radiometrically corrected granulates in Top-Of-Atmosphere radiance covering  $(25 \times 23) \text{ km}^2$ . Sentinel-2 also offers Level-1C imagery which consists of tiles of  $100 \times 100 \text{ km}^2$  that provide geo-coded Top-Of-Atmosphere reflectance with a sub-pixel multi-spectral and multi-date registration. It additionally includes cloud masks, supplying an indication about the presence of cirrus clouds. Moreover, it provides Level-2A, a tile of  $100 \times 100 \text{ km}^2$  which provides Bottom-Of-Atmosphere reflectance in cartographic geometry [41]. Since December 2018 Level-2A products are generated globally and are ready for users to download. Level-2A can be also generated using Level-1C with Sentinel-2 Toolbox.

ESA frequently provides specific access tools, algorithms, and software in support of the use and processing of their satellites, such as the Sentinel-2 Toolbox within the Sentinel Application Platform (SNAP, <https://sentinel.esa.int/web/sentinel/toolboxes/sentinel-2>). On SNAP Sen2Cor is the atmospheric correction processor to generate Level-2A data; it also includes algorithms for cloud and cloud shadow detection, cirrus detection, and slope effect correction. Beyond SNAP, Sentinel-2 imagery can also be atmospherically corrected using MAJA (developed by CESBIO/CNES) or i-COR (developed by VITO), among other options. In addition, the further pre-processing chain of Level-2A products can be done on SNAP: resampling and subsetting followed by calculations using the bands. Besides vegetation indices, more deterministic biophysical parameters, such as LAI or FAPAR, can be calculated on SNAP with Sentinel-2 data following the algorithms developed by Weiss and colleagues [42]. It is relevant to point that as an experimental Sentinel-2 approach done with simulated Sentinel-2 imagery in 2015 [43], even before its launch, it already showed potential for retrieving LAI and with actual Sentinel-2 data it is currently fully operational.

### 3. Sentinel-2 for Precision Agriculture

The ESA Sentinel-2 satellite's potential applications in agriculture are multi-fold, as described above with regard to its specific spatial, temporal, and spectral capacities. Regarding the definition of the International Society of Precision Agriculture, precision agriculture is a management strategy that uses temporal and spatial information to support management decisions according to estimated variability for improved resource use efficiency and sustainability of agricultural production. This section gathers examples of Sentinel-2 capacity to sense variations at several levels of agricultural production and the subsequent management decisions that can be taken to address any disruption on crop development. The examples discussed are summarized in Table 3.



One very notable use of Sentinel-2 features at field monitoring is the ESA Sen2-Agri system which generates modules of time series along the growing season. These different products consist of a monthly cloud-free composite of surface reflectance at 10–20 m resolution; monthly dynamic cropland masks, delivered from the agricultural mid-season onwards; a cultivated crop type map at 10 m resolution for main crop groups; and time series of vegetation status indicators, NDVI and LAI [44]. This field monitoring tool has been successfully used not only in developed countries with homogenous and well delineated fields, but as well in challenging (in terms of farm fields boundaries) countries such as Mali, successfully differentiating crop types (80% of overall accuracy) [45]. In Europe S2-Agri is going to be used among other initiatives within the Sen4-CAP (<http://esa-sen4cap.org/>) to advance towards an integrated administration and control system of the Common Agriculture Policy (CAP), which will be reformulated in 2020. CAP subsidies play a central role in European Union policies as they account for around 30% of EU total budget [46]. Hence, the European Union is taking advantage of Sentinel-2 features to monitor croplands aiming to optimize resources by determining actual crop type and cropland accurate area. Following the coming 2020 CAP framework, several case studies in European regions for cropland classification using Sentinel-2, such as the one published by Campos-Taberner et al. [47] for Valencia (Spain), will provide scientific tools necessary for the upcoming policies. The precise agricultural monitoring will follow adequate fertilization or irrigation guidelines, among others.

The advantage of Sentinel-2 multi-temporal data for crop-type classification is very relevant as it reaches up to 91–95% overall accuracies in various crop classifications [48] while single-date images show limited results. Besides the revisit time, the higher spatial and spectral resolution also improves area detection and estimations as shown in small size rice fields classification in Taiwan [49]. Additionally, in complex landscapes Sentinel-2 imagery was able to map sugarcane fields in China with over 90% of accuracy [50]. Regarding classification techniques, beside pixel based, object based has showed to be an accurate classification approach using Sentinel-2 imagery [51]. In this sense, object-based dynamic time warping was demonstrated to be more efficient in classifying crops than the pixel-based approach when using multitemporal Sentinel-2 imagery [52], nonetheless random forest seems to be more efficient for crop type classification when crop spectral variability is high [53]. In general terms, the combination of Sentinel-2 imagery and various algorithmic approaches, such as random forest, k-nearest neighbor, and support vector machine have also shown optimal classification accuracy, resulting with over 90% of overall accuracy for the various approaches [54]. Understanding the context of the crop, namely crop type spectral features and surface extension, is central to accurately making classification method decisions when using Sentinel-2 imagery.

Sentinel-2-derived research has approached yield predictions in two different fashions. On the one hand empirical models using fitting techniques of the obtained data sets (VIs and climatological data correlated with actual yields) show robust results with machine learning and random forest techniques [55,56] as well as relatively high correlation, with an  $R^2$  over 0.70, when using simple regressions with the most suitable VIs (i.e., NDVI) [57,58]. On the other hand, the radiative transfer models approach uses the calculated biophysical parameters such as LAI or FAPAR to model crop growth and predict yields [59–62]. The estimation of yields is a central issue in order to ensure food security and stabilize market demands. These performance models from Sentinel-2 imagery developed once images are available are not only aimed to provide final grain yield estimations, but they can also be used to monitor sudden changes in crop potential yields. This is especially central in taking agricultural management decisions as information is worth a little if it comes too late.

Soil is the element that facilitates the assimilation of both fertilizers and water through the roots, namely the two central factors for a proper development of crops. Studies of Sentinel-2 potential for soil monitoring have been undertaken. Soil organic carbon (SOC) can be successfully mapped in croplands using Sentinel-2 bands [63–66]. SOC prediction models were found to be overall better when formulated using the visible B4, B5, near infrared B8A, and two SWIR bands (B11, B12) [67]. Furthermore, cropland soil degradation with classification methods have also been studied [68,69].

Detecting agricultural fields affected by soil degradation or poor SOC can drive the intervention of farmers in order to prevent the failure of their fields. Thus, the decision to apply organic matter to improve soil structure or to improve field edges buffers (e.g., revegetating or reshaping them) can be driven with these Sentinel-2 data.

Irrigation and water requirements of crops are related to cultivars evapotranspiration (ET), where thermal bands are very suitable to estimate this parameter. However, in the case of Sentinel-2 only optical bands are available. In this sense, different strategies have been performed in order to determinate evapotranspiration and propose proper irrigations based on Sentinel-2 information. The combination of Sentinel-2 data with models such as FAO-56 Penman–Monteith and Aquacrop in the case of tomato have shown to be successful [70,71]. Moreover, the combination of FAO-56 Penman–Monteith with Sentinel-2 derived LAI have also shown relatively high accuracy in ET retrieval [72]. Several studies also take advantage of the red and red-edge bands to predict the crop coefficient (Kc) and thereafter use it together with  $ET_o$  [73,74]. The combination of accurate crop type mapping and ET models allows establishing regional and field-level water requirements to drive a sustainable use of water resources.

Regarding precision fertilization management, two different approaches, both fairly similar to those described above for yield, have been developed. On one hand, the use of radiative transfer models to obtain chlorophyll content [75–77], and on the other hand, using empirically correlated vegetation indexes and bands to assess the nitrogen status of the crop [78–81]. This strategy generally includes the three Sentinel-2 RE bands (705, 740, and 783 nm, Figure 2), which are very important for the retrieval of chlorophyll content [82]. The Belgian collaborative agriculture platform (<http://www.belcam.info/>) is taking advantage of these Sentinel-2 features to provide freely nitrogen status information at parcel level for farmers growing potato and winter wheat in Belgium. Information on the nitrogen absorbed by the crops and nitrogen nutrition index is provided; this information allows Belgian farmers to decide on precise N applications. Furthermore, geolocalized fertilizer spreaders can make applications on fields following Sentinel-2 sensed data in order to balance the applied amount of fertilizer depending on the crops needs per zone.

Regarding biotic stresses several approaches have been studied. Despite the insufficient resolution of Sentinel-2, 10 or 20 m, to analyze crops at leaf and individual plant scale, different strategies have been developed. On the one hand, the use of VIs to monitor effects of pests at canopy level in crops have been used in coffee [83] and wheat [84], as well as new VIs proposed with Sentinel-2 multispectral bands that have shown to sense wheat yellow rust using bands 4, 5, and 7, the so-called Red Edge Disease Stress Index (REDSI) [85]. On the other hand, classification algorithms using 2, 3, 4, and 8 bands have shown to correctly separate cotton root rot affected areas with 94.1% accuracy [86]. Furthermore, the combination of UAV and Sentinel-2 for western swamphen monitoring in rice crops have shown a relative high potential (57% of accuracy) for detecting infested fields [87]. The detection of biotic stresses can drive precise application of pesticides for both regional and field-level, furthermore the current drones' on-board carrying capacity has increased and there are various drone-based spraying experiences for crop protection against biotic stresses [88,89]. In this case, the application within the field could be driven by Sentinel-2-derived information on the most affected sites by the biotic stresses.

Concerning abiotic stresses, heavy metal stress in rice fields produced by cadmium (Cd), lead (Pb), and mercury (Hg) have shown to be effectively sensed using red edge features of Sentinel-2 [90–92]. In the case of salinity, Sentinel-2 bands 2 and 4 showed higher correlation ( $R^2 = 0.67$ ) with the electrical conductivity of fields than Landsat-8 bands [93]. Drought monitoring at 20 cm soil-depth, using combined soil moisture models, and Sentinel-2-derived NDVI has shown to correlate strongly, with determination coefficients ( $R^2$ ) ranging between 65% and 83% [94,95]. Detecting areas affected by abiotic stresses can guide determination of suitable zones for agriculture or, on the other hand, in need for phytoremediation activities or the implementation of halophyte cultivars.

**Table 3.** Examples of remote sensing information, derived from Sentinel-2, used for crop monitoring and management and the methodological approach used.

S2 for Agriculture	Sensed Trait	Method	Ref.
<b>Plant stress</b>	Cotton rot root ( <i>Phymatotrichopsis omnivore</i> )	Classification of affected areas with trained algorithms using bands 2, 3, 4, and 8	[86]
	Rice crops and Western Swampphen ( <i>Porphyrio porphyrio</i> )	S2 imagery calibrated with an UAV	[87]
	Hessian fly ( <i>Mayetiola destructor</i> ) infestation in Wheat	NDVI decrease with increasing infestation, ground, airborne, and satellite multiplatform	[84]
	Coffee leaf rust ( <i>Hemileia vastarix</i> )	Field spectral data and S2-derived VIs	[83]
	Wheat yellow rust ( <i>Puccinia striiformis</i> )	Proposed index: Red Edge Disease Stress Index (REDSI), bands 4, 5, and 7	[85]
	Metal stress in rice	Red-edge S2 bands	[90–92]
<b>Management scale</b>	Drought	S2 VI and OPTRAM soil moisture monitoring	[94,95]
	Salinity	S2 visible bands, blue and red, are sensitive to soil salinity	[93]
	Fertilization	N in crops	Biophysical retrieval of canopy chlorophyll content, also with VIs within the red-edge. Assessment of the nutritional status, NNI
	Water	Irrigation and hydric requirements (cotton, tomato, wheat, and maize)	S2 vegetation parameters, surface albedo, and crop height for FAO-56 Penman-Monteith ET estimation; red and red-edge bands to predict crop coefficients (Kc). Irrigated and rain-fed cropland differentiation. Combination of S2 and Aquacrop.
	Fields monitoring	Cropland assessments	VIs for cropping practices assessment; regional and nation-scale cropland and crop type classification with S2, time series and retrieval of biophysical and vegetation radiometric indexes (sen-2Agri)
	Soils	Soil features	Determining soil OM with VIs, and S2 bands. The wavelength of the OM spectral feature in the visible is close to S2 red band. Classification of soils degradation.
<b>Yield prediction</b>	Empirical models	VIs and yields, together with climatological data to build a dataset. Fitting techniques (regressions, random forest, machine learning) to predict yields.	[63–69]
	Radiative transfer models (RTM)	Coupling with crop functioning models, FAPAR, LAI, SLA, and light use efficiency.	[55–58] [59–62]

#### 4. Sentinel-2: Comparative Advantages and Future Work

The Sentinel-2 spatial resolution and number and positioning of spectral reflectance bands excel at cropland and crop type classification, as well as monitoring of crops phenology, namely by enabling the measurement of growth (biomass) and health (stress symptoms) separately. Furthermore, as reviewed, red-edge bands are also effective for monitoring nitrogen status, and therefore optimizing fertilization practices. The literature-proven VIs using the red-edge bands have proven robust indicators of plant N status that can be useful for agricultural monitoring. Due to Sentinel-2 red-edge bands characteristics, the biophysical parameters estimation obtained with these multispectral indexes have higher accuracy than other satellites [96]. Hence, the revisit time of 5 days, the red-edge band, and the higher resolution sets an improved management perspective.

The combination of available sensors also remains an excellent opportunity for improvement. For instance, Sentinel-2 has limitations towards estimating evapotranspiration and water stress due to its lack of any thermal bands, and as such data fusion with other complimentary sensors may generate opportunities to coordinate data with available thermal bands such as NASA's MODIS or Landsat-8 [97]. However, the spatial resolution of their thermal bands (1 km and 100 m, respectively) might be a downside for some agricultural applications that require more detail. Other possible approaches are the combination of optical imagery of Sentinel-2 and thermal bands from the sensors on board Sentinel-3 to obtain crop evapotranspiration, which is promising for bridging the gap and monitoring evapotranspiration with a higher resolution [98]. Besides that, the combination of Sentinel-2 optical bands and Sentinel-1 C-band Synthetic Aperture Radar has successfully been used for mapping soil moisture and irrigation demands [99].

Another challenge is pest monitoring, where pest early warning systems are useful to predict outbreaks and are generally based on weather information, and intensive and time-consuming field work. Being able to use Sentinel-2, or other remote sensing instruments, to assess pest distribution would be of great importance, considering the current globalization of invasive pests and their sudden effects. Nonetheless, a bottleneck is found here, on the one hand the pixel resolution makes it very difficult to monitor pests and generally only when damage is significant it can be sensed. On the other hand, it is a challenge to differentiate the reflectance patterns of plants due to biotic stresses and see what the exact cause of the stress is. The combination of ground, drone, and satellite sensors might be a way to go, with some regional and technical limitations that need to be overcome.

Soils are generally not considered at the expense of other parameters. However, the importance of SOC, which is gaining interest within the sustainable agriculture paradigm, can be monitored with Sentinel-2. This can help to overcome this fact and move forward into precision agriculture and soil sustainability. Regarding soils, monitoring cropland pollution due to heavy metals is important to ensure food safety. It is relevant to see that three papers reviewed in this issue had rice and China as the case study [90–92] where serious pollution problems in this sense happen [100].

Sen2-Agri has set the structure for a coordinated field monitoring platform which will likely lead to a multi-country precision agriculture initiative. In developing countries, on the other hand, this platform, and similar ones using Sentinel-2 data, can contribute to food safety and help farmers in managing their crops. Due to the small fields, intercropping practices and terrain-derived issues crop monitoring strategies need to be tuned in order to successfully manage these areas.

#### 5. Conclusions

Several previous review papers exist on the use of remote sensing for agriculture [101–103], however, to our knowledge there has not yet been a specific review article on Sentinel-2 features for agriculture. We are convinced that may prove beneficial to researchers and agrarian institutions entering this field to understand the features and applications of Sentinel-2 in crops monitoring alone and as compared to other previous and current satellites. The comparison of the Sentinel-2 A + B constellation with other orbiting satellites with similar features highlights its improved capacities within precision agriculture. In this sense, we detail several applications on the detection and

management of biotic and abiotic stresses, crop water requirements, crop type classifications, and the estimation of crop yields and nitrogen status. Furthermore, besides the promising technical features of Sentinel-2, it is worth mentioning the public availability of its imagery. By making this data accessible to all interested institutions, the potential benefits of precision agriculture are greater and may better provide the mechanisms for advising farmers towards a more productive and sustainable management, thus, contribute to economic and ecological sustainability worldwide.

**Author Contributions:** J.S., M.L.B., J.L.A. and S.C.K. conceived the review; J.S. and M.L.B. performed the literature research; M.L.B. made the figures; S.C.K. and J.L.A. analyzed the scientific rigor; S.C.K. coordinated the structure and scope of the topic; J.S. and S.C.K. wrote the paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Spanish projects AGL2016-76527-R and IRUEC PCIN-2017-063 from the Ministerio de Economía y Competitividad (MINECO, Spain) and by the support of Catalan Institution for Research and Advanced Studies (ICREA, Generalitat de Catalunya, Spain), through the ICREA Academia Program.

**Acknowledgments:** This study was supported in part by the Spanish project AGL2016-76527-R from the Ministerio de Economía y Competitividad (MINECO, Spain) and by the support of Catalan Institution for Research and Advanced Studies (ICREA, Generalitat de Catalunya, Spain), through the ICREA Academia Program.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Berthet, E.T.; Bretagnolle, V.; Lavorel, S.; Sabatier, R.; Tichit, M.; Segrestin, B. Applying ecological knowledge to the innovative design of sustainable agroecosystems. *J. Appl. Ecol.* **2019**, *56*, 44–51. [\[CrossRef\]](#)
- General Assembly Transforming Our World: The 2030 Agenda for Sustainable Development*; United Nations: New York, NY, USA, 2015.
- Matton, N.; Canto, G.S.; Waldner, F.; Valero, S.; Morin, D.; Inglada, J.; Arias, M.; Bontemps, S.; Koetz, B.; Defourny, P. An Automated Method for Annual Cropland Mapping along the Season for Various Globally-Distributed Agrosystems Using High Spatial and Temporal Resolution Time Series. *Remote Sens.* **2015**, *7*, 13208–13232. [\[CrossRef\]](#)
- Pierce, F.J.; Nowak, P. Aspects of Precision Agriculture. *Adv. Agron.* **1999**, *67*, 1–85.
- Araus, J.L.; Kefauver, S.C. Breeding to adapt agriculture to climate change: Affordable phenotyping solutions. *Curr. Opin. Plant Biol.* **2018**, *45*, 237–247. [\[CrossRef\]](#)
- Mulla, D.J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst. Eng.* **2013**, *114*, 358–371. [\[CrossRef\]](#)
- Yang, J.; Gong, P.; Fu, R.; Zhang, M.; Chen, J.; Liang, S.; Xu, B.; Shi, J.; Dickinson, R. The role of satellite remote sensing in climate change studies. *Nat. Clim. Chang.* **2013**, *3*, 875–883. [\[CrossRef\]](#)
- Zhang, N.; Wang, M.; Wang, N. Precision agriculture—A worldwide overview. *Comput. Electron. Agric.* **2002**, *36*, 113–132. [\[CrossRef\]](#)
- Adams, J.B.; Gillespie, A.R. *Remote Sensing of Landscapes with Spectral Images: A Physical Modeling Approach*; Cambridge University Press: Cambridge, UK, 2006; pp. 1–362.
- Avery, T.E.; Berlin, G.L. *Fundamentals of Remote Sensing and Airphoto Interpretation*; Macmillan: London, UK, 1992.
- Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [\[CrossRef\]](#)
- Rautiainen, M.; Stenberg, P. Application of photon recollision probability in coniferous canopy reflectance simulations. *Remote Sens. Environ.* **2005**, *96*, 98–107. [\[CrossRef\]](#)
- Disney, M.; Lewis, P.; Saich, P. 3D modelling of forest canopy structure for remote sensing simulations in the optical and microwave domains. *Remote Sens. Environ.* **2006**, *100*, 114–132. [\[CrossRef\]](#)
- Steddom, K.; Bredehoeft, M.W.; Khan, M.; Rush, C.M. Comparison of visual and multispectral radiometric disease evaluations of *Cercospora* leaf spot of sugar beet. *Plant Dis.* **2005**, *89*, 153–158. [\[CrossRef\]](#)
- Aparicio, N.; Villegas, D.; Araus, J.L.; Casadesús, J.; Royo, C. Relationship between growth traits and spectral vegetation indices in durum wheat. *Crop Sci.* **2002**, *42*, 1547–1555. [\[CrossRef\]](#)
- Peñuelas, J.; Filella, I. Reflectance assessment of mite effects on apple trees. *Int. J. Remote Sens.* **1995**, *16*, 2727–2733. [\[CrossRef\]](#)



17. Blazquez, C.H.; Edwards, G.J. Spectral reflectance of healthy and diseased watermelon leaves. *Ann. Appl. Biol.* **1986**, *108*, 243–249. [[CrossRef](#)]
18. Sharp, E.L. Monitoring Cereal Rust Development with a Spectral Radiometer. *Phytopathology* **1985**, *75*, 936. [[CrossRef](#)]
19. Miller, J.R.; Hare, E.W.; Wu, J. Quantitative characterization of the vegetation red edge reflectance 1. An inverted-gaussian reflectance model. *Int. J. Remote Sens.* **1990**, *11*, 1755–1773. [[CrossRef](#)]
20. Curran, P.J.; Dungan, J.L.; Macler, B.A.; Plummer, S.E. The effect of a red leaf pigment on the relationship between red edge and chlorophyll concentration. *Remote Sens. Environ.* **1991**, *35*, 69–76. [[CrossRef](#)]
21. Curran, P.J.; Dungan, J.L.; Gholz, H.L. Exploring the relationship between reflectance red edge and chlorophyll content in slash pine. *Tree Physiol.* **1990**, *7*, 33–48. [[CrossRef](#)]
22. Kokaly, R.F. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sens. Environ.* **2001**, *75*, 153–161. [[CrossRef](#)]
23. Guerif, M.; Houles, V.; Baret, F. Remote sensing and detection of nitrogen status in crops. Application to precise nitrogen fertilization. In Proceedings of the 4th International Symposium on Intelligent Information Technology in Agriculture, Beijing, China, 26–29 October 2007.
24. Chen, P.; Haboudane, D.; Tremblay, N.; Wang, J.; Vigneault, P.; Li, B. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sens. Environ.* **2010**, *114*, 1987–1997. [[CrossRef](#)]
25. Gitelson, A.A.; Merzlyak, M.N. Signature analysis of leaf reflectance spectra: Algorithm development for remote sensing of chlorophyll. *J. Plant Physiol.* **1996**, *148*, 494–500. [[CrossRef](#)]
26. Liang, S.; Li, X.; Wang, J. *Advanced Remote Sensing*; Academic Press: Cambridge, MA, USA, 2012; ISBN 9780123859549.
27. Homolová, L.; Malenovsky, Z.; Clevers, J.G.P.W.; García-Santos, G.; Schaepman, M.E. Review of optical-based remote sensing for plant trait mapping. *Ecol. Complex.* **2013**, *15*, 1–16. [[CrossRef](#)]
28. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
29. SUHET. *Sentinel-2 User Handbook*; European Commission: Brussels, Belgium, 2015.
30. Le Maire, G.; François, C.; Dufrêne, E. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sens. Environ.* **2004**, *89*, 1–28. [[CrossRef](#)]
31. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sens.* **2017**, *2017*, 1353691. [[CrossRef](#)]
32. Bannari, A.; Morin, D.; Bonn, F.; Huete, A.R. A review of vegetation indices. *Remote Sens. Rev.* **1995**, *13*, 95–120. [[CrossRef](#)]
33. Hunt, E.R.; Cavigelli, M.; Daughtry, C.S.T.; McMurtrey, J.E.; Walthall, C.L. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis. Agric.* **2005**, *6*, 359–378. [[CrossRef](#)]
34. Hunt, E.R.; Doraiswamy, P.C.; McMurtrey, J.E.; Daughtry, C.S.T.; Perry, E.M.; Akhmedov, B. A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *21*, 103–112. [[CrossRef](#)]
35. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [[CrossRef](#)]
36. Dash, J.; Curran, P.J. Evaluation of the MERIS terrestrial chlorophyll index (MTCI). *Adv. Sp. Res.* **2007**, *39*, 100–104. [[CrossRef](#)]
37. Gitelson, A.A.; Viña, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*. [[CrossRef](#)]
38. Vincini, M.; Frazzi, E.; D'Alessio, P. A broad-band leaf chlorophyll vegetation index at the canopy scale. *Precis. Agric.* **2008**, *9*, 303–319. [[CrossRef](#)]
39. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 83–92. [[CrossRef](#)]

40. Congedo, L. *Semi-Automatic Classification Plugin Semi-Automatic Classification Plugin Documentation*; SCP: Rome, Italy, 2017; pp. 3, 159, 160, 154, 159, 206.
41. ESA. *The Copernicus Space Component: Sentinels Data Products List*; ESA: Paris, France, 2014; pp. 1–18.
42. Weiss, M.; Baret, F. *S2ToolBox Level 2 Products: LAI, FAPAR, FCOVER—Version 1.1*; INRA: Avignon, France, 2016; 53p.
43. Verrelst, J.; Rivera, J.P.; Veroustraete, F.; Muñoz-Mari, J.; Clevers, J.G.P.W.; Camps-Valls, G.; Moreno, J. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods—A comparison. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 260–272. [[CrossRef](#)]
44. Defourny, P.; Bontemps, S.; Bellemans, N.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Nicola, L.; Rabaute, T.; et al. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sens. Environ.* **2019**, *221*, 551–568. [[CrossRef](#)]
45. Lambert, M.J.; Traoré, P.C.S.; Blaes, X.; Baret, P.; Defourny, P. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens. Environ.* **2018**, *216*, 647–657. [[CrossRef](#)]
46. Commission, E. EU Budget: The Common Agricultural Policy beyond 2020. Available online: [https://ec.europa.eu/commission/presscorner/detail/en/MEMO\\_18\\_3974](https://ec.europa.eu/commission/presscorner/detail/en/MEMO_18_3974) (accessed on 26 April 2020).
47. Campos-Taberner, M.; García-Haro, F.J.; Martínez, B.; Sánchez-Ruiz, S.; Gilabert, M.A. A copernicus sentinel-1 and sentinel-2 classification framework for the 2020+ European common agricultural policy: A case study in València (Spain). *Agriculture* **2019**, *9*, 556. [[CrossRef](#)]
48. Vuolo, F.; Neuwirth, M.; Immitzer, M.; Atzberger, C.; Ng, W.T. How much does multi-temporal Sentinel-2 data improve crop type classification? *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *72*, 122–130. [[CrossRef](#)]
49. Son, N.T.; Chen, C.F.; Chen, C.R.; Guo, H.Y. Classification of multitemporal Sentinel-2 data for field-level monitoring of rice cropping practices in Taiwan. *Adv. Sp. Res.* **2020**, *65*, 1910–1921. [[CrossRef](#)]
50. Wang, M.; Liu, Z.; Ali Baig, M.H.; Wang, Y.; Li, Y.; Chen, Y. Mapping sugarcane in complex landscapes by integrating multi-temporal Sentinel-2 images and machine learning algorithms. *Land Use Policy* **2019**, *88*, 104190. [[CrossRef](#)]
51. Cai, Y.; Lin, H.; Zhang, M. Mapping paddy rice by the object-based random forest method using time series Sentinel-1/Sentinel-2 data. *Adv. Sp. Res.* **2019**, *64*, 2233–2244. [[CrossRef](#)]
52. Csillik, O.; Belgiu, M.; Asner, G.P.; Kelly, M. Object-based time-constrained dynamic time warping classification of crops using Sentinel-2. *Remote Sens.* **2019**, *11*, 1257. [[CrossRef](#)]
53. Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2018**, *204*, 509–523. [[CrossRef](#)]
54. Noi, P.T.; Kappas, M. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. *Sensors* **2018**, *18*, 18.
55. Gómez, D.; Salvador, P.; Sanz, J.; Casanova, J.L. Potato Yield Prediction Using Machine Learning Techniques and Sentinel 2 Data. *Remote Sens.* **2019**, *11*, 1745. [[CrossRef](#)]
56. Hunt, M.L.; Blackburn, G.A.; Carrasco, L.; Redhead, J.W.; Rowland, C.S. High resolution wheat yield mapping using Sentinel-2. *Remote Sens. Environ.* **2019**, *233*, 111410. [[CrossRef](#)]
57. Fieuzal, R.; Bustillo, V.; Collado, D.; Dedieu, G. Combined use of multi-temporal Landsat-8 and sentinel-2 images for wheat yield estimates at the intra-plot spatial scale. *Agriculture* **2020**, *10*, 327. [[CrossRef](#)]
58. Toscano, P.; Castrignanò, A.; Di Gennaro, S.F.; Vonella, A.V.; Ventrella, D.; Matese, A. A precision agriculture approach for durum wheat yield assessment using remote sensing data and yield mapping. *Agriculture* **2019**, *9*, 437. [[CrossRef](#)]
59. Battude, M.; Al Bitar, A.; Morin, D.; Cros, J.; Huc, M.; Marais Sicre, C.; Le Dantec, V.; Demarez, V. Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sens. Environ.* **2016**, *184*, 668–681. [[CrossRef](#)]
60. Novelli, F.; Vuolo, F. Assimilation of sentinel-2 leaf area index data into a physically-based crop growth model for yield estimation. *Agriculture* **2019**, *9*, 255. [[CrossRef](#)]

61. Habyarimana, E.; Piccard, I.; Catellani, M.; De Franceschi, P.; Dall'Agata, M. Towards predictive modeling of sorghum biomass yields using fraction of absorbed photosynthetically active radiation derived from sentinel-2 satellite imagery and supervised machine learning techniques. *Agronomy* **2019**, *9*, 203. [\[CrossRef\]](#)
62. He, L.; Mostovoy, G. Cotton Yield Estimate Using Sentinel-2 Data and an Ecosystem Model over the Southern US. *Remote Sens.* **2019**, *11*, 2000. [\[CrossRef\]](#)
63. Castaldi, F.; Hueni, A.; Chabrilat, S.; Ward, K.; Buttafuoco, G.; Bomans, B.; Vreys, K.; Brell, M.; van Wesemael, B. Evaluating the capability of the Sentinel 2 data for soil organic carbon prediction in croplands. *ISPRS J. Photogramm. Remote Sens.* **2019**, *147*, 267–282. [\[CrossRef\]](#)
64. Castaldi, F.; Chabrilat, S.; Don, A.; van Wesemael, B. Soil Organic Carbon Mapping Using LUCAS Topsoil Database and Sentinel-2 Data: An Approach to Reduce Soil Moisture and Crop Residue Effects. *Remote Sens.* **2019**, *11*, 2121. [\[CrossRef\]](#)
65. Vaudour, E.; Gomez, C.; Loiseau, T.; Baghdadi, N.; Loubet, B.; Arrouays, D.; Ali, L.; Lagacherie, P. The Impact of Acquisition Date on the Prediction Performance of Topsoil Organic Carbon from Sentinel-2 for Croplands. *Remote Sens.* **2019**, *11*, 2143. [\[CrossRef\]](#)
66. Gholizadeh, A.; Žižala, D.; Saberioon, M.; Borůvka, L. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. *Remote Sens. Environ.* **2018**, *218*, 89–103. [\[CrossRef\]](#)
67. Vaudour, E.; Gomez, C.; Fouad, Y.; Lagacherie, P. Sentinel-2 image capacities to predict common topsoil properties of temperate and Mediterranean agroecosystems. *Remote Sens. Environ.* **2019**, *223*, 21–33. [\[CrossRef\]](#)
68. Rosero-Vlasova, O.A.; Vlassova, L.; Pérez-Cabello, F.; Montorio, R.; Nadal-Romero, E. Modeling soil organic matter and texture from satellite data in areas affected by wildfires and cropland abandonment in Aragón, Northern Spain. *J. Appl. Remote Sens.* **2018**, *12*, 1. [\[CrossRef\]](#)
69. Žižala, D.; Juřicová, A.; Zádorová, T.; Zelenková, K.; Minařík, R. Mapping soil degradation using remote sensing data and ancillary data: South-East Moravia, Czech Republic. *Eur. J. Remote Sens.* **2019**, *52*, 108–122. [\[CrossRef\]](#)
70. Vanino, S.; Nino, P.; De Michele, C.; Falanga Bolognesi, S.; D'Urso, G.; Di Bene, C.; Pennelli, B.; Vuolo, F.; Farina, R.; Pulighe, G.; et al. Capability of Sentinel-2 data for estimating maximum evapotranspiration and irrigation requirements for tomato crop in Central Italy. *Remote Sens. Environ.* **2018**, *215*, 452–470. [\[CrossRef\]](#)
71. Dalla Marta, A.; Chirico, G.B.; Falanga Bolognesi, S.; Mancini, M.; D'Urso, G.; Orlandini, S.; De Michele, C.; Altobelli, F. Integrating Sentinel-2 Imagery with AquaCrop for Dynamic Assessment of Tomato Water Requirements in Southern Italy. *Agronomy* **2019**, *9*, 404. [\[CrossRef\]](#)
72. Pasqualotto, N.; D'Urso, G.; Bolognesi, S.F.; Belfiore, O.R.; Van Wittenberghe, S.; Delegido, J.; Pezzola, A.; Winschel, C.; Moreno, J. Retrieval of evapotranspiration from sentinel-2: Comparison of vegetation indices, semi-empirical models and SNAP biophysical processor approach. *Agronomy* **2019**, *9*, 663. [\[CrossRef\]](#)
73. Rozenstein, O.; Haymann, N.; Kaplan, G.; Tanny, J. Estimating cotton water consumption using a time series of Sentinel-2 imagery. *Agric. Water Manag.* **2018**, *207*, 44–52. [\[CrossRef\]](#)
74. Rozenstein, O.; Haymann, N.; Kaplan, G.; Tanny, J. Validation of the cotton crop coefficient estimation model based on Sentinel-2 imagery and eddy covariance measurements. *Agric. Water Manag.* **2019**, *223*, 105715. [\[CrossRef\]](#)
75. Clevers, J.G.P.W.; Kooistra, L.; van den Brande, M.M.M. Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. *Remote Sens.* **2017**, *9*, 663. [\[CrossRef\]](#)
76. Nutini, F.; Confalonieri, R.; Crema, A.; Movedi, E.; Paleari, L.; Stavrakoudis, D.; Boschetti, M. An operational workflow to assess rice nutritional status based on satellite imagery and smartphone apps. *Comput. Electron. Agric.* **2018**, *154*, 80–92. [\[CrossRef\]](#)
77. Delloye, C.; Weiss, M.; Defourny, P. Retrieval of the canopy chlorophyll content from Sentinel-2 spectral bands to estimate nitrogen uptake in intensive winter wheat cropping systems. *Remote Sens. Environ.* **2018**, *216*, 245–261. [\[CrossRef\]](#)
78. Vizzari, M.; Santa, F.; Benincasa, P. Sentinel 2-based nitrogen VRT fertilization in wheat: Comparison between traditional and simple precision practices. *Agronomy* **2019**, *9*, 663. [\[CrossRef\]](#)
79. Ramoelo, A.; Cho, M.A. Explaining leaf nitrogen distribution in a semi-arid environment predicted on sentinel-2 imagery using a field spectroscopy derived model. *Remote Sens.* **2018**, *10*, 269. [\[CrossRef\]](#)



80. Chemura, A.; Mutanga, O.; Odindi, J.; Kutwayo, D. Mapping spatial variability of foliar nitrogen in coffee (*Coffea arabica* L.) plantations with multispectral Sentinel-2 MSI data. *ISPRS J. Photogramm. Remote Sens.* **2018**, *138*, 1–11. [\[CrossRef\]](#)
81. Söderström, M.; Piikki, K.; Stenberg, M.; Stadig, H.; Martinsson, J. Producing nitrogen (N) uptake maps in winter wheat by combining proximal crop measurements with Sentinel-2 and DMC satellite images in a decision support system for farmers. *Acta Agric. Scand. Sect. B Soil Plant Sci.* **2017**, *67*, 637–650. [\[CrossRef\]](#)
82. Delegido, J.; Verrelst, J.; Alonso, L.; Moreno, J. Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. *Sensors* **2011**, *11*, 7063–7081. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Chemura, A.; Mutanga, O.; Dube, T. Separability of coffee leaf rust infection levels with machine learning methods at Sentinel-2 MSI spectral resolutions. *Precis. Agric.* **2017**, *18*, 859–881. [\[CrossRef\]](#)
84. Bhattarai, G.P.; Schmid, R.B.; McCornack, B.P. Remote Sensing Data to Detect Hessian Fly Infestation in Commercial Wheat Fields. *Sci. Rep.* **2019**, *9*, 6109. [\[CrossRef\]](#)
85. Zheng, Q.; Huang, W.; Cui, X.; Shi, Y.; Liu, L. New spectral index for detecting wheat yellow rust using sentinel-2 multispectral imagery. *Sensors* **2018**, *18*, 868. [\[CrossRef\]](#)
86. Song, X.; Yang, C.; Wu, M.; Zhao, C.; Yang, G.; Hoffmann, W.C.; Huang, W. Evaluation of Sentinel-2A satellite imagery for mapping cotton root rot. *Remote Sens.* **2017**, *9*, 906. [\[CrossRef\]](#)
87. Pla, M.; Bota, G.; Duane, A.; Balagué, J.; Curcó, A.; Gutiérrez, R.; Brotons, L. Calibrating Sentinel-2 Imagery with Multispectral UAV Derived Information to Quantify Damages in Mediterranean Rice Crops Caused by Western Swamphen (*Porphyrio porphyrio*). *Drones* **2019**, *3*, 45. [\[CrossRef\]](#)
88. Meivel, S.M.; Professor, A.; Maguteeswaran, R.; Gandhiraj, N.B.; Srinivasan, G. Quadcopter UAV Based Fertilizer and Pesticide Spraying System. *Int. Acad. Res. J. Eng. Sci.* **2016**, *1*, 2414–2422.
89. Façal, B.S.; Costa, F.G.; Pessin, G.; Ueyama, J.; Freitas, H.; Colombo, A.; Fini, P.H.; Villas, L.; Osório, F.S.; Vargas, P.A.; et al. The use of unmanned aerial vehicles and wireless sensor networks for spraying pesticides. *J. Syst. Archit.* **2014**, *60*, 393–404. [\[CrossRef\]](#)
90. Zhang, Z.; Liu, M.; Liu, X.; Zhou, G. A new vegetation index based on multitemporal sentinel-2 images for discriminating heavy metal stress levels in rice. *Sensors* **2018**, *18*, 2172. [\[CrossRef\]](#)
91. Liu, M.; Wang, T.; Skidmore, A.K.; Liu, X. Heavy metal-induced stress in rice crops detected using multi-temporal Sentinel-2 satellite images. *Sci. Total Environ.* **2018**, *637–638*, 18–29. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Sun, N.; Wang, P.; Huang, F.; Li, B. Developing an integrated index based on phenological metrics for evaluating cadmium stress in rice using Sentinel-2 data. *J. Appl. Remote Sens.* **2018**, *12*, 1. [\[CrossRef\]](#)
93. Davis, E.; Wang, C.; Dow, K. Comparing Sentinel-2 MSI and Landsat 8 OLI in soil salinity detection: A case study of agricultural lands in coastal North Carolina. *Int. J. Remote Sens.* **2019**, *40*, 6134–6153. [\[CrossRef\]](#)
94. Mananze, S.; Pôças, I. Agricultural drought monitoring based on soil moisture derived from the optical trapezoid model in Mozambique. *J. Appl. Remote Sens.* **2019**, *13*, 1–16. [\[CrossRef\]](#)
95. West, H.; Quinn, N.; Horswell, M.; White, P. Assessing vegetation response to soil moisture fluctuation under extreme drought using sentinel-2. *Water* **2018**, *10*, 838. [\[CrossRef\]](#)
96. Xie, Q.; Dash, J.; Huete, A.; Jiang, A.; Yin, G.; Ding, Y.; Peng, D.; Hall, C.C.; Brown, L.; Shi, Y.; et al. Retrieval of crop biophysical parameters from Sentinel-2 remote sensing imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *80*, 187–195. [\[CrossRef\]](#)
97. Mokhtari, A.; Noory, H.; Pourshakouri, F.; Haghighatmehr, P.; Afrasiabian, Y.; Razavi, M.; Fereydooni, F.; Sadeghi Naeni, A. Calculating potential evapotranspiration and single crop coefficient based on energy balance equation using Landsat 8 and Sentinel-2. *ISPRS J. Photogramm. Remote Sens.* **2019**, *154*, 231–245. [\[CrossRef\]](#)
98. Guzinski, R.; Nieto, H. Evaluating the feasibility of using Sentinel-2 and Sentinel-3 satellites for high-resolution evapotranspiration estimations. *Remote Sens. Environ.* **2019**, *221*, 157–172. [\[CrossRef\]](#)
99. Bousbih, S.; Zribi, M.; Hajj, M.; El Baghdadi, N.; Lili-Chabaane, Z.; Gao, Q.; Fanise, P. Soil moisture and irrigation mapping in a semi-arid region, based on the synergetic use of Sentinel-1 and Sentinel-2 data. *Remote Sens.* **2018**, *10*, 1953. [\[CrossRef\]](#)
100. Sodango, T.H.; Li, X.; Sha, J.; Bao, Z. Review of the spatial distribution, source and extent of heavy metal pollution of soil in China: Impacts and mitigation approaches. *J. Heal. Pollut.* **2018**, *8*, 53–70. [\[CrossRef\]](#)
101. Atzberger, C. Correction: Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sens.* **2013**, *5*, 949–981. [\[CrossRef\]](#)

102. Bégué, A.; Arvor, D.; Bellon, B.; Betbeder, J.; de Aballeyra, D.; Ferraz, R.P.D.; Lebourgeois, V.; Lelong, C.; Simões, M.; Verón, S.R. Remote sensing and cropping practices: A review. *Remote Sens.* **2018**, *10*, 99. [[CrossRef](#)]
103. Weiss, M.; Jacob, F.; Duveiller, G. Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* **2020**, *236*, 111402. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



## **Annex 2.** Pèrdua de cobertura forestal a l'àrea protegida del volcà de Toluca (Mèxic) arran del canvi a una categoria menys restrictiva l'any 2013

**Annex 2.** Forest cover loss in the Nevado de Toluca volcano protected area (Mexico) after the change to a less restrictive category in 2013

Andrea González-Fernández<sup>1,2</sup>, Joel Segarra<sup>3</sup>, Armando Sunny<sup>4</sup>, Stephane Couturier<sup>1\*</sup>

<sup>1</sup>Laboratory for Geospatial Analysis (LAGE), Geography Institute, National Autonomous University of Mexico (UNAM), Circuito Exterior s/n, Coyoacán, Cd. Universitaria, 04510 Mexico City, Mexico

<sup>2</sup>Autonomous Metropolitan University - campus Lerma, Hidalgo Pte. 46, Col. La Estación, 52006 Lerma, Estado de México, Mexico

<sup>3</sup>Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, Avinguda Diagonal, 643, 08028 Barcelona, Spain


<sup>4</sup>Center for Research in Applied Biological Sciences, Autonomous University of the State of Mexico, Instituto Literario 100, Colonia Centro, 50000 Toluca, State of Mexico, Mexico

\*Correspondence

Publicat a la revista Biodiversity and Conservation



# Forest cover loss in the Nevado de Toluca volcano protected area (Mexico) after the change to a less restrictive category in 2013

Andrea González-Fernández<sup>1,2</sup> · Joel Segarra<sup>3</sup> · Armando Sunny<sup>4</sup> ·  
Stephane Couturier<sup>1</sup> 

Received: 30 March 2021 / Revised: 26 December 2021 / Accepted: 2 January 2022 /

Published online: 11 February 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2022

## Abstract

In 2013, the protection status of the Nevado de Toluca Volcano was changed from National Park to a less restrictive category. Much controversy has arisen surrounding this decision as the new category allows forest harvesting practices in most of the natural protected area. We assessed the forest cover loss in the Nevado de Toluca protected area and a 20 km buffer zone around it, before and after the change in protection status, and we established a map of forest types in view of assessing biodiversity vulnerability to the protection change in the forest harvesting areas. The forest type classification was carried out applying a Random Forest algorithm to Sentinel-2 imagery. The forest cover loss was assessed using the Landsat-based Global Forest Watch product during the period 2001–2019. Forest cover loss increased after the change of the Nevado de Toluca protection status; 39.7 hectares/year ( $\pm 16.7$  ha/year) between 2001 and 2013 versus 106.0 hectares/year ( $\pm 40.2$  ha/year) between 2014 and 2019. Clear-cut deforestation occurred in 49.6 hectares during year 2018. According to our forest map, 7171 ha (54.8%) of the *Abies religiosa* forest and 6511 ha (32.5%) of the pine forests of the Nevado de Toluca protected area are inside the forest harvesting zonation. Also, forest cover loss around the natural protected area has increased, especially in 2018, indicating that allowing forest harvesting practices in the natural protected area has not reduced the pressure in the surrounding forests either. We ponder the consequences of the category change and give some suggestions for the future of the forests in the Nevado de Toluca protected area.

**Keywords** *Abies religiosa* forests · Natural protected areas · Forest harvesting · Global forest cover change · Logging · Satellite imagery

---

Communicated by Daniel Sanchez Mata.

---

This article belongs to the Topical Collection: Forest and plantation biodiversity.

---

✉ Stephane Couturier  
andres@igg.unam.mx

Extended author information available on the last page of the article

## Introduction

Forests play a critical role in the Earth's terrestrial C sinks and exert strong control on the evolution of atmospheric CO<sub>2</sub> (Pan et al. 2011) they are also extremely important in the regulation of the hydrological cycle as the sum emission of water vapour from forests (transpiration and evaporation) typically surpasses that from other vegetation and even from open water (Sheil 2018). Moreover, forest soils tend to be deeper, more porous, and more permeable than other soils, increasing moisture infiltration and storage and reducing run-off (Malmer et al. 2010). Unfortunately, 2.3 million square kilometres of forest have been lost worldwide between 2000 and 2012 (Hansen et al. 2013), following the consumption trends of globalized commodities to a large extent (Couturier 2019). Online platforms like Global Forest Watch and the freely available global forest cover change (GFCC) product (Hansen et al. 2013) provide local data on forests around the world and facilitate near real-time rapid assessments by the informed citizenship, preventing land use change from going unnoticed and strengthening local initiatives for forest restoration (Osorno-Covarrubias et al. 2018).

Tropical mountains play an important role as refuges for cold and temperate affinity species which can persist through global climate fluctuations by undertaking comparatively short elevational shifts instead of long latitudinal movements (Sandel et al. 2011). Mountains with these characteristics have been identified as important centres for generation and maintenance of biodiversity during the last glacial periods (Mastretta-Yanes et al. 2015) and may effectively shelter many cold and temperate affinity species into the next century given the temperature increase due to climate change (Loarie et al. 2009). Forests and alpine grasslands that inhabit these mountains are distributed in “sky-islands” pattern: highlands of cold climates separated by warmer lowlands (Mastretta-Yanes et al. 2015). Mexican highlands, in addition to being tropical mountains, are also located in the transition zone between the Nearctic and Neotropical biogeographical realms (Morrone and Márquez 2001), representing a hotspot of temperate biodiversity (Myers et al. 2000) with approximately 50% of all *Pinus* species and 30% of all *Quercus* species (Valencia 2004). Mexican highlands also contain *Abies* forests; one of the most threatened and fragmented forest types in Mexico which occupy a relatively small area (143,579 ha) of the Mexican surface (Martínez-Méndez et al. 2016).

The Nevado de Toluca Volcano is a tropical Mexican mountain located in the Trans-Mexican Volcanic Belt and contains approximately 10,000 ha of well-preserved *Abies religiosa* forest (northwestern part of the volcano), which represents one of the largest continuous extensions of this ecosystem in Mexico (Mastretta-Yanes et al. 2014). The other dominant forest type is pine (*P. hartwegii*, *P. montezumae*, and *P. pseudostrobus* to a lesser extent). There are also oak forests (*Q. laurina* and *Q. rugosa*), alder forests (*Alnus jorullensis*), mixed *Abies-Pinus* and mixed *Pinus-Quercus* forests (CONANP 2016). In this natural protected area, several species are considered at risk according to Nom-059-Semarnat-2010 (SEMARNAT 2010): 6 fungus (e.g. *Amanita muscaria*, *Psathyrella spadicea* and species of the *Morchella* genus), 19 plants (e.g. *Comarostaphylis discolor* and *Juniperus sabinoides*), 13 birds (e.g. *Accipiter cooperii*, *Falco peregrinus* and *Xenospiza bailei*), 10 reptiles (e.g. *Barisia imbricata*, *Sceloporus grammicus*, and *Thamnophis scalaris*), 10 amphibians (e.g. *Ambystoma rivulare*, *Hyla plicata*, and *Lithobates montezumae*) and 8 mammals (e.g. *Cratogeomys fumosus* and *Glaucomys volans*) (Sánchez-Jasso et al. 2013; CONANP 2016). Moreover, *Abies religiosa* forests of the volcano are of special importance as they contain a critically endangered salamander species endemic

to this volcano (*Pseudoeurycea robertsi*, IUCN SSC Amphibian Specialist Group 2016; Sunny et al. 2021a, b) as well as other threatened or near threatened salamander species (*Aquiloerycea cephalica*, *Isthmura bellii* and *P. leprosa*; Bille 2009; IUCN SSC Amphibian Specialist Group 2016; Sunny et al. 2021a, b). These forests also receive the migratory colonies of the emblematic monarch butterfly (Saunders et al. 2018), being essential for the conservation of this species. In addition, this volcano supplies water to two of the main Mexican river basins; the “Balsas” basin to the south-west and the “Lerma-Chapala-Santiago” basin to the north-east (Rojas-Merced et al. 2007).

In 2013, the protection status of the Nevado de Toluca Volcano was downgraded from National Park to Flora and Fauna Protected Area (“Área de Protección de Flora y Fauna Nevado de Toluca” in Spanish), a less restrictive category (DOF 2013). In the study justifying the status change, it was argued that the protected area had suffered a “widespread deterioration of such a magnitude that it was expected to completely lose its function in the northern part within a decade or two” (CONANP 2013, cited by Mastretta Yanes et al. 2014). Nevertheless, some studies suggest that the arguments used to justify the change were based on erroneous, insufficient, and biased data (Mastretta-Yanes et al. 2014; Lebreton and Imbernon 2017), and the discourse about the degradation of the Nevado de Toluca protected area was exaggerated in view of legitimizing a political decision (Depraz et al. 2017; Lebreton and Imbernon 2017). In the justification study (CONANP 2013), excessive illegal logging is presented as a consequence of the highly restrictive protection status of the protected area, therefore one of the main objectives of the change was to regulate logging under sustainable restricted use schemes. Nevertheless, a controversy arose around the scientific evidence supporting the category change, and irregularities were reported in the public consultation that preceded the change (Héritier and Lebreton 2017; Lebreton and Imbernon 2017). The Flora and Fauna Protected Area new category allows selective logging in most of the natural protected area, which could in fact result in an incentive for former actors of illegal logging to accelerate deforestation in the area.

In view of the above, the objective of this study is threefold: (1) to assess forest cover loss in the Nevado de Toluca protected area and a 20 km buffer zone around it before and after the protection status was changed; (2) to map forest cover types in the Nevado de Toluca protected area at 10 m spatial resolution and document the proportion of forest types within the forest harvesting areas, that could possibly be more vulnerable to the protection category change; and (3) to discuss the results in terms of forest management in the protected area.

## Materials and methods

### Study site

The Nevado de Toluca Volcano (NTV, 18° 59′–19° 18′ N, 99° 40′–99° 59′ W), the fourth highest peak in Mexico (4680 m a.s.l.), is in the Trans-Mexican Volcanic Belt, 22 km South-west of the city of Toluca. The Nevado de Toluca protected area has a temperate semi-cold climate with an average annual temperature ranging between 5 and 12 °C. Annual precipitation averages 1200–1800 mm (CONABIO 2000). The main land cover types in the area around the volcano are fir (*Abies religiosa*) and pine (*Pinus hartwegii*, *Pinus montezumae* and *Pinus pseudostrobus*) forests. There are also small patches of broadleaved forests (oak and alder) in the eastern part of the volcano, and alpine grasslands at the highest elevations. The protected

area has an extension of 53,912 ha above the contour line 3000 m a.s.l., covering the municipalities of Toluca, Zinacantepec, Almoloya de Juárez, Amanalco, Temascaltepec, Coatepec Harinas, Villa Guerrero, Villa Victoria, Calimaya and Tenango del Valle. It is considered a priority terrestrial region for biodiversity conservation (CONABIO 2000); 627 plant species as well as 175 vertebrate species and 209 macromycete fungi species (CONANP 2013) were registered in the area. It was declared a National Park in 1936 but at that moment, 54 rural communities had already collective land rights over almost 89% of the total area. Because of the agrarian reforms of the 1910 Mexican Revolution, this situation is common in natural protected areas and generates an ambiguous regulatory status regarding land and forest management (Depraz et al. 2017).

### Forest cover loss assessment

The GFCC product (Hansen et al. 2013) is a widely used open-access map to assess forest cover loss worldwide (Margono et al. 2014; Ceccherini et al. 2020), and in Mexico (Sims and Alix-Garcia 2017; Osorno-Covarrubias et al. 2018). This 30 m resolution product is based on a nearly complete set of imagery from the Landsat missions 4, 5, 7, and 8 in the two last decades. The mapping of deforestation events is achieved following hierarchical decision trees classifiers trained on the compared values of vegetation indices from 1 year to another. A yearly estimate of deforestation (a “Year of gross forest cover loss event”, or “loss year” GIS layer) is provided with pixels encoded as 0 (no loss) or a value ranging from 1 to 19, corresponding to the occurrence of gross deforestation (deforestation regardless of regeneration of the forest in other areas) in year 2001, 2002, ..., 2018, or 2019. We downloaded and analysed the GFCC product in the Nevado de Toluca protected area and for a buffer area outside and within 20 km of the protected area. This was done before and after the change in protection status from National Park to Flora and Fauna Protected Area; that is, from 2001 to 2013 and from 2014 to 2019. An accuracy assessment of the product was performed for the periods 2001–2013 (before the change) and 2014–2019 (after the change). For both periods, 50 pixels of the “loss” class and 500 pixels of the “no loss” class of the product were evaluated against high resolution imagery extracted from the Google Earth archive in (or nearest to) years 2000, 2014 and 2019. Error margins at 95% probability of the deforestation rates were derived from the set of error pixels and match pixels according to the metrics introduced by Stehman and Czaplewski (1998) for the accuracy assessment of land cover maps:

$$S(\hat{A}_k) = A \times S(\hat{p}_k)$$

where  $\hat{A}_k$  is the estimated area of the mapped change class  $k$  (in our case: the forest “loss” class)  $A$  is the total map area,  $\hat{p}_k$  is the estimated area proportion of change class  $k$  derived from the error matrix (see notations in Olofsson et al. 2014: Eqs. (9, 10 and 11), page 52), and  $S$  is the standard error function of the estimators.  $S(\hat{p}_k)$ , the standard deviation of the area proportion of deforestation, is calculated according to the following formula (Cochran, 1977, cited by Olofsson et al. 2014):

$$S(\hat{p}_k) = \sqrt{\sum_i W_i^2 \frac{n_{ik} \left(1 - \frac{n_{ik}}{n_i}\right)}{n_{i-1}}}$$

where  $n_{ik}$  is the sample count at cell  $(i, k)$  in the error matrix and  $W_i$  is the area proportion of map class  $i$ . The error margin of the annual deforestation rate equated the standard



deviation  $S(\hat{A}_k)$  divided by the number of years of the assessed period. Accuracy assessment analysis was carried out with the package ‘mapaccuracy’ (Costa 2021) for R (version 4.0.2; R Development Core Team 2020).

Two Sentinel-2 scenes acquired in two dates (28-11-2016 and 28-12-2017) were downloaded from the Copernicus Open Access Hub (<https://scihub.copernicus.eu/>) for the purpose of visually confirming detected large deforestation spots.

## Forest type classification

In a previous study by Franco-Maass et al. (2006), a forest type map of the Nevado de Toluca National Park was produced for year 2000. This map is vector (polygon) based and derived from aerial photointerpretation of forest stands (“*rodales*”) at 1:37,000 scale. In our research, we aim to build instead a raster-based map of forest types derived from 10 m spatial resolution Sentinel 2 imagery using a machine learning approach. A potential advantage of our approach is that the methodology is less labour intensive and more standardized when repeated over time by forester experts. Our methodology consists in four steps: the acquisition of satellite imagery, the extraction of a “non-forest” layer, the identification of four forest classes in the field, and the application and validation of a Random Forest (RF) classification algorithm.

In a first step, Sentinel-2 multispectral imagery was downloaded from the Copernicus Open Access Hub (<https://scihub.copernicus.eu/>). A single date image (23-12-2018) was downloaded as a level 2A bottom of atmosphere reflectance image without cloud cover, for use for forest type classification. Two SPOT 6 images (same swath) covering the Nevado de Toluca protected area in year 2015 was previously purchased; it was used in this study to differentiate forested from non-forested land as explained in the following paragraph.

The second step consisted in building a “non-forest” vector layer for the study area: the clearly defined agricultural land use along the northern edge and the volcano crater were delineated. In the remaining area, the fused panchromatic—multispectral SPOT 6 images (bands in the blue, green, and red spectral domains) yielded three bands in the visible spectrum at 1.5 m spatial resolution, whose R–G–B colour composite allows to clearly distinguish forest from non-forest in the intricate and fragmented forest landscape of the study area. A Normalized Green Red Difference Index (NGRDI) was extracted from these SPOT 6 fused panchromatic—multispectral bands and the resulting image was classified with an upper threshold of 0.1 to extract a “non-forest” (mostly pastureland) vector layer. All mentioned layers were joined into a “non-forest” vector layer used subsequently as a mask for the classification of forest types with the 10 m resolution Sentinel 2 image.

A total of 4 forest classes were considered: “*Pinus ssp.* forest”, “*Abies religiosa* forest”, “*Cupressus lusitanica* forest”, and “broadleaf forest”. These classes corresponded to relatively distinct tones on colour composites based on B3 (green), B4 (red) and B8 (near infrared) bands of the Sentinel 2 image on forest stands visited in the field and were used in previous classification schemes of the Nevado de Toluca protected area. We did not classify mixed forest types and followed a forest type majority approach. The documentation and georeferentiation of large sites of each of the 4 forest classes was done during a set of field visits in 2018, incorporating a diversity of sites along a gradient of slopes (flat, hilly and sharp mountainous terrains) for the most common forest types (pine and fir forests). The sites were spotted previously on PlanetScope high resolution imagery.

The fourth step was the application and validation of a pixel-based supervised Random Forest (RF) classifier. RF is a machine learning classifier that consists of a combination of

decision tree classifiers. Each decision tree is generated using a random vector and casts a unit vote for the most popular class (Breiman 2001). RF applied to Sentinel-2 imagery was found effective for forest type classification (Immitzer et al. 2016; Persson et al. 2018; Soleimannejad et al. 2019). The number of trees was set to 500 (Rodríguez-Galiano et al. 2012). Seeds for the RF classifier were reflectance values of the Sentinel 2 pixels that corresponded to sites visited in the field after delineation on screen. We used the reflectance values of bands B3 (green), B4 (red) and B8 (near infrared) of Sentinel 2. The number of pixels assigned to a class was approximately proportional to the frequency of this class in the entire protected area according to existing forest type maps. 12,000 pixels were used as the seeds to develop the RF classifier, two third for training and one third for testing. A confusion matrix was computed from the validation pixels to measure the accuracy of the RF classification. The classification algorithm was finally applied to the entire study area using the "non-forest" layer as a mask. The RF classifier was computed in ArcGis Pro 2.3.0.

## Results

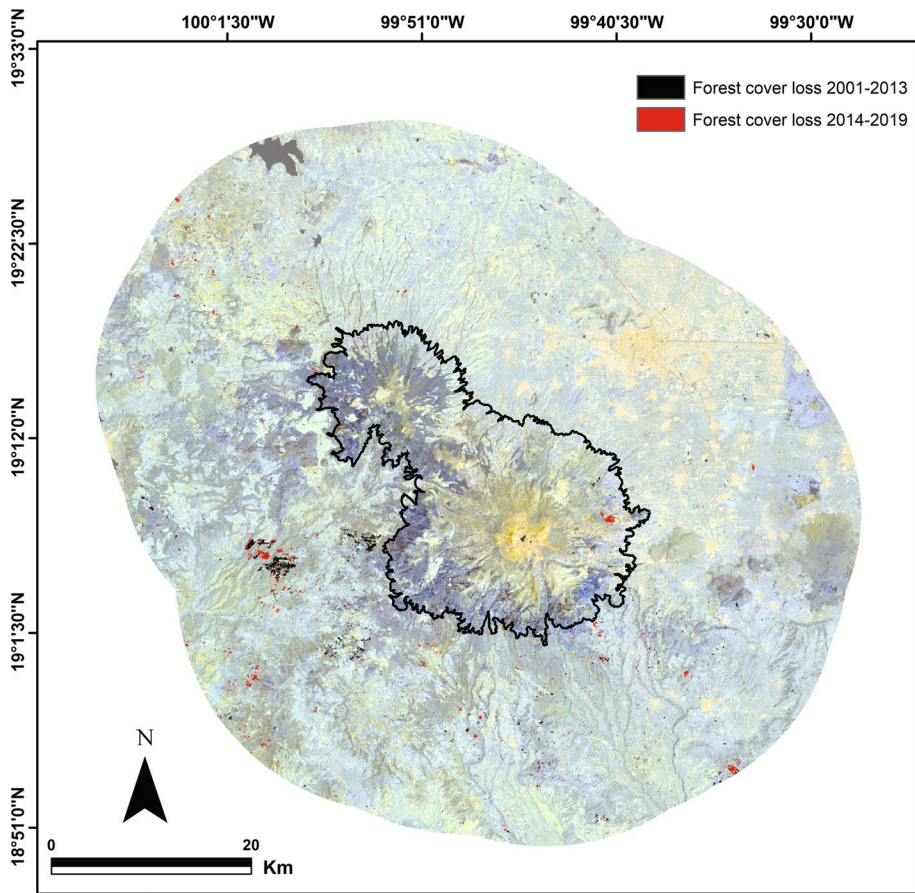
### Forest cover loss in the Nevado de Toluca protected area

According to the GFCC map, the loss in forest cover between 2001 and 2019 represents 0.64% of the total forest in the Nevado de Toluca protected area (Fig. 1). The pattern of gross forest cover loss is somehow regular along the 19-year period except a peak in year 2018 (Fig. 2). The estimated deforestation rates following accuracy assessment (Table 1) are higher;  $39.7 \pm 16.7$  ha/year during 2001–2013 (before the category change) and  $106.0 \pm 40.2$  ha/year during 2014–2019 (after category change).

The increase reported by the GFCC map was partly due to forest cover loss observed in 2018 (66.3 ha of forest was lost: Fig. 2). Most of this deforestation (49.6 ha) occurred in a forest stand in the Calimaya municipality; Red–Green–Blue colour composites at 10 m spatial resolution of Sentinel 2 images were used to illustrate the neighbourhood where this forest cover loss had occurred (Fig. 3); deforestation in this area started in 2017 where a forest patch of 1.6 ha was entirely logged and continued in 2018 when a contiguous patch of 49.6 ha was entirely logged. During a field visit to the deforested site in 2019, we took pictures (Figs. 4 and 5) and observed that *Cupressus lusitanica* was one of the main tree species logged. Forest cover loss in the buffer zone around the Nevado de Toluca protected area also increased after the change in the protection status, especially in 2018 when 394.8 ha were lost (Fig. 6).

### The 2018 forest type map of the Nevado de Toluca protected area

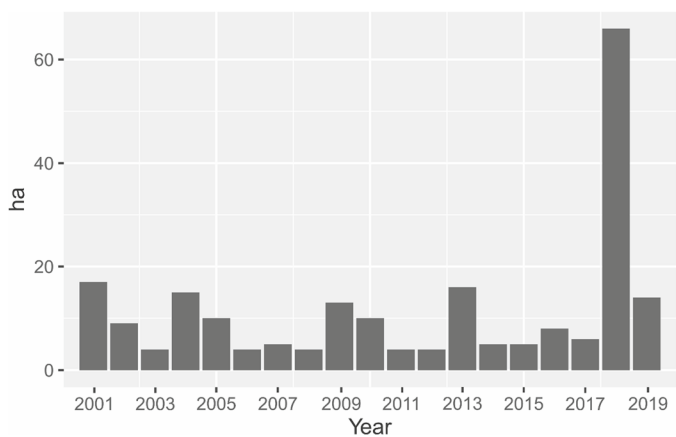
The forest type map of the Nevado de Toluca protected area at 10 m resolution is displayed in Fig. 7. The area mapped as forested land was 34,214.6 hectares. The overall accuracy of the forest type classification was 89.7% (Table 2), showing an acceptable biodiversity mapping performance of the Random Forest classification method in our study area. Probably due to their reduced occurrence in the protected area and limited number of training pixels, *Cupressus lusitanica* forest and broadleaved communities were mapped with less accuracy for the map producer than coniferous communities. *Abies religiosa* forest, the most relevant forest type in terms of conservation and



**Fig. 1** Forest cover loss between 2001 and 2013 (in black) and between 2014 and 2019 (in red) according to the Global Forest Cover Change map in the Nevado de Toluca protected area and within a 20 km-wide buffer zone around it

biodiversity in the Nevado de Toluca protected area, yielded high producer and user accuracies. According to our forest type map, 7171 ha (54.8%) of the *Abies religiosa* forest and 6511 ha (32.5%) of the pine forest of the Nevado de Toluca protected area are inside the forest harvesting area (“sustainable use of natural resources” zonation) (Table 3).

The 10 m resolution forest type map presented in this study is spatially more precise than previous forest type maps in the Nevado de Toluca protected area, developed by Lebreton and Imbernon (2017) and Franco-Maass et al. (2006). The spatial distribution of pine forest pixels in our 2018 map (Fig. 7) compares well with the spatial pattern of the density map elaborated for pine forests in 2016–2017 by Rojas-García et al. (2019). Their map was derived from a high intensity (1.25%) sampling of tree density in the field throughout the entire set of pine forests.

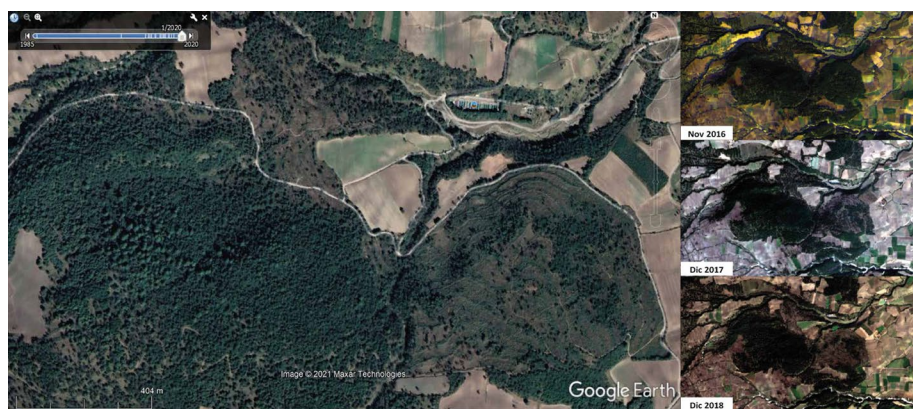


**Fig. 2** Forest cover loss per year (in ha) from 2001 to 2019 in the Nevado de Toluca protected area according to the Global Forest Cover Change map

**Table 1** Error matrix of the forest type classification using Random Forest on Sentinel-2 imagery

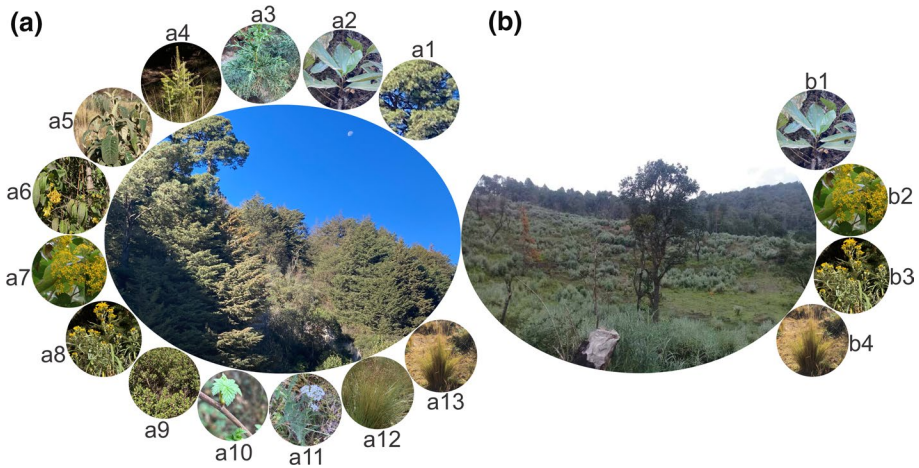
Forest type	<i>Abies religiosa</i>	<i>Pinus</i> sp.	<i>Cupressus lusitanica</i>	Broadleaf	UA [%]
<i>Abies religiosa</i>	1023	80	0	0	92.7
<i>Pinus</i> sp.	190	2330	40	80	88.3
<i>Cupressus lusitanica</i>	0	0	100	0	100.0
Broadleaf	0	20	0	120	85.7
PA [%]	84.3	95.9	71.4	66.7	<b>89.7</b>

The number in bold case indicates the overall classification accuracy of the map. UA indicates user's accuracy and PA producer's accuracy



**Fig. 3** Satellite image coverage of the forested land in the Nevado de Toluca Flora and Fauna Protected Area where clear-cut deforestation occurred in 2017 and 2018. On the right hand side, Red–Green–Blue colour composites at 10 m spatial resolution of Sentinel 2 images from November 2016 to December 2018 (the period when deforestation events occurred). On the left hand side, recent Google Earth image of the area (year 2020)





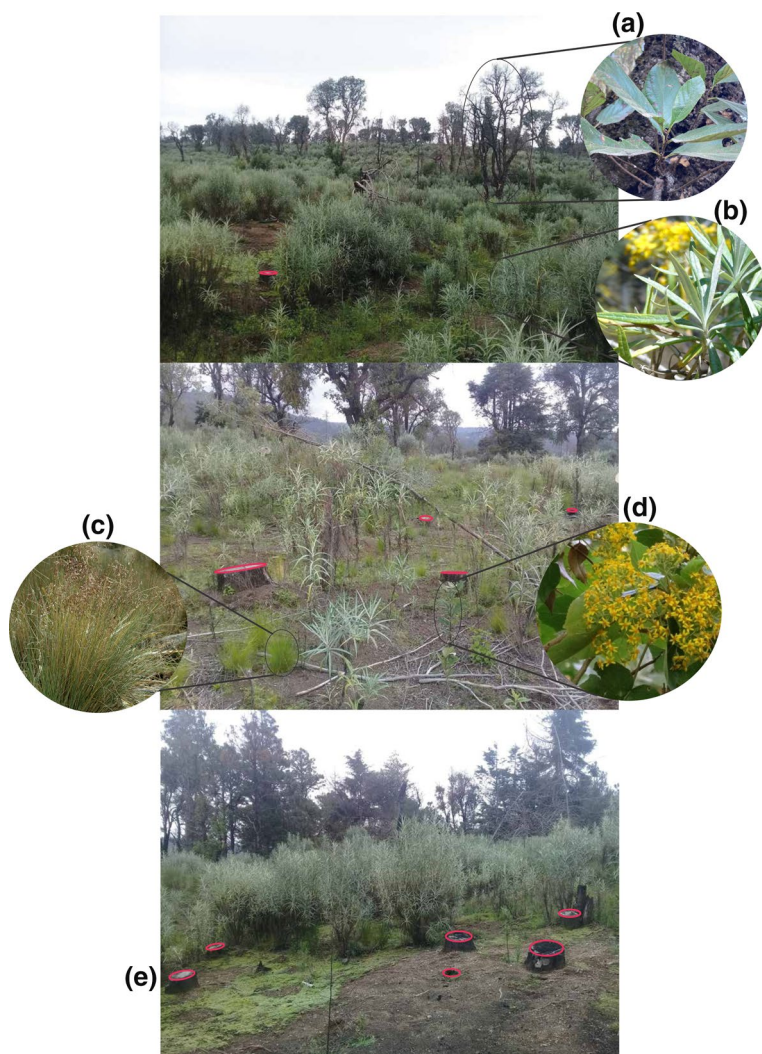
**Fig. 4** Scenery of the forested land (a) and deforested plot in 2018 (b) in the Calimaya municipality. The flora encountered in the forested land includes: (a1) *Pinus* spp., (a2) *Quercus* sp., (a3) *Cupressus lusitanica*, (a4) *Abies religiosa*, (a5) *Buddleja cordata*, (a6) *Cestrum parqui*, (a7) *Roldana angulifolia*, (a8) *Senecio cinerarioides*, (a9) *Baccharis conferta*, (a10) *Alchemilla procumbens*, (a11) *Achillea millefolium*, (a12) *Calamagrostis tolucensis* and (a13) *Agrostis tolucensis*. The flora in the deforested plot includes (b1) *Quercus* sp., (b2) *Roldana angulifolia*, (b3) *Senecio cinerarioides* and (b4) *Agrostis tolucensis*

## Discussion

### Forest cover loss increased inside and outside the protected area

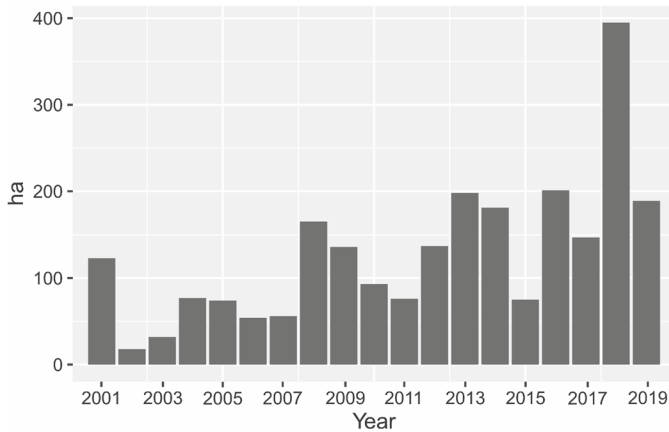
We found a significant increase in forest cover loss after the change of the Nevado de Toluca protection status to “Flora and Fauna Protected Area” ( $39.7 \pm 16.7$  ha/year to  $106.0 \pm 40.2$  ha/year; the confidence intervals do not overlap). This forest cover loss was estimated at 30 m spatial resolution, which is sufficiently precise to include deforestation according to the FAO definition (the removal of contiguous forest patches of minimum 0.5 hectare; FAO 2010). For the first time in the Nevado de Toluca protected area, gross annual deforestation is estimated with confidence intervals (error margins), using international standard estimators (Olofsson et al. 2014; Couturier et al. 2012).

According to the GFCC map, 0.34% of the total forested land was lost between 2001 and 2013 (8.9 ha/year or 0.026% annual forest loss rate). After accuracy assessment, the deforestation area estimator (Olofsson et al. 2014) indicates that between 0.87 and 2.14% of the forested land was lost between 2001 and 2013 (23.0–56.4 ha/year or 0.067–0.16% annual forest loss rate). The GFCC “forest loss” layer is widely reported as underestimating deforestation rates, and the underestimation is larger if deforestation occurs in smaller plots (Hamunyela et al. 2020; Milodowski et al. 2017; Shimizu et al. 2020). Accuracy assessment studies of the GFCC “forest loss” layer on temperate forests report underestimation of GFCC (e.g. Linke et al. 2017; Tortini et al. 2019). The underestimation in our study appears larger than in the above mentioned studies; nevertheless, in contrast with the Nevado de Toluca area, logging practices in temperate forested areas of the USA and Canada occur in the form of clear-cutting of plots mostly larger than 1 ha. The large underestimation found in our study (8.9 ha/year according

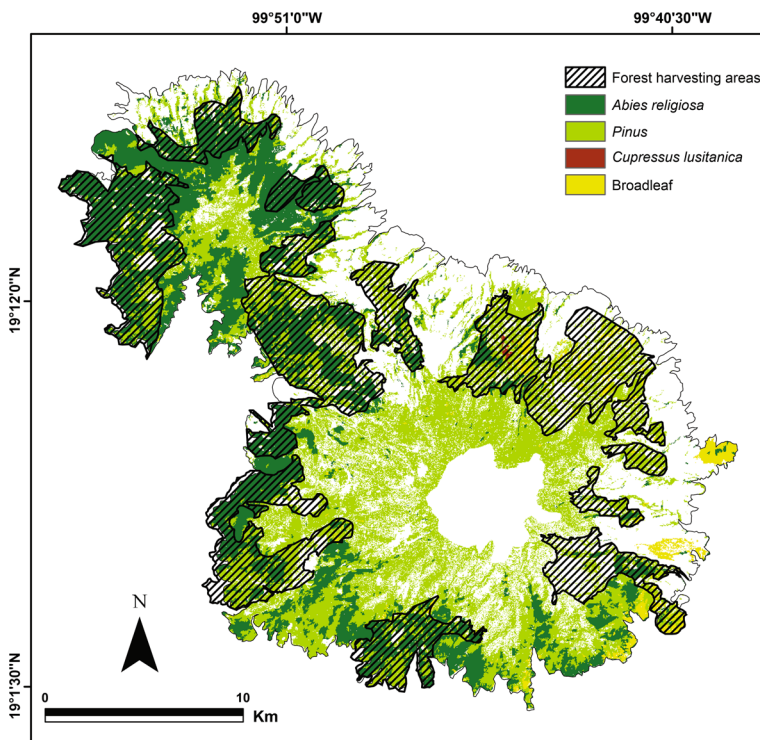


**Fig. 5** Pictures of the clear-cut deforested stand in the Calimaya municipality. The flora in the deforested stand includes: **a** *Quercus* sp., **b** *Senecio cinerarioides*, **c** *Agrostis tolucensis*, and **d** *Roldana angulifolia*. Stumps from clear-cut deforestation **e** are delineated in red colour

to the GFCC “forest loss” value in place of our estimated 23.0–56.4 ha/year during 2001–2013) appears in line with the results of Yamada et al. (2020) based on a study in temperate forests of Japan with disturbance patterns concentrated in very small plots. In their study, only 11% of deforestation plots smaller than 0.89 ha were detected by the GFCC “forest loss” layer (Fig. 8 of Yamada et al. 2020), which meant that the actual deforestation rate in those plots was reported about nine times higher than the GFCC “forest loss” value. In our study, deforestation is almost exclusively due to selective logging, typically detectable in plots of less than 0.5 ha.



**Fig. 6** Forest cover loss per year (in ha) within the 20 km-wide buffer zone around the Nevado de Toluca protected area between 2001 and 2019 according to the Global Forest Cover Change map



**Fig. 7** Forest type map of the Nevado de Toluca protected area at 10 m resolution. Areas where forest harvesting is allowed under the new protection status are represented with stripes. Fir (*Abies religiosa*) forest is shown in dark green, pine (*Pinus spp.*) forest in green, broadleaf in yellow and cedar (*Cupressus lusitanica*) in brown



**Table 2** Area distribution of forest types in the Nevado de Toluca protected area (according to the forest type map elaborated in this study)

Forest type	Total forest area (ha)	Margins of uncertainty (ha)	% inside forest harvesting areas
<i>Abies religiosa</i>	13094.91	2055.90	54.77
<i>Pinus sp</i>	20024.42	821.00	32.52
<i>Cupressus lusitanica</i>	29.33	8.39	100
<i>Broadleaf</i>	1065.98	354.97	42.66

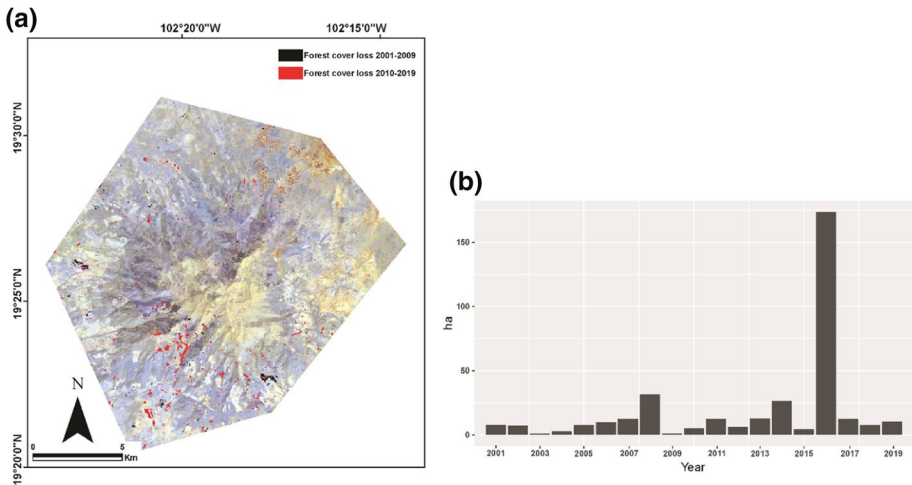
Forest harvesting areas refer to areas under the “Sustainable use of natural resources” zonation within the protected area

**Table 3** Error matrices of the Global Forest Cover Change (GFCC) maps (period 2001–2013 and period 2014–2019) derived from the accuracy assessment carried out in this study. Annual deforestation rates stated in the manuscript were derived from the estimated areas of forest loss in both periods. The GFCC “forest loss” value is much smaller than the estimated value after accuracy assessment (large underestimation of deforestation by GFCC)

2001–2013	Forest loss	No change	Forest gain	Total	User Acc	Mapped area (ha)
Forest loss	35	15	0	50	0.7	115.89
No change	4	492	4	500	0.98	53870.31
Total	39	507	4	550		53986.20
Prod Acc	0.90	0.97				
Estimated area (ha)	515.56	53470.64				
Confidence interval (ha)	216.69	216.69				
Overall Acc					0.93	
2014–2019	Forest loss	No change	Forest gain	Total	User Acc	Mapped area (ha)
Forest loss	46	4	0	50	0.92	104.30
No change	5	494	1	500	0.99	53881.91
Total	51	498	1	550		53986.20
Prod Acc	0.90	0.99				
Estimated area (ha)	635.85	53350.35				
Confidence interval (ha)	240.51	240.51				
Overall Acc					0.98	

Acc Accuracy; Prod Producer

The size of the validation sites in our study were 30×30 m (0.09 ha). A potential limitation of our verification protocol is the geometric location of our 30×30 m verification sites on the original Landsat imagery (reference imagery of the GFCC product): In mountainous settings, geometric errors related to differentiated georeferencing of the Landsat imagery and of the high-resolution imagery may occur and affect the accuracy assessment process. This difficulty is hard to overcome because conspicuous reference features (to correct for geometric inconsistencies) in homogenous forest contexts are sometimes difficult to spot on the 30 m (too coarse) resolution Landsat imagery. Annual forest loss maps at 10 m resolution (near to crown scale) derived from similar algorithms than GFCC applied



**Fig. 8** **a** Forest cover loss between 2001 and 2009 (in black) and between 2010 and 2019 (in red) and **b** forest cover loss per year (in ha) from 2001 to 2019 according to the Global Forest Cover Change map in the Pico de Tancitaro protected area which was downgraded from National Park to “Flora and Fauna Protected Area” in 2009

to Sentinel-2 imagery (for example) instead of Landsat imagery should be more adapted to the estimation of small plot deforestation (forest degradation) typical of selective logging practices in Mexico.

Any of the deforestation rates mentioned above for the 2001–2013 period is much lower than 30 m resolution deforestation hotspots at the national level (CONAFOR 2020). This finding is contrary to the premise of “widespread deterioration of such a magnitude that [the protected area] was expected to completely lose its function in the northern part within a decade or two”, according to the preliminary justification documents for recategorization (CONANP 2013, cited by Mastretta Yanes et al. 2014). High deforestation rates (over 150 ha/year) were derived between 1972 and 2000 by Franco-Maass et al. (2006), using vector-based forest density maps and a linguistic classification scale (including “fragmented”, “semi-dense”, “dense”) as a guide for aerial photointerpretation. Although this type of map is very valuable for landscape planning or forest inventory sampling at one date, it has been documented that the delineation of density classes is highly interpret-dependent and the comparison over time of density maps generated by unrelated interpreters yields too much uncertainty for forest monitoring and assessment in Mexico (Couturier et al. 2010, 2009) and elsewhere (Rommel et al. 2005). Moreover, the 1972–2000 deforestation rate (without error margins) was calculated with the interpretation that the 2000 “fragmented” class was a non-forest class (Franco-Maass et al. 2006, p. 26), which is probably a major source of uncertainty in their estimate (Mastretta-Yanes et al. 2014).

Another premise about strict protection status is that deforestation activities are displaced to other areas in the surroundings (Meyfroidt and Lambin 2009), and that downgrading the protection status should alleviate the pressure on surrounding resources. In our case, on the contrary, forest cover loss has increased in the surroundings as well (with most deforestation on year 2018), which means the logging activity within the Flora and Fauna Protected Area did not seem to release the pressure on resources outside either. The number of permits for “sustainable use of the natural resources” dramatically increased starting

2017 (<https://datos.gob.mx/busca/dataset/autorizaciones-y-aprovechamientos-forestales>). These data suggest that the announcement of the less restrictive new status “Flora and Fauna Protected Area” may have attracted instead more local and external actors involved in logging activities. The cumulative effect of this appeal could represent an excessive threat to conservation.

The deforestation rate found during the 13 years of the National Park status ( $39.7 \pm 16.7$  ha/year) is higher than the gross deforestation rate of 15.9 ha/year estimated by Lebreton and Imbernon (2017) based on the delineation of aerial photographs for 2000–2014. By contrast, the deforestation rate after the change of protection status was found to be  $106.0 \pm 40.2$  ha/year, significantly higher than the deforestation rate before the change. We conclude that after five years of operation, the new protection status is likely to have failed its alleged purpose of reducing deforestation.

The decrease in forest cover is partly due to the deforestation occurred during year 2018, prevailing a clear-cut deforestation event of a large forest stand (51.2 hectares) in the Calimaya municipality and neighbourhoods. Only selective logging (no clear-cut deforestation) is allowed in the forest harvesting areas according to the management plan of the Flora and Fauna Protected Area (DOF 2016). No such clear-cut deforestation had occurred in the period 2001–2013 under the National Park status, which means the former status ensured overall forest stand persistence, whereas the new status has not ensured overall forest stand persistence. A hypothesis may hold that the Calimaya deforestation event is exceptional and does not reflect trends of harvesting practices in Flora and Fauna Protected Areas. In view of testing this hypothesis, we observed remote sensing data on the Pico de Tancítaro protected area in the neighbouring state of Michoacán, which was also downgraded from National Park to Flora and Fauna Protected Area in 2009. The GFCC product in the Pico de Tancítaro revealed a sudden forest cover loss increase to 174 hectares during 2016 (a set of several entirely logged forest stands), when the forest cover loss average between 2001 and 2019 (excluding 2016) is on average 9.9 ha/year (Fig. 8). This observation tends to favor the counter hypothesis of non-persistent forest stands and excessive harvesting practices under the Flora and Fauna Protected Area status, in contradiction with the ecological integrity of forests. In response to questions about the deforestation event in Calimaya, the CONANP authorities affirm that a harvesting authorization was emitted for this forest stand, on the grounds that *Cupressus lusitanica* is not native of the protected area. The only authorization that appears in the public dataset for this municipality, however, is for a later time period (<https://datos.gob.mx/busca/dataset/autorizaciones-y-aprovechamientos-forestales>). In any case, the use of permits should follow principles of a “sustainable use of natural resources” by which it is doubtful that a clear-cut of a 51.2 hectares forest stand is ecologically benign and compatible with the biodiversity conservation priority objective of national protected areas.

### Implications of the results for the 2013 justification of the new protection category

In Mexican environmental regulations, the modification of a protection decree is allowed if the objectives of the decree are impossible to sustain (article 62 of the LGEEPA regulations on natural protected areas, LGEEPA 2020a). The affirmation of “high deforestation rates that were compromising the integrity of the ecosystems and species of the Nevado de Toluca Volcano region” in the justification study (CONANP 2013, cited by Mastretta Yanes et al. 2014) was a major reason invoked for the protection status change in 2013.

The gross yearly deforestation rate of 0.067–0.16% annual forest loss rate found between 2001 and 2013 (before the change of the protection status) invalidates this major reason. Gross deforestation may be partly compensated by regeneration in other areas, which was not assessed in our study; Lebreton and Imbernon (2017) report a net gain (although without error margins) of forested land between 2000 and 2014 using a spatially detailed forest—non forest photointerpretation approach. In line with recommendations in Couturier et al. (2012), this study should stand as a baseline for deforestation estimation with error margins in the Nevado de Toluca protected area. In addition to the deforestation rate criterion, we further propose the indicator of forest stand persistence as a key criterion for environmental policy assessments in protected areas in Mexico, with the purpose of ensuring the ecological integrity of forest patches. According to our data, the “Flora and Fauna Protected Area” management plan failed in ensuring forest stand persistence in 2017 and 2018 while the “National Park” management plan did not fail throughout 2001–2013.

Another reason invoked for the category change in 2013 is the widespread practice of selective logging in the protected area and the subsequent widespread degradation of the forest. The evidence of selective logging and the density of tree removal were established on a total of about 50 forest plots in four sites of the southern part of the Nevado de Toluca National Park (Endara-Agramont et al. 2012, 2013). The use of forest resources including timber has indeed been a historical revindication of local *ejidos* and communities in the National Park and caused debates because of the contradiction with the strict protection status of the National Park (Mastretta-Yanes et al. 2014). This is a problem which needs to be tackled when deciding on a protection status, as discussed in the next subsections. Nevertheless, the justification for the decision on the new category in 2013 was related to the unsustainable character of selective logging at that time. This unsustainable character cannot be deduced from the one date sampling (date which is not mentioned in the publications) on which both studies are based, nor from the forest density study by Franco-Maass et al. (2006) because of major uncertainties in the comparison of photo interpreted density data, as discussed in the previous subsection. In view of the above, we agree with Lebreton and Imbernon (2017) that the recategorization decision in 2013 was not based on solid, reproduceable, scientific grounds.

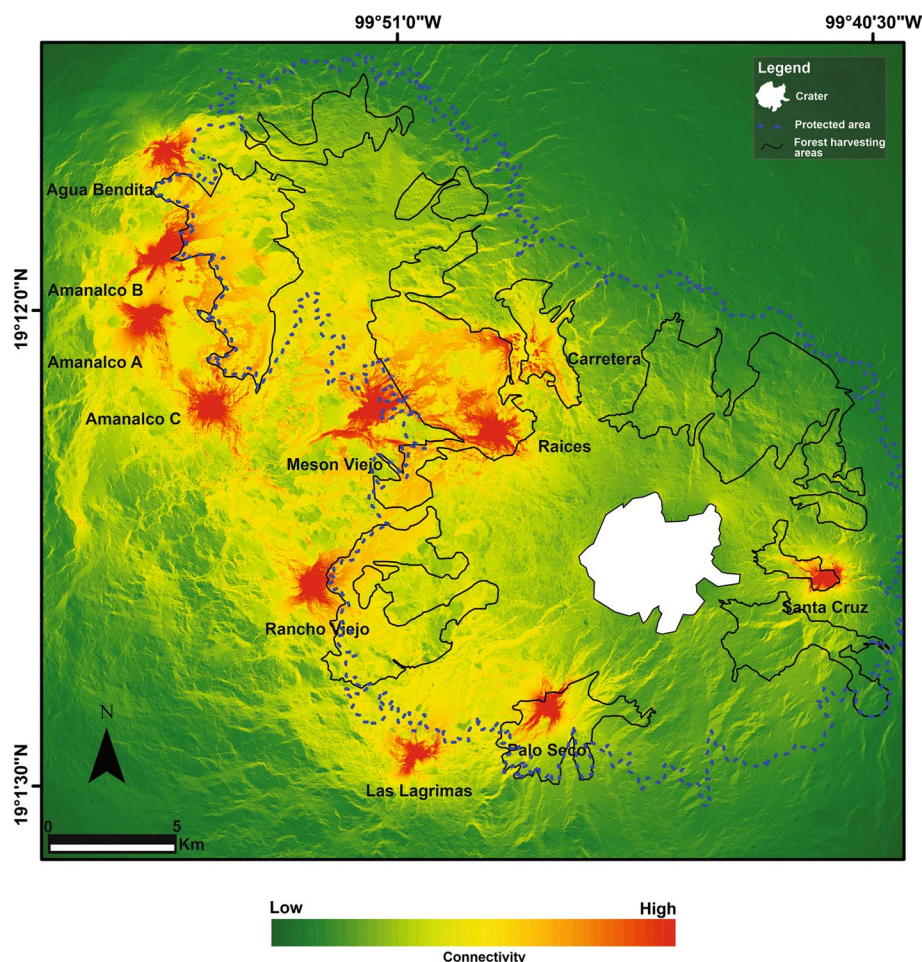
### Potential threats of the 2013 new category to the Nevado de Toluca protected area

As devised earlier, the announcement of the less restrictive new status “Flora and Fauna Protected Area” may have attracted more local and external actors involved in logging activities in the past few years. According to the Google Earth image archive, at least two new sawmills were installed inside the protected area after 2014. Under the Flora and Fauna Protection Area status, more flexibility is allowed to external investors with a high financial and commercial capacity: private lands may be purchased, and land tenure may shift from ejidal to private (Depraz et al. 2017).

Only selective logging (no clear-cut deforestation) is allowed in the forest harvesting areas according to the management plan of the Flora and Fauna Protected Area (DOF 2016). In the context of the clear-cut deforestation events that have happened, the management plan is not explicit on how, and with what qualified human resource, the sustainability condition will be enforced in the forest harvesting areas (Mastretta Yanes et al. 2014).

According to our 2018 map, 7171 ha (54.8%) of the *Abies religiosa* forest (one of the most extended *Abies* patches in Mexico and of upmost biodiversity importance in the region) and 6,511 ha (32.5%) of the pine forest of the Nevado de Toluca protected area are

inside the forest harvesting area (“sustainable use of natural resources” zonation) (Fig. 7; Table 3). Regarding ecosystem function, services or biodiversity, unsustainable selective logging is not as detrimental as deforestation. Nevertheless, selective logging extending deeper into the core of remaining intact forest areas, causes widespread collateral damage to the remaining trees (Asner et al. 2005). In this sense, it could threaten forest sustainability in a greater way than bold deforestation (Broadbent et al. 2008). At least, selective logging adds stress to the already fragile state of *Abies religiosa* forest due to climate change (Sáenz-Romero et al. 2012). Selective logging can also be detrimental to forest specialist fauna species, like the *Pseudoeurycea robertsi* salamander that depends on the availability of core forest areas in the landscape (González-Fernández et al. 2019). Some *Pseudoeurycea robertsi* populations and important connectivity areas between these populations overlap with forest harvesting areas. The areas with higher connectivity include the



**Fig. 9** Connectivity map of the *P. robertsi* populations in the Nevado de Toluca protected area (modified from Sunny et al. 2021a, b). High connectivity corridors between 11 populations of *P. robertsi* can be seen in orange and red tones. Forest harvesting areas were delineated in black in order to appreciate the potential threat to the biodiversity of the forests in the Flora and Fauna Protected Area



north-western (Amanalco) and central part (Raíces and Mesón Viejo) of the Nevado de Toluca protected area as well as Rancho Viejo (Fig. 9). For example, the deforestation of Mesón Viejo would break most biological corridors, affecting the connectivity of the entire landscape (Sunny et al. 2021a, b).

Finally, the Nevado de Toluca protected area is part of the strategic set of protected areas to the South of the Central Valleys in the Mexican altiplano, including the Cumbres del Ajusco National Park, the Iztacchiuatl-Popocatepetl National Park and the La Malinche National Park. The forests of these protected areas play several key environmental roles to preserve the livelihood of the most populated region of the country, including the mitigation of climate change, the recharge of vital aquifers for the agricultural valleys and for the population (López-García and Navarro-Cerrillo, 2021; López-García, 2019). Of particular importance to the food security of the altiplano cities is the fertile Toluca—Atla-comulco valley, North to the Nevado de Toluca, which depends on the moist climate and the quality of aquifers. Considering the above, a precautionary principle should be held to avoid the risks of maintaining such strategical forested land under a low protection status. Beyond the observed increase in deforestation shown in this study, cumulative effects of the announcement of the less restrictive protection area may provoke new clear-cut deforestation events in the near future.

### **Suggestions for the environmental public policy in the Nevado de Toluca protected area**

Two suggestions are emitted. First, selective logging should be monitored regardless of the protection status; a substantial part of the CONANP management plan should focus on the conditions, resources, and governance strategies for enforcing this monitoring task. The study by Rojas-García et al. (2019) sets a reproducible methodology, based on intense sampling, to monitor tree and stump density over time. The methodology could be complemented by a probabilistic estimator of area change with error margins (Olofsson et al. 2014; Couturier et al. 2012). In parallel, our study sets a reproducible methodology to map forested land at 10 m resolution, which could allow for forest density monitoring. The mentioned independent self-consistent datasets could constitute the core of scientifically ground forest density monitoring, additional to the deforestation monitoring at 30 m resolution, with the possible effect of reducing the likeliness of controversial decisions.

The other suggestion is to reconsider the Flora and Fauna Protection Area status. According to Mexican environmental legislation, the Biosphere Reserve category allows the sustainable use of resources, but only by local communities (article 48: LGEEPA 2020b), unlike the Flora and Fauna Protected Area category which allows resource extraction by external actors (article 54: LGEEPA 2020b). A change to a Biosphere Reserve would be more consistent with the ecological importance of the Nevado de Toluca protected area within the system of strictly protected areas in the Central Valleys, in terms of ecosystem services, biodiversity, endemism and emblematic and threatened species (Heredia-Bobadilla et al. 2016). Among compensating financing sources for the management plan and forest conservation, eco-tourism options (Franco-Maass et al. 2006) could be extended to biodiversity watching and other low impact activities, and financial returns from hydrological benefits by the industry and agriculture in the valley could be considered. Payment for hydrological services from the citizenship of the neighbouring city of Toluca has already been explored (Brunett et al. 2010). The Toluca Valley aquifer is already overexploited,

with problems like groundwater table decline, subsidence and drying of springs and ponds, affecting Toluca and Mexico City's water supply (Esteller and Diaz-Delgado 2002).

The evidence shown in this study tends to confirm previous analysis that the “Flora and Fauna Protected Area” status may not be in favour of the environment and local communities. Toscana-Aparicio and Granados-Ramírez (2015) mentioned that the rural communities in Loma Alta and Raíces are being more oppressed by the new regulation than before. Depraz et al. (2017) conclude that the change of protection category reflects in fact public policies for more access to natural resources in the central megalopolis region. With the current protection category, this access and exploitation of the forest cover may occur through large scale commercial activities beyond the harvesting capacities of local communities.

## Conclusions

In this study, deforestation trends at 30 m resolution in the Nevado de Toluca protected area are derived from the application of an international standard area estimator to the GFCC map (the “forest loss” layer). In the justification study for the recategorization of the Nevado de Toluca National Park in 2013, two scientific claims were based on the forest cover situation: First, the trend of illegal selective logging would lead to the complete functional loss of some forested ecosystems in one to two decades. Second, the excessive illegal logging is associated to the highly restrictive protection status of the protected area, and a legal scheme with sustainable forest use zonation would lower selective logging in the protected area.

According to our results, the forest loss for the period 2001–2013 under the National Park protection status was  $39.7 \pm 16.7$  ha/year for a total forested land of 34,214 ha (between 0.067% and 0.160% annual forest loss rate). Additionally, large clear-cut deforestation did not occur. While these results confirm the existence of selective logging, the forest loss rate was clearly insufficient for the complete functional loss of forested ecosystems, casting doubts on the validity of the first hypothesis.

Furthermore, the forest loss during the following 2014–2019 period, under the Flora and Fauna Protected Area status, was found  $106.0 \pm 40.2$  ha/year, significantly higher than under the former protection status. Additionally, one large clear-cut deforestation of a 51.2 ha forest stand occurred in 2017–2018. These results are clearly not favourable to hypothesis 2 and indicate, on the contrary, that forest stand persistence is not respected in the management plan of the Flora and Fauna Protected Area, and that selective logging is more intense than under the National Park protection status. Since forest loss also increased in a surrounding 20 km buffer zone, a likely hypothesis is that more numerous logging actors inside and outside the communities of the Nevado de Toluca have been attracted by the economic opportunities of timber production following the announcement of the recategorization to a less restrictive protection, that the volume of timber harvested per year is unsustainable within the protected area and that this poses a much greater threat than before on the biodiversity and ecological integrity of the forests.

In this research, a new forest type map was elaborated for the Nevado de Toluca protected area. This 2018 map, at 10 m resolution, is spatially more detailed than previous maps and could allow for forest cover density monitoring better suited to the needs of monitoring fine scale selective logging in the protected area. It is derived from a machine learning classification technique applied to Sentinel 2 imagery, which can be



easily reproduced for the purpose of rapid assessment of environmental policies in the protected area. According to this map, more than half of the *Abies religiosa* forests are included in forest harvesting areas. Because of the exceptional biodiversity importance of the *Abies* forest (a major continuous patch in the country) and the strategic value of the protected area for the aquifers in the central valleys, it is of special concern what may happen in the following years. An assessment on the substitution of the *Abies* forests by other forest covers should be conducted in future studies. Indeed, the reduction of *Abies* forests and substitution by pine forests was reported in Sierra Nevada (Montoya et al. 2020).

We suggest the Nevado de Toluca protected area should return to a higher protection category. For example, the Biosphere Reserve category allows the forest use by local communities but avoids the establishment of private sawmills of external actors within the protected area. The conditions for the implementation of selective logging monitoring should be a key part of the forest management plan of the Nevado de Toluca protected area in the future. Both criteria of forest cover loss and forest stand persistence should be considered to assess this management plan. Considering the above, we make the following recommendations for the protected area:

- (1) Forest cover loss and forest stand persistence should be monitored and publicly reported every year;
- (2) Forest density should be monitored every 5 years using high intensity sampling (e.g. Rojas-García et al. 2019), and using 10 m resolution forest maps derived from a Random Forest classifier, as an independent assessment; Forests density should be publicly reported every 5 years;
- (3) The “Flora and Fauna Protection Area” category should be reconsidered, and the Biosphere Reserve category should be considered as a candidate for a new protection status.

A similar recategorization scheme has been proposed by environmental national authorities and in the international debate as a model to apply to other National Parks (e.g., the Popocatepetl-Iztaccihuatl National Park in Mexico; Ceballos 2011). Our study sheds lights on the risks this policy option could represent for forest conservation at the national level.

Studies have documented a widespread trend, since 2000, for Protected Area Downgrading, Downsizing, and Degazettement (PADDD) (Mascia and Pailler 2011; Qin et al. 2019); between 1892 and 2018, 73 countries enacted 3749 PADDD events (2705 downgrades, 698 downsizes, and 346 degazettements), removing 519,857 km<sup>2</sup> from protection and tempering regulations in an additional 1,659,972 km<sup>2</sup> (Golden Kroner et al. 2019). The methods and recommendations presented here could hold for a wide range of protected areas around the world affected by this trend.

**Acknowledgements** We are grateful to the editor and two anonymous reviewers for their very constructive and useful comments. We are grateful to Daniel Carrillo for taking us to the deforested area. A.G.F. is on her post-doctoral stay at Universidad Nacional Autónoma de México (UNAM), receiving a scholarship from UNAM (CJIC/CTIC/5052/2021). During a previous stage of this research, she also received a post-doctoral scholarship by Universidad Autónoma de México (UAM), campus Lerma (PODEP: 511-6/2019-15932). A.S received financial support from the Secretary of Research and Advanced Studies (SYEA) of the Autonomous University of the State of Mexico (Grant: 4732/2019CIB). J.S. is recipient of a FPI doctoral fellowship from Ministerio de Ciencia e Innovación (Spain)(Grant PRE2020-091907). This research was elaborated under five academic projects: UNAM funded project PAPIIT IN304722 “De la historia ambiental a la defensa territorial: Contribución a la conservación y restauración de la biodiversidad forestal en México mediante Tecnologías de la Información Geográfica”; CONACYT funded project A1-S-34633 “De

la caracterización química y molecular al aprovechamiento sustentable de especies silvestres de *Lupinus*"; CONACYT funded project LN-CONACYT-2021-315858 "Laboratorio Nacional de Observación de la Tierra (LANOT) 2021"; Instituto de Geografía UNAM supported project "Estimación robusta de tasas de deforestación en apoyo al Sistema Satelital de Monitoreo Forestal (SaMoF)", and Universitat Autònoma de Barcelona (UAB) supported project "Instituciones, actores locales y TIG libres en problemáticas forestales: hacia una mayor justicia ambiental en México".

**Funding** This research received public funding through the following institutions: DGAPA UNAM (Project PAPIIT IN304722); CONACYT (Projects A1-S-34633 and LN-CONACYT-2021–315858), Secretary of Research and Advanced Studies (SYEA) in the Autonomous University of the State of Mexico (Grant 4732/2019CIB); Universidad Autònoma de Barcelona (UAB); Instituto de Geografía (UNAM).

**Data availability** The primary data underlying this research are from free access space agency repositories (Sentinel-2 and Landsat imagery), from the French spatial agency (SPOT-6 image) and from the free access Global Forest Watch website.

**Code availability** QGIS was used as the Geographic Information System for data analysis. ArcGIS Pro 2.3.0 was used as the Remote Sensing platform for forest type classification.

## Declarations

**Conflicts of interest** The authors declare no known conflicts of interest/competing interests with respect to the process and results of this research.

## References


- Asner GP, Knapp DE, Broadbent EN, Oliveira PJ, Keller M, Silva JN (2005) Selective logging in the Brazilian Amazon. *Science* 310:480–482. <https://doi.org/10.1126/science.1118051>
- Bille T (2009) Field observations on the salamanders (Caudata: Ambystomatidae, Plethodontidae) of Nevado de Toluca, Mexico. *Salamandra* 45:155–164
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
- Broadbent EN, Asner GP, Keller M, Knapp DE, Oliveira PJ, Silva JN (2008) Forest fragmentation and edge effects from deforestation and selective logging in the Brazilian Amazon. *Biol Conserv* 141:1745–1757. <https://doi.org/10.1016/j.biocon.2008.04.024>
- Brunett E, Baró JE, Cadena E, Esteller MV (2010) Pago por servicios ambientales hidrológicos: caso de estudio Parque Nacional del Nevado de Toluca, México. *CIENCIA Ergo-Sum* 17(3):286–294
- Ceballos G (ed) (2011) Propuesta de recategorización y de decreto del parque nacional Nevado de Toluca, Toluca, SEMARNAT/Government of the State of México
- Ceccherini G, Duveiller G, Grassi G, Lemoine G, Avitabile V, Pilli R, Cescatti A (2020) Abrupt increase in harvested forest area over Europe after 2015. *Nature* 583(7814):72–77. <https://doi.org/10.1038/s41586-020-2438-y>
- Cochran WG (1977) Sampling techniques, 3rd edn. Wiley, New York
- CONABIO, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (2000) Regionalización. [http://www.conabio.gob.mx/conocimiento/regionalizacion/doctos/rtp\\_109.pdf](http://www.conabio.gob.mx/conocimiento/regionalizacion/doctos/rtp_109.pdf). Accessed 5 Nov 2018
- CONAFOR (2020) Estimación de la tasa de deforestación mediante el enfoque de muestreo. Work report 14th October 2020. Comisión Nacional Forestal. <https://www.gob.mx/conafor/documentos/estimacion-de-la-tasa-de-deforestacion-bruta-en-mexico-para-el-periodo-2001-2018-mediante-el-metodo-de-muestreo?idiom=es>. Accessed 18 Aug 2021.
- CONANP (Comisión Nacional de Áreas Naturales Protegidas) (2013) Estudio previo justificativo para la modificación de la declaratoria del Parque Nacional Nevado de Toluca. Accessed 20 Jan 2019.
- CONANP (Comisión Nacional de Áreas Naturales Protegidas) (2016) Programa de Manejo Área de Protección de Flora y Fauna Nevado de Toluca. [https://simec.conanp.gob.mx/pdf\\_libro\\_pm/104\\_libro\\_pm.pdf](https://simec.conanp.gob.mx/pdf_libro_pm/104_libro_pm.pdf)
- Costa H (2021) Unbiased Thematic Map Accuracy and Area. R package version 0.1.0.
- Couturier S (2019) The global scar on Congo forests. *Nat Sustain* 2:547–548. <https://doi.org/10.1038/s41893-019-0315-1>

- Couturier S, Mas JF, Cuevas G, Benítez J, Vega-Guzmán A, Coria-Tapia V (2009) An accuracy index with positional and thematic fuzzy bounds for land-use/ land-cover maps. *Photogr Eng Rem Sens* 75(7):789–805. <https://doi.org/10.14358/PERS.75.7.789>
- Couturier S, Mas JF, López-Granados E, Benítez J, Coria-Tapia V, Vega-Guzmán A (2010) Accuracy assessment of the Mexican National Forest Inventory map: a study in four ecogeographical areas. *Singapore J Trop Geogr* 31(2):163–179. <https://doi.org/10.1111/j.1467-9493.2010.00399.x>
- Couturier S, Núñez JM, Kolb M (2012) Measuring tropical deforestation with error margins: a method for REDD monitoring in South-Eastern Mexico. In: Sudarshana P, Nageswara-Rao M, Soneji JR (eds) *Tropical forests*. Intech Open Access Publishing, London, pp 269–296
- Depraz S, Sanial E, Catalán AKR, Rojas AS (2017) Less protection for better conservation? A politicised relationship between a city and its protected area in the vicinity of Nevado de Toluca (Mexico). *J Urban Res*. <https://doi.org/10.4000/articulo.3261>
- DOF, Diario Oficial de la Federación Mexicana (2013) Decreto que reforma, deroga y adiciona diversas disposiciones del diverso publicado el 25 de enero de 1936, por el que se declaró Parque Nacional la montaña denominada “Nevado de Toluca” que fue modificado por el diverso publicado el 19 de febrero de 1937. [http://dof.gob.mx/nota\\_detalle\\_popup.php?codigo=5315889](http://dof.gob.mx/nota_detalle_popup.php?codigo=5315889). Accessed 5 Nov 2018.
- DOF, Diario Oficial de la Federación Mexicana (2016) Acuerdo por el que se da a conocer el Resumen del Programa de Manejo del Área Natural Protegida con categoría de Área de Protección de Flora y Fauna Nevado de Toluca. Diario Oficial de la Federación del 21 de octubre de 2016. [http://dof.gob.mx/nota\\_detalle.php?codigo=5457780&fecha=21/10/2016](http://dof.gob.mx/nota_detalle.php?codigo=5457780&fecha=21/10/2016). Accessed 5 Nov 2018
- Endara-Agramont AR, Franco-Maass S, Nava-Bernal G, Hernández JI, Fredericksen TS (2012) Effect of human disturbance on the structure and regeneration of forests in the Nevado de Toluca National Park, Mexico. *J for Res* 23(1):39–44. <https://doi.org/10.1007/s11676-012-0226-8>
- Endara-Agramont AR, Calderón-Contreras R, Nava-Bernal G, Franco-Maass S (2013) Analysis of fragmentation processes in high-mountain forests of the centre of Mexico. *Am J Plant Sci* 4:697–704
- Esteller MV, Díaz-Delgado C (2002) Environmental effects of aquifer overexploitation: a case study in the Highlands of Mexico. *Environ Manage* 29(2):266–278. <https://doi.org/10.1007/s00267-001-0024-0>
- FAO (2010) *Global Forest Resources Assessment 2010*. Main Report, Forestry paper no 163, Rome, p 340
- Franco-Maass S, Regil-García HH, González-Esquivel C, Nava-Bernal G (2006) Cambio de uso del suelo y vegetación en el Parque Nacional Nevado de Toluca, México, en el periodo 1972–2000. *Investig Geogr* 61:38–57
- Golden Kroner RE, Qin S, Cook CN, Krithivasan R, Pack SM, Bonilla OD, Cort-Kansinall KA, Coutinho B, Feng M, García MIM, He Y (2019) The uncertain future of protected lands and waters. *Science* 364(6443):881–886
- González-Fernández A, Arroyo-Rodríguez V, Ramírez-Corona F, Manjarrez J, Aguilera-Hernández A, Sunny A (2019) Local and landscape drivers of the number of individuals and genetic diversity of a microendemic and critically endangered salamander. *Land Ecol* 34:1989–2000. <https://doi.org/10.1007/s10980-019-00871-2>
- Hamunyela E, Brandt P, Shirima D, Thanh DHT, Herold M, Roman-Cuesta RM (2020) Space-time detection of deforestation, forest degradation and regeneration in montane forests of Eastern Tanzania. *Int J Appl Earth Obs Geoinf* 88:102063
- Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SAA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, Kommareddy A, Egorov A, Chini L, Justice CO, Townshend JRG (2013) High-resolution global maps of 21st-century forest cover change. *Science* 342:850–853
- Heredia-Bobadilla RL, Monroy-Vilchis O, Zarco-González MM, Martínez-Gómez D, Mendoza-Martínez GD, Sunny A (2016) Genetic structure and diversity in an isolated population of an endemic mole salamander (*Ambystoma rivulare* Taylor, 1940) of central Mexico. *Genetica* 144:689–698. <https://doi.org/10.1007/s10709-016-9935-9>
- Héritier S, Lebreton C (2017) From ‘revolutionary’ to contested park. Mobilization and conflicts in the recategorization process of the Nevado de Toluca National Park (Mexico). *J Urban Res*. <https://doi.org/10.4000/articulo.3279>
- Immitzer M, Vuolo F, Atzberger C (2016) First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens* 8(3):166. <https://doi.org/10.3390/rs8030166>
- IUCN SSC Amphibian Specialist Group (2016) *Pseudoeurycea robertsi*. The IUCN Red List of Threatened Species. <https://www.iucnredlist.org/species/59393/53983925>. Accessed 12 Feb 2018
- Lebreton C, Imbernon J (2017) Controversia científica y democrática: el caso de la recategorización del Parque Nacional del Nevado de Toluca, México. *Soc Ambient* 14:5–30
- LGEEPA (2020a) Artículo 62 del Reglamento de la Ley General del Equilibrio Ecológico y la Protección al Ambiente en materia de áreas naturales protegidas
- LGEEPA (2020b) Artículos 48, 54 de la Ley General del Equilibrio Ecológico y la Protección al Ambiente

- Linke J, Fortin MJ, Courtenay S, Cormier R (2017) High-resolution global maps of 21st-century annual forest loss: Independent accuracy assessment and application in a temperate forest region of Atlantic Canada. *Remote Sens Environ* 188:164–176
- Loarie SR, Duffy PB, Hamilton H, Asner GP, Field CB, Ackerly DD (2009) The velocity of climate change. *Nature* 462(7276):1052–1055. <https://doi.org/10.1038/nature08649>
- López-García J (2019) Changes in forest cover in Sierra Nevada, Mexico, 1994–2015. *J Maps* 15(2):418–424
- López-García J, Navarro-Cerrillo RM (2021) Changes in the constituents of the “Bosque de Agua” of the Sierra Cruces-Ajusco-Chichinautzín, Mexico, an area with payment for environmental services. *Environ Earth Sci* 80(20):1–14
- Malmer A, Murdiyarso D, Ilstedt U (2010) Carbon sequestration in tropical forests and water: a critical look at the basis for commonly used generalizations. *Glob Chang Biol* 16:599–604. <https://doi.org/10.1111/j.1365-2486.2009.01984.x>
- Margono BA, Potapov PV, Turubanova S, Stolle F, Hansen MC (2014) Primary forest cover loss in Indonesia over 2000–2012. *Nat Clim Change* 4(8):730–735. <https://doi.org/10.1038/nclimate2277>
- Martínez-Méndez N, Aguirre-Plante E, Eguarte LE, Jaramillo-Correa JP (2016) Modelado de nicho ecológico de las especies del género *Abies* (Pinaceae) en México: Algunas implicaciones taxonómicas y para la conservación. *Bot Sci* 94:5–24. <https://doi.org/10.17129/botsci.508>
- Mascia MB, Pailler S (2011) Protected area downgrading, downsizing, and degazettement (PADDD) and its conservation implications. *Conserv Lett* 4(1):9–20
- Mastretta-Yanes A, Cao R, Nicasio-Arzeta S, Quadri P, Escalante-Espinosa T, Arredondo L, Piñero D (2014) ¿Será exitosa la estrategia del cambio de categoría para mantener la biodiversidad del Nevado de Toluca? In: Eguarte LE (ed) Reflexiones sobre la ecología y la conservación en México. Instituto de Ecología-Universidad Nacional Autónoma de México, Mexico City, pp 7–17
- Mastretta-Yanes A, Moreno-Letelier A, Piñero D, Jorgensen TH, Emerson BC (2015) Biodiversity in the Mexican highlands and the interaction of geology, geography and climate within the Trans-Mexican Volcanic Belt. *J Biogeogr* 42(9):1586–1600. <https://doi.org/10.1111/jbi.12546>
- Meyfroidt P, Lambin EF (2009) Forest transition in Vietnam and displacement of deforestation abroad. *Proc Natl Acad Sci USA* 106:16139–16144. <https://doi.org/10.1073/pnas.0904942106>
- Milodowski DT, Mitchard ETA, Williams M (2017) Forest loss maps from regional satellite monitoring systematically underestimate deforestation in two rapidly changing parts of the Amazon. *Environ Res Lett* 12:094003
- Montoya E, Guzmán-Plazola RA, López-Mata L (2020) Fragmentation dynamics in an *Abies religiosa* forest of central Mexico. *Can J for Res* 50:680–688. <https://doi.org/10.1139/cjfr-2019-0235>
- Morrone JJ, Márquez J (2001) Halffter’s Mexican transition zone, beetle generalized tracks, and geographical homology. *J Biogeogr* 28:635–650. <https://doi.org/10.1046/j.1365-2699.2001.00571.x>
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403:853–858. <https://doi.org/10.1038/35002501>
- Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA (2014) Good practices for estimating area and assessing accuracy of land change. *Rem Sens Environ* 148:42–57. <https://doi.org/10.1016/j.rse.2014.02.015>
- Osorno-Covarrubias J, Couturier S, Piceno Hernández M (2018) Measuring from space the efficiency of local forest management: the successful case of the indigenous community of Cherán, Mexico. *IOP Conf Ser Earth Environ Sci* 151:012010. <https://doi.org/10.1088/1755-1315/151/1/012010>
- Pan Y, Birdsey R, Fang J, Houghton R, Kauppi PE, Kurz WA, Phillips OL, Shvidenko A, Lewis SL, Canadell JG, Ciais P (2011) A large and persistent carbon sink in the world’s forests. *Science* 333:988–993. <https://doi.org/10.1126/science.1201609>
- Persson M, Lindberg E, Reese H (2018) Tree species classification with multi-temporal Sentinel-2 data. *Remote Sens* 10(11):1794. <https://doi.org/10.3390/rs10111794>
- Qin S, Golden Kroner RE, Cook C, Tesfaw AT, Braybrook R, Rodríguez CM, Poelking C, Mascia MB (2019) Protected area downgrading, downsizing, and degazettement as a threat to iconic protected areas. *Conserv Biol* 33(6):1275–1285
- R Development Core Team (2020) R: a language and environment for statistical computing. R Foundation for statistical computing, Vienna
- Rommel TK, Csillag F, Mitchell S, Wulder M (2005) Integration of forest inventory and satellite imagery: a Canadian status assessment and research issues. *For Ecol Manage* 207:405–428. <https://doi.org/10.1016/j.foreco.2004.11.023>
- Rodríguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, Rigol-Sánchez JP (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J Photogramm Remote Sens* 67:93–104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>

- Rojas-García F, Fredericksen TS, Lozada SV, Agramont AR (2019) Impact of timber harvesting on carbon storage in montane forests of central Mexico. *New for* 50(6):1043–1061. <https://doi.org/10.1007/s11056-019-09714-z>
- Rojas-Merced EE, Valdez-Pérez ME, Mireles-Lezama P, Reyes-Enríquez A, Pastor-Medrano J (2007) Estimación de la producción de agua superficial del Parque Nacional Nevado de Toluca, para el año 2006. *Quivera Rev Estud Territ* 9(1):159–176
- Sáenz-Romero C, Rehfeldt GE, Duval P, Lindig-Cisneros RA (2012) *Abies religiosa* habitat prediction in climatic change scenarios and implications for monarch butterfly conservation in Mexico. *For Ecol Manage* 275:98–106. <https://doi.org/10.1016/j.foreco.2012.03.004>
- Sánchez-Jasso JM, Aguilar-Miguel X, Medina-Castro JP, Sierra-Domínguez G. (2013) Riqueza específica de vertebrados en un bosque reforestado del Parque Nacional Nevado de Toluca, México. *Revista Mexicana de Biodiversidad* 84(1):360–373. <https://doi.org/10.7550/rmb.29473>
- Sandel B, Arge L, Dalsgaard B, Davies RG, Gaston KJ, Sutherland WJ, Svenning JC (2011) The influence of late quaternary climate-change velocity on species endemism. *Science* 334:660–664. <https://doi.org/10.1126/science.1210173>
- Saunders SP, Ries L, Oberhauser KS, Thogmartin WE, Zipkin EF (2018) Local and cross-seasonal associations of climate and land use with abundance of monarch butterflies *Danaus plexippus*. *Ecography* 41:278–290. <https://doi.org/10.1111/ecog.02719>
- SEMARNAT (2010) Norma Oficial Mexicana NOM-059-SEMARNAT-2010, Protección ambiental-especies nativas de México de flora y fauna silvestres-Categorías de riesgo y especificaciones para su inclusión, exclusión o cambio. Lista de especies en riesgo. Diario Oficial de la Federación. [http://dof.gob.mx/nota\\_detalle.php?codigo=5173091&fecha=30/12/2010](http://dof.gob.mx/nota_detalle.php?codigo=5173091&fecha=30/12/2010). Accessed 5 Nov 2018
- Sheil D (2018) Forests, atmospheric water and an uncertain future: the new biology of the global water cycle. *For Ecosyst* 5:1–22. <https://doi.org/10.1186/s40663-018-0138-y>
- Shimizu K, Ota T, Mizoue N (2020) Accuracy assessments of local and global forest change data to estimate annual disturbances in temperate forests. *Remote Sens* 12(15):2438. <https://doi.org/10.3390/rs12152438>
- Sims KR, Alix-García JM (2017) Parks versus PES: evaluating direct and incentive-based land conservation in Mexico. *J Environ Econ Manage* 86:8–28. <https://doi.org/10.1016/j.jeem.2016.11.010>
- Soleimannejad L, Ullah S, Abedi R, Dees M, Koch B (2019) Evaluating the potential of Sentinel-2, Landsat-8, and IRS satellite images in tree species classification of Hyrcanian forest of Iran using Random Forest. *J Sustain* for 38(7):615–628. <https://doi.org/10.1080/10549811.2019.1598443>
- Stehman SV, Czaplewski RL (1998) Design and analysis for thematic map accuracy assessment: fundamental principles. *Rem Sens Environ* 64(3):331–344
- Sunny A, Domínguez-Vega H, Caballero-Viñas C, Ramírez-Corona F, Suárez-Atilano M, González-Fernández A (2021a) A Salamander tale: relative abundance, morphometrics and microhabitat of the critically endangered Mexican salamander *Pseudoeurycea robertsi* (Taylor, 1939). *Herpetozoa* 34:35–47
- Sunny A, González-Fernández A, López-Sánchez M, Ramírez-Corona F, Suárez-Atilano M, González-Fernández A (2021b) Genetic diversity and functional connectivity of a critically endangered salamander. *Biotropica*. <https://doi.org/10.1111/btp.13025>
- Tortini R, Mayer AL, Hermosilla T, Coops NC, Wulder MA (2019) Using annual Landsat imagery to identify harvesting over a range of intensities for non-industrial family forests. *Landsc Urban Plan* 188:143–150
- Toscana-Aparicio A, Granados-Ramírez R (2015) Recategorización del Parque Nacional Nevado de Toluca. *Polít Cult* 44:79–105
- Valencia S (2004) Diversidad del género *Quercus* (Fagaceae) en México. *Bol Soc Bot México* 75:33–53. <https://doi.org/10.17129/botsoci.1692>
- Yamada Y, Ohkubo T, Shimizu K (2020) Causal analysis of accuracy obtained using high-resolution global forest change data to identify forest loss in small forest plots. *Remote Sens* 12:2489. <https://doi.org/10.3390/rs12152489>

## Authors and Affiliations

**Andrea González-Fernández<sup>1,2</sup> · Joel Segarra<sup>3</sup> · Armando Sunny<sup>4</sup> ·  
Stephane Couturier<sup>1</sup> **

<sup>1</sup> Laboratory for Geospatial Analysis (LAGE), Geography Institute, National Autonomous University of Mexico (UNAM), Circuito Exterior s/n, Coyoacán, Cd. Universitaria, 04510 Mexico City, Mexico

<sup>2</sup> Autonomous Metropolitan University - campus Lerma, Hidalgo Pte. 46, Col. La Estación, 52006 Lerma, Estado de México, Mexico

<sup>3</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, Avinguda Diagonal, 643, 08028 Barcelona, Spain

<sup>4</sup> Center for Research in Applied Biological Sciences, Autonomous University of the State of Mexico, Instituto Literario # 100, Colonia Centro, 50000 Toluca, State of Mexico, Mexico





**Annex 3.** La gestió dels paisatges mediterranis muntanyosos abandonats: els efectes del guarà en el control de la biomassa i diversitat florística de les pastures

**Annex 3.** Managing abandoned Mediterranean mountain landscapes: The effects of donkey grazing on biomass control and floral diversity in pastures

Joel Segarra <sup>a, b, \*</sup>, Jordi Fernàndez-Martínez <sup>c</sup>, Jose Luis Araus <sup>a, b</sup>

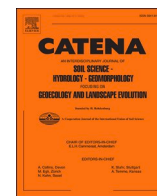
<sup>a</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain

<sup>b</sup> AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198 Lleida, Catalonia, Spain

<sup>c</sup> Laboratorio de Ruralización, Aldea de Puy de Cinca, 22439, Secastilla, Aragón, Spain

\*Correspondence

Publicat a la revista Catena



# Managing abandoned Mediterranean mountain landscapes: The effects of donkey grazing on biomass control and floral diversity in pastures

Joel Segarra<sup>a,b,\*</sup>, Jordi Fernàndez-Martínez<sup>c</sup>, Jose Luis Araus<sup>a,b</sup>

<sup>a</sup> Integrative Crop Ecophysiology Group, Plant Physiology Section, Faculty of Biology, University of Barcelona, 08028 Barcelona, Catalonia, Spain

<sup>b</sup> AGROTECNIO (Center for Research in Agrotechnology), Av. Rovira Roure 191, 25198 Lleida, Catalonia, Spain

<sup>c</sup> Laboratorio de Ruralización, Aldea de Puy de Cinca, 22439, Secastilla, Aragón, Spain

## ARTICLE INFO

### Keywords:

Land abandonment  
Pastures  
Floral diversity  
Sentinel-2  
Catalan donkey

## ABSTRACT

Traditional Mediterranean Mountain landscapes in Spain have suffered dramatic environmental and social changes over the last seven decades. The loss of these landscapes has had consequences for biodiversity, soil erosion, landscape quality and ecosystem services as croplands and pastures were mainly converted into forests and scrublands. Many animal breeds present in traditional land uses such as the Catalan donkey (*Equus asinus* var. *catalana*), are also at risk of extinction, but can provide environmental services while recovering traditional landscapes. In Puy de Cinca, a village in mountainous northeast Spain, we studied how grazing by Catalan donkeys reduces pasture biomass and the effects on plant diversity in pastures. We used Sentinel-2 satellite imagery to calculate the Normalized Difference Vegetation Index (NDVI), a biomass sensitive spectral index, throughout the grazing period to monitor pasture biomass and compared it to pastures without grazing. We also calculated several plant diversity indicators in pastures with and without grazing donkeys. Furthermore, we studied land use changes over the last seven decades using old (1956) and current (2018) aerial images, with forest, agricultural lands, and trails interconnecting the village being mapped to understand landscape changes. The results indicated a great increase in the forested area ( $348.3 \pm 17.0$  ha). Meanwhile, a severe decrease in cropland area (73 %) and trail length (62.5 %) was also observed. Concerning the effect of donkey grazing, biomass was lower in pastures with grazing donkeys, with NDVI values decreasing once donkeys started grazing. Nevertheless, plant diversity was higher in pastures with grazing donkeys than in abandoned pastures. This study demonstrated the capacity of low-to-moderate-intensity donkey grazing to improve plant diversity and reduce biomass in pastures. Furthermore, the study of land use changes allowed an understanding of landscape dynamics, which can help address the social and environmental recovery of the village.

## 1. Introduction

Mountainous areas across Spain have suffered dramatic social and environmental changes over the last few generations. The changes that the country experienced from the mid-nineteenth to the twentieth century have contributed to population growth and therefore an increase in demand for agricultural products, which has required more land to be put under cultivation (Collantes and Pinilla, 2012; Garrabou i Segura, 1985). Along with agricultural expansion, a considerable increase in the extraction of forest resources during industrialization and urbanization contributed to a generalized deforestation (Iriarte-Goni, 2008). Since the mid-twentieth century, the mechanization and intensification of farming in the most suitable lowlands made mountain farming

economically unviable, and resulted in a generalized abandonment of the hillslopes (García-Ruiz et al., 1996). In the case of the Spanish Pyrenees and the Iberian Range, more than 90 % of the originally cultivated land has been abandoned (Ruiz-Flaño, 1993). Moreover, in some mountain areas in Spain, dams were built and the most productive areas ended under water (García-Ruiz, 1977). Many villages also ended under water, and others were abandoned by the locals who moved to larger towns or metropolitan areas because they could not earn their livelihoods in the countryside.

These changes have degraded many traditional mountain landscapes, which were territories shaped by the social, cultural and physical realities of their historical communities (Greider and Garkovich, 2010). The loss of traditional landscapes has had consequences for biodiversity,

\* Corresponding author.

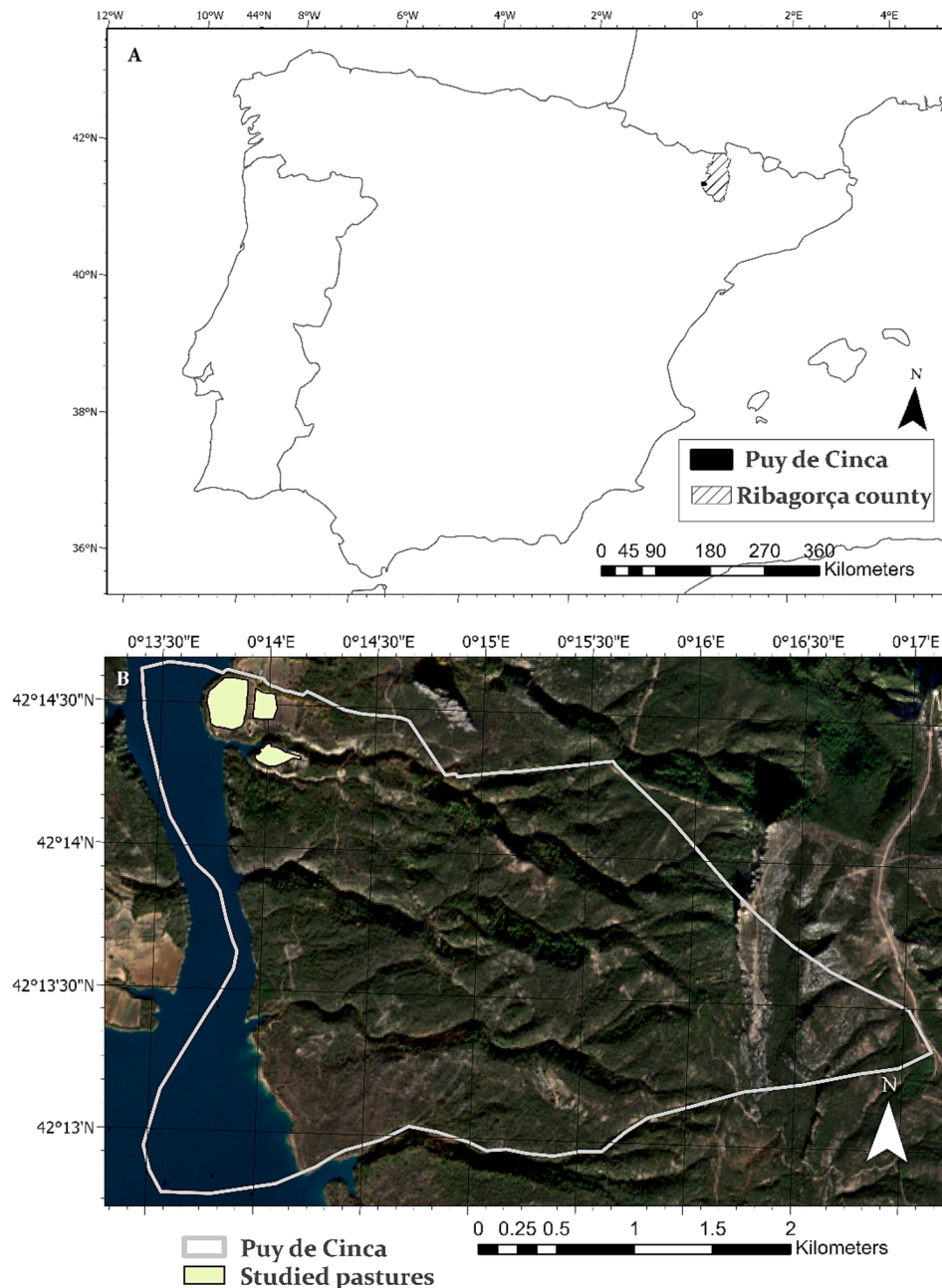
E-mail address: [joel.segarra@ub.edu](mailto:joel.segarra@ub.edu) (J. Segarra).

<https://doi.org/10.1016/j.catena.2023.107503>

Received 14 February 2023; Received in revised form 30 August 2023; Accepted 2 September 2023

Available online 6 September 2023

0341-8162/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

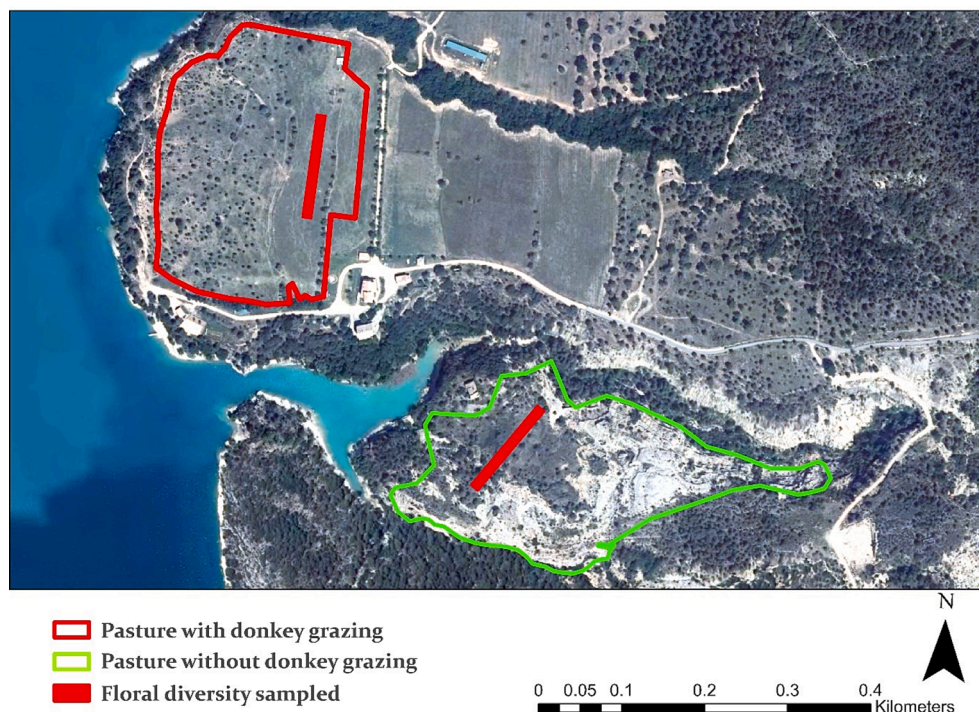


**Fig. 1.** The location of Puy de Cinca and Ribagorça county in northeast Spain is shown in A, while B shows the extent of Puy de Cinca and the location of the studied pastures overlies a false color Sentinel-2 image captured on the 21st of November 2020.

soil erosion, landscape quality and ecosystem services (Geri et al., 2010; Navarro and Pereira, 2015). Croplands and pastures were mainly converted into scrublands and forests which, alongside reforestation and afforestation based fundamentally on conifers (especially *Pinus* spp.) since the 1950s (Gomez Mendoza and Mata Olmo, 1992), has resulted in fuel accumulations that have increased the fire risk in most of these unstructured landscapes (Duguy et al., 2007; Jiménez-Olivencia et al., 2021). These traditional mountain landscapes are particularly vulnerable to global and climate change (Guiot and Cramer, 2016; Zamora et al., 2007), however, sustainability scenarios to mitigate and adapt to climate change imply conservation and recovery of these landscapes (Malek et al., 2018).

In contrast with other areas of the world, in Europe there is evidence of declines in the previously high biological value of long established agroecosystems (Donald et al., 2001) following agricultural

intensification and abandonment of marginal and mountainous lands. Within the Mediterranean Basin, traditional land use systems had a high biological diversity, which probably never existed in pristine oak woodlands but did exist in systems that were moderately modified such as traditional agro-silvo-pastoral systems (Blondel and Aronson, 1995). For instance, pastures account for 47 % of plant species of the total flora of the Aragonese Pyrenees in Spain, but pasture barely occupies 10 % of the region's area (García and Gómez, 2007). At the same time, many animal breeds present in traditional Mediterranean land uses such as *Equus asinus* var. *catalana*, known as the Catalan donkey, are also at risk of extinction (Bartolomé et al., 2020; Gutierrez et al., 2005; Jordana i Vidal and Folch Lopez, 1998). Of all livestock, donkeys have experienced the greatest effects of rural depopulation and agricultural industrialization, and most European donkey breeds are disappearing (Kugler et al., 2007). At the beginning of the twentieth century, the number of



**Fig. 2.** Location of floral diversity sampling in a pasture with donkey grazing and an abandoned pasture, the transects followed during the sampling of floral diversity are indicated. The background image is an aerial orthophoto composition generated by the Spanish National Plan of Aerial Orthophotography in 2018.

Catalan donkeys was around 50,000, but by the 1990s only 100 Catalan donkeys were left (Jordana, 2010). Some effort has been made since then to recover the declining population of this landrace and currently the total population of Catalan donkeys is around 1,000. Therefore, providing donkeys with an environmental service role and strengthening their function in recovering traditional mountain landscapes could help recover the donkey's declining population.

Allowing large herbivores to roam freely, within the carrying capacity of an area, can contribute, to some extent, towards biodiversity conservation, and it is sustainable economically (Schwabe et al., 2013; Timmermann et al., 2015) because forage supply is generally not needed. Moreover, it is sustainable environmentally as this practice limits the input of additional nutrients into the habitat. Some authors have also observed an increase in floral diversity in pastures supporting large herbivores (Köhler et al., 2016), while others argue that grazing might have negative effects such as soil erosion and biodiversity loss in some cases (Papanastasis et al., 2002), especially due to overgrazing on poor soils. Besides biodiversity, large herbivores, such as donkeys, feed on some forest species and can play a role in fire prevention by controlling the forest understory (Bartolomé et al., 2020). Donkeys have been used successfully to control some flammable grasses in the Mediterranean scrub (Gulías et al., 2016). In other mountainous areas in Spain, different livestock species have also been used in forest lands to prevent fire risk and to restore pastures (Álvarez-Martínez et al., 2016; Casasús et al., 2004). In this sense, grazing can contribute to managing forest understory vegetation to prevent fire, and thus restore pastures, increase biodiversity (Verdú et al., 2000), and provide ecosystem services such as the activation of rural economies (Gutiérrez-Peña et al., 2016).

The study we present aims to assess the effects of Catalan donkey grazing on biomass control and floral diversity in pastures for an improved environmental management of an abandoned mountain landscape. We focus on a mountainous area (Puy de Cinca, Aragon, northeast Spain), which has a deeply degraded traditional landscape due to the abandonment of settlements in the mid-twentieth century after the construction of a dam during Francoist Spain. However, since the

end of the 1990s the local farmers union has addressed the rehabilitation of some settlements and promoted the reestablishment of local animal landraces. Currently, there is a significant Catalan donkey population used for transportation by some locals, as well as to reduce forest understory, recover pastures and the landrace itself. In this study, we have used an integrative approach by combining the study of traditional landscape changes, plant diversity assessment, and remote sensing that, to our knowledge, has not yet been used for monitoring donkey grazing effects on pasture biomass. With this aim we have structured the study around three research questions. First, how much has the traditional landscape changed since the mid twentieth century (1956–2018) in Puy de Cinca? Second, does the grazing of Catalan donkeys reduce pasture biomass and phytovolume? Third, is floral diversity in pastures with grazing donkeys greater than in abandoned pastures?

## 2. Materials and methods

### 2.1. Study area

In Aragonese, the local language, Puy de Cinca (known as Pui de Cinca in Aragonese) means an elevated territory above the Cinca river. After the construction of a dam in 1960s, the village was abandoned either due to the existence of certain pressures or a strategy to encourage the affected population to favor voluntary expropriation (Buil, 1993). Significant parts of the village were flooded because of the construction of the Grado dam on the Cinca river, and other villages that were located close to the river basin were also affected requiring inhabitants to relocate to other towns and cities. Puy de Cinca is characterized by two main settlements; on the one hand, the main historic village of Puy de Cinca, and on the other, a settlement of disperse traditional farmhouses called Aldea de Puy de Cinca. This municipality, as well as most of those that comprise the Aragon highlands, underwent reforestation with *Pinus* spp. after the 1950s (Grasa, 2017).

In the late 1990s Aldea de Puy de Cinca started a rehabilitation process since the owner of the land, the company managing the region's hydraulic infrastructure (Confederación Hidrológica del Ebro) agreed to





Fig. 3. Sampling 1 m<sup>2</sup> to determine floral diversity.

a stewardship with the Union of Farmers of Aragon (Union de Agricultores y Ganaderos de Aragon). Several buildings were restored and 20 ha of olive trees, 2 ha of vineyards and some pastures were rehabilitated. Moreover, the Union of Farmers of Aragon also promoted the conservation of animal landraces local to the Pyrenees such as the Catalan Donkey, the Churra Tensina sheep or the Pyrenean goat (Azon-Pardo, 1998). Nowadays, the village has four inhabitants who run an inn where environmental courses are held. The village occupies about 1,000 ha, of which about 90 ha area are now inundated (Fig. 1).

In the study, we assessed land use changes across the entirety of Puy de Cinca (Ribagorça county, northeast Spain), an area located in the transition between the Mediterranean climate and mountainous temperate climate typical of the Pyrenees (Fig. 1). In regard to grazing, we studied three pastures: a donkey-grazed pasture (from June to September 2020), a pasture that was recovered in previous years but was not grazed during the studied season, and an abandoned pasture that had not been completely restored and had undergone minimal grazing in recent times. During the study season, the latter abandoned was not grazed. The two recovered pastures had been recovered in the previous six years. The locations of the pastures are shown in Fig. 1, and detailed information of each pasture and the methodology implemented is described in the subsequent sections.

## 2.2. Floral diversity

The floral diversity was studied in two pastures of similar sizes encompassing 6.5 ha of the pastured area, and 5 ha of the abandoned area (Fig. 2). The pasture grazed by donkeys (42°14'32.43" N, 0°13'53.31" E at 465 m.a.m.s.l.) during the study season (June to September 2020) carried eight Catalan donkeys (*Equus asinus* var. *cat-alana*) on a permanent basis (aged between two and 12 years), which can be considered a low-to-moderate grazing intensity with approximately 1.2 animals per ha, since low, moderate and high grazing intensities correspond to 0.5, 1.5, and 2.5 animal unit/ha, respectively (Török et al., 2016). This pastured area had been visited by donkeys since 2014 and therefore its abandonment was considered reversed. The donkeys started to graze the pasture in June 2020 so, the floral diversity was assessed throughout the seasons until spring (2021) because the results of floral diversity depend on the flowering stages. A GPS collar (Digitanimal, Madrid, Spain), which is specially designed for animals, was hung around the neck of one adult donkey according to the animal welfare regulations (11/2003 Spanish law on animal welfare). The GPS collar was checked regularly, as well as the general well-being of all donkeys following the regulations for Equidae in Spain (804/2011 Spanish Royal Decree). GPS devices can help monitor the welfare of animals and behavioral parameters (Herlin et al., 2021). The GPS collar determined the grazing patterns and locations of the selected donkey to enable better management of the herd, but the data were not included in the study because most donkeys were not geolocated. From the

observation of GPS points and by site managers, we can argue that grazing was distributed relatively evenly throughout the pasture. The second pasture was abandoned in the 1960 s (42°14'22.86" N, 0°14'02.94" E, at 472 m.a.m.s.l.), and has a rugged terrain that makes its restoration more difficult. It was sparsely grazed by free-roaming donkeys nearly a decade ago, so the vegetation has experienced minimal alteration. During the study season no donkey grazing was permitted in this area in order to avoid effects on the plant community. In each pasture, either grazed and not grazed, 1 m<sup>2</sup> squares were placed every 10 m in a 100 m transect, totaling 10 squares, and covering a significant part of the pastures (Fig. 2). Samples were taken in each of the squares throughout the transect (Fig. 3). All species were collected and classified.

To assess the floral diversity, a detrended correspondence analysis (DCA) (Hill and Gauch, 1980) was used to explore the main gradients of plant community variability in the grazed and non-grazed pastures. DCA was calculated in R using the vegan package (Oksanen et al., 2015) with data obtained from plant diversity indicators based on Albuquerque et al. (2010). The absolute frequency (Fa, Equation (1)), the relative frequency (Fr, equation (2)), the absolute dominance (Doa, equation (3)), the relative dominance (Dor, equation (4)) and the importance value (IV, Equation (5)) were calculated from the obtained samples for each species:

$$Fa_i = (n_i/N) \cdot 100 \quad (1)$$

Where  $Fa_i$  = Absolute frequency of species  $i$ ;  $n_i$  = number of sample units where species  $i$  was found;  $N$  = total number of sample units.

$$Fr_i = \left( \frac{Fa_i}{\sum Fa} \right) \cdot 100 \quad (2)$$

Where  $Fr_i$  = Relative frequency of species  $i$ ;  $Fa_i$  = Absolute frequency of species  $i$ ;  $\sum Fa$  = summation of the absolute frequencies of the species.

$$Doa_i = \left( \frac{\sum PC_i}{N} \right) \cdot 100 \quad (3)$$

Where  $Doa_i$  = Absolute dominance of species  $i$ ;  $PC_i$  = percentage of coverage of the species per sampled unit;  $N$  = total number of sample units.

$$Dor_i = \left( \frac{Doa_i}{\sum Doa} \right) \cdot 100 \quad (4)$$

Where  $Dor_i$  = Relative dominance of species  $i$ ;  $Doa_i$  = absolute dominance of species  $i$ ;  $\sum Doa$  = total of the absolute dominance of the species.

$$IV_i = (Fr_i + Dor_i)/2 \quad (5)$$

Where  $IV_i$  = Importance value index of species  $i$ ;  $Fr_i$  = Relative frequency of species  $i$ ;  $Dor_i$  = Relative dominance of species  $i$ .

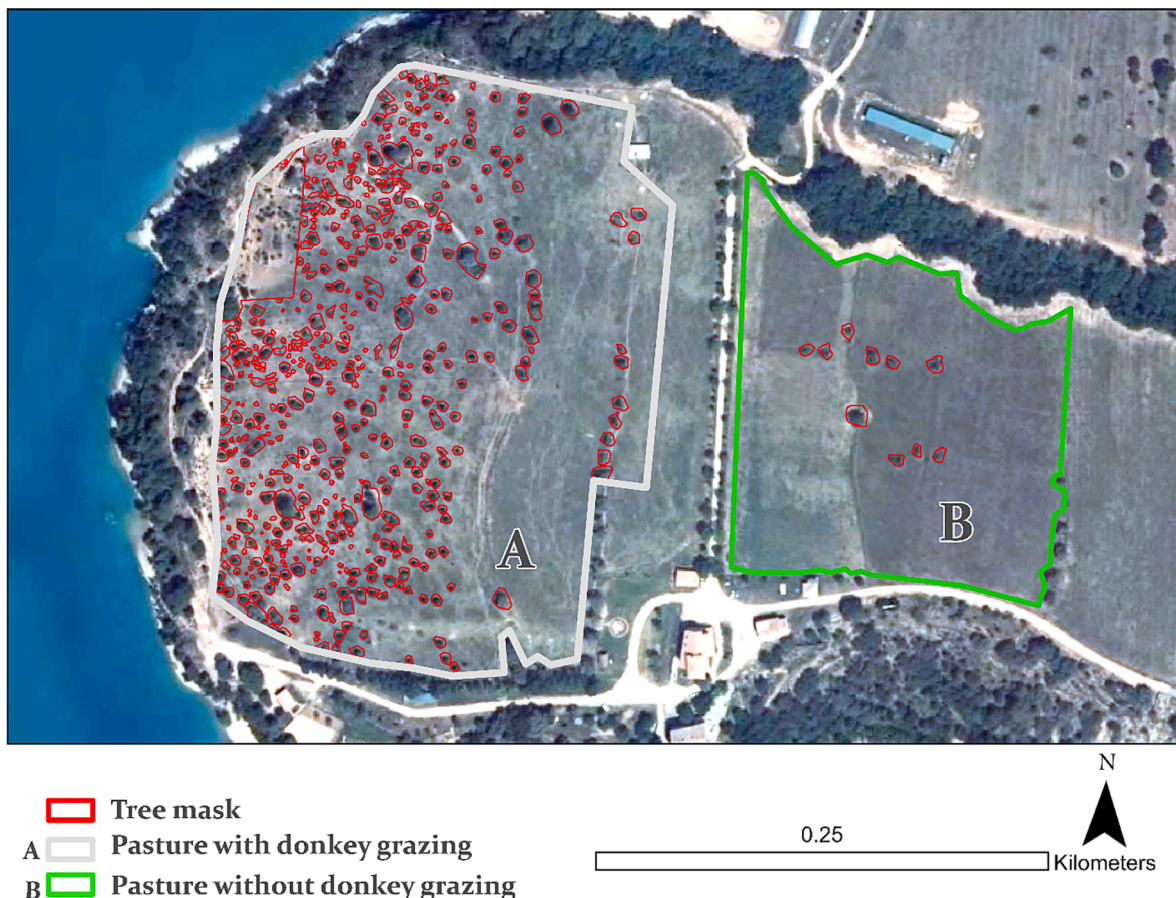
Species diversity was also calculated for each habitat using the Shannon-Wiener index ( $H'$ , equation (6) (Shannon and Weaver, 1949)), and Pielou's equity index ( $J'$ , equation (7) (Pielou, 1966)), according to the formulas:

$$H' = - \sum p_i (\ln p_i) \quad (6)$$

Where  $H'$  = Shannon-Wiener diversity index;  $p_i$  = relative dominance of species  $i$ ;  $\ln p_i$  = Neperian logarithm of  $p_i$ .

$$J' = \frac{H'}{\ln S} \quad (7)$$

Where  $J'$  = Pielou's equity index;  $H'$  = Shannon-Wiener diversity index;  $S$  = total number of sampled species;  $\ln S$  = Neperian logarithm of  $S$ . The Pielou index measures the proportion of the observed diversity in relation to the maximum expected diversity. Its value ranges from 0 to 1, so that 1 corresponds to situations where all species are equally



**Fig. 4.** Pasture with donkey grazing indicated with an A and the control pasture B without donkey grazing are shown. Tree mask is also delineated in red. The background image is an aerial orthophoto composition generated by the Spanish National Plan of Aerial Orthophotography in 2018. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

abundant.

To interpret the floristic similarity between the plots, the Jaccard similarity index ( $IS_j$ , equation (8) (Jaccard, 1908)), was used as a measure of distance. The similarity index ( $IS_j$ ) was also used to compare the analyzed habitats.

$$IS_{ja, b} = C / A + B + C \quad (8)$$

Where  $IS_j$  a, b = floristic similarity by means of the Jaccard index between habitats a and b (pastured and not pastured); C = number of species common to both habitats; A = number of species exclusive to habitat a; B = number of species exclusive to habitat b.

### 2.3. Remote sensing data and processing

To assess land use changes in Puy de Cinca, we used two aerial scenes, one from 1956 and the other from 2018 at 0.5 m spatial resolution. The images were downloaded from the regional Aragon government geodata server for the municipality of Secastilla where Puy de Cinca is located ([https://idearagon.aragon.es/fichaDescarga/fichaDescarga\\_22214.html](https://idearagon.aragon.es/fichaDescarga/fichaDescarga_22214.html)). The 1956 image is now publicly available, a black-and-white photograph originally captured by a military aircraft. The 2018 scene was captured within the National Plan of Aerial Orthophotography from the Spanish Government; it is a red-green-blue orthophoto. In both cases, forest areas were mapped by establishing a threshold pixel value, 100 points distributed in the forest landscape were visually spotted and averaged to obtain the pixel value thresholds to differentiate forested areas from other land uses in 1956 and 2018. In the case of the 1956 photograph, pixels between 0 and 100 values were considered forested areas, while in the 2018 orthophoto, pixels between

15 and 85 values were considered forested areas. This processing was carried out in ArcGIS Pro 3.1.0. by selecting a logical expression that selects a subset of raster cells within the established thresholds, and subsequently the forested areas in both periods were quantified with the corresponding confidence intervals according to Olofsson et al. (2014). Moreover, agricultural lands in 1956 were visually delineated and compared to current declared agricultural activities within geographic information systems of the common agrarian policy of the European Union, obtained at <https://idearagon.aragon.es/descargas>. We decided to visually delineate the agricultural areas because the spectral resolution of the black-and-white 1956 scene did not provide enough information to calculate an accurate classification algorithm. We therefore assessed the agricultural land from previous documents and direct consultation with locals to focus on specific areas; riverbanks and terraced hills close to the settlements. Trails were delineated based on 1956 aerial images and compared to the current trails. A visual example of what was delineated as trails in the 1956 aerial images is shown in the appendix (Fig. A1).

The classification of forest changes in the two scenes (1956 and 2018) was re-classified into four categories to improve the assessment of forested area and its classification accuracy: no forest, namely pixels that currently do not host forest and did not host forest in the 1956 scene; forest loss, where forest pixels identified for 1956 were lost by 2018; forest gain, where forest pixels gained by 2018 contrasted with forest absence in 1956; and stable forest, with forest pixels that were stable across the studied time period. To assess the accuracy of the classification of forest changes, a confusion matrix was computed by assessing 150 random points on ArcGIS Pro 3.1.0. The accuracy assessment obtained from the confusion matrix was used to calculate the areas for each



**Table 1**

Phytosociological parameters of the identified species and those not identified in donkey-grazed pasture. The parameters are ordered from the highest to the lowest importance value (IV). Absolute frequency (Fa), relative frequency (Fr), absolute dominance (Doa), and relative dominance (Dor) are also shown.

Species	Fa (%)	Fr (%)	Doa (%)	Dor (%)	IV (%)
<i>Medicago minima</i>	100	5.78	61.5	19.15	12.46
<i>Cynodon dactylon</i>	100	5.78	55	17.12	11.45
<i>Medicago rigidula</i>	100	5.78	32.1	9.99	7.89
<i>Plantago lanceolata</i>	100	5.78	21.5	6.69	6.24
<i>Knautia arvensis</i>	100	5.78	18.6	5.79	5.79
<i>Sanguisorba minor</i>	90	5.20	17.2	5.35	5.28
<i>Salvia verbenaca</i>	100	5.78	14.5	4.51	5.15
<i>Dactylis glomerata</i>	100	5.78	9.4	2.93	4.35
<i>Crepis capillaris</i>	80	4.62	12.7	3.95	4.29
<i>Sherardia arvensis</i>	80	4.62	4.6	1.43	3.03
<i>Hypochaeris radicata</i>	70	4.05	4.4	1.37	2.71
<i>Urospermum picroides</i>	80	4.62	1.6	0.50	2.56
<i>Aegilops geniculata</i>	40	2.31	8.8	2.74	2.53
<i>Potentilla reptans</i>	30	1.73	10.3	3.21	2.47
<i>Trifolium pratense</i>	20	1.16	10.5	3.27	2.21
<i>Bromus erectus</i>	50	2.89	3.1	0.97	1.93
<i>Torilis nodosa</i>	40	2.31	4.7	1.46	1.89
<i>Galium album</i>	30	1.73	5	1.56	1.65
<i>Cenarium tenuiflorum</i>	40	2.31	1.1	0.34	1.33
Not identified 1	40	2.31	0.8	0.25	1.28
<i>Eringium campestre</i>	40	2.31	0.8	0.25	1.28
<i>Hypericum perforatum</i>	20	1.16	4.3	1.34	1.25
<i>Brachypodium phoenicoides</i>	10	0.58	6	1.87	1.22
<i>Malva sylvestris</i>	30	1.73	1	0.31	1.02
<i>Anagallis foemina</i>	30	1.73	0.8	0.25	0.99
<i>Ononis repens</i>	30	1.73	0.7	0.22	0.98
<i>Euphorbia falcata</i>	20	1.16	1.8	0.56	0.86
<i>Pethroragia sp.</i>	20	1.16	0.7	0.22	0.69
<i>Koeleria vallesiana</i>	20	1.16	0.6	0.19	0.67
<i>Lotus corniculatus</i>	20	1.16	0.5	0.16	0.66
Not identified 2	10	0.58	2	0.62	0.60
<i>Silene vulgaris</i>	10	0.58	2	0.62	0.60
<i>Alyssum alyssoides</i>	10	0.58	1	0.31	0.44
Not identified 3	10	0.58	0.3	0.09	0.34
<i>Medicago sativa</i>	10	0.58	0.3	0.09	0.34
<i>Geranium molle</i>	10	0.58	0.2	0.06	0.32
<i>Carex sp.</i>	10	0.58	0.2	0.06	0.32
<i>Hippocrepis unisiliquosa</i>	10	0.58	0.2	0.06	0.32
Dead plant	10	0.58	0.2	0.06	0.32
<i>Achillea millefolium</i>	10	0.58	0.2	0.06	0.32

of the four forest classes with corresponding confidence intervals. Error margins at 95 % probability for each class were derived from the set of error pixels and matched pixels according to the following equations (equation (9):

$$S(\hat{A}_k) = A \times S(\hat{p}_{\bullet k}) \quad (9)$$

Where  $\hat{A}_k$  is the estimated area of the mapped change class k (in our case it has four classes), A is the total map area,  $\hat{p}_{\bullet k}$  is the estimated area proportion of change in class k derived from the error matrix (see notations in Olofsson et al., (2014): Eqs. (9), 10 and 11), page 52), and S is the standard error function of the estimators.  $S(\hat{p}_{\bullet k})$ , shown in equation (10), is the standard deviation of the area proportion of the corresponding class. An approximate 95 % confidence interval is obtained as  $\hat{A}_k \pm 1.96 \times S(\hat{A}_k)$ .

$$S(\hat{p}_{\bullet k}) = \sqrt{\sum_i W_i^2 \frac{n_{ik} \left(1 - \frac{n_{ik}}{n_i}\right)}{n_i - 1}} \quad (10)$$

As an indicator of biomass cover in pastured and non-pastured pastures, Sentinel-2 images were downloaded free of clouds on the Copernicus Open Access Hub (<https://scihub.copernicus.eu/>), for the grazing period June to September 2020, during the donkeys' presence, and we obtained 12 images (24 June, 29 June, 19 July, 29 July, 8 August, 23 August, 27 August, and 12 September). Moreover, we also downloaded

**Table 2**

Phytosociological parameters of the identified species in the abandoned pasture without grazing donkeys. The parameters are ordered from the highest to the lowest value of importance (IV). Absolute frequency (Fa), relative frequency (Fr), absolute dominance (Doa), and relative dominance (Dor) are also shown.

Species	Fa (%)	Fr (%)	DoA (%)	Dor (%)	IV (%)
<i>Brachypodium phoenicoides</i>	70	11.29	43.4	31.40	21.35
<i>Knautia arvensis</i>	70	11.29	19	13.75	12.52
<i>Galium album</i>	90	14.52	13	9.41	11.96
<i>Dorycnium hirsutum</i>	70	11.29	14	10.13	10.71
<i>Odontites luteus</i>	60	9.68	11.3	8.18	8.93
<i>Eringium campestre</i>	40	6.45	8	5.79	6.12
<i>Pilosella officinarum</i>	30	4.84	9.8	7.09	5.96
<i>Plantago lanceolata</i>	40	6.45	4.3	3.11	4.78
<i>Agrimonia eupatoria</i>	40	6.45	0.4	0.29	3.37
<i>Catananche caerulea</i>	20	3.23	4.7	3.40	3.31
<i>Sanguisorba minor</i>	30	4.84	2.2	1.59	3.22
<i>Prunella laciniata</i>	10	1.61	2	1.45	1.53
<i>Genista scorpius</i>	10	1.61	2	1.45	1.53
<i>Lavandula latifolia</i>	10	1.61	2	1.45	1.53
<i>Potentilla reptans</i>	10	1.61	1.5	1.09	1.35
<i>Hippocrepis unisiliquosa</i>	10	1.61	0.3	0.22	0.91
<i>Medicago sativa</i>	10	1.61	0.3	0.22	0.91

images (9) in previous and subsequent months to see the evolution of the pastures throughout the year (31 January, 10 February, 20 February, 21 March, 20 May, 30 May, 17 October, 21 November, and 26 December). These images were downloaded as Sentinel-2 2A bottom-of-reflectance products, which are already corrected for atmospheric interference. Two bands were used, red (Band 4) and near-infrared (Band 7) at 10 m spatial resolution to obtain a biomass indicator of pastured and non-pastured pastures. Subsequently, the normalized difference vegetation index (NDVI) (Rouse et al., 1974) was used because it is widely applied as a biomass indicator for forest cover and pastures (Maselli et al., 1998). The NDVI (Equation (11)) was calculated for all the dates mentioned in the two pastures shown in Fig. 4. Moreover, due to the presence of large trees within the pastures, trees were masked to differentiate them from pasture pixels and avoid interferences in NDVI calculations (Fig. 4). In the pastured area (A in Fig. 4), the same eight donkeys mentioned in the description of the floral diversity assessment were present and grazing, while in the non-pastured area, no donkeys were present. We decided to use these two pastures to determine biomass when grazing because the control pasture had some recovery in previous years (it had been grazed in the previous 2019 season). Meanwhile, the abandonment of the pasture used for the floristic diversity diffculted observing remotely the pastured area due to the occurrence of shrubs and trees.

$$NDVI = \frac{(NearInfraredB8 - RedB4)}{(NearInfraredB8 + RedB4)} \quad (11)$$

### 3. Results

A total of 56 floral species (Tables 1 and 2) were detected in pastures with and without donkey grazing, from which 39 grew in the donkey-grazed pasture and 17 grew in the abandoned pasture. Hence, the number of species found in the pasture with donkeys was 56 % greater than the number of species in the abandoned pasture. In the donkey-grazed pasture three species could not be identified as the plants did flower or bear fruit, and therefore could not be classified. From the total number of species registered, 34 were identified to the specific level and represented 18 botanical families. In the donkey-grazed pasture the 18 botanical families were present, but the abandoned pasture only had 9. The most abundant families were Asteraceae (7), Poaceae (6) and Fabaceae-Faboidae (9).

The species with the highest IV in the donkey-grazed pasture were *Medicago minima*, *Cynodon dactylon*, *Medicago rigidula*, *Plantago lanceolata* and *Knautia arvensis* (Table 1). Meanwhile, in the abandoned pasture the species that had the highest IV were: *Brachypodium*



**Table 3**  
Number of species (S), Shannon-Wiener (H') biodiversity index and Pielou (J') index for the donkey-grazed pasture and the abandoned pasture.

Habitat	S	H'	J'
Donkey-grazed pasture	39	2.77	0.76
Abandoned pasture	17	2.21	0.78

**Table 4**  
Confusion matrix for the forest change map 1956–2018 shown in Fig. 5. Users accuracy (UA), producers accuracy (PA), and overall accuracy (OA) are shown, with 150 pixels being used for the accuracy assessment.

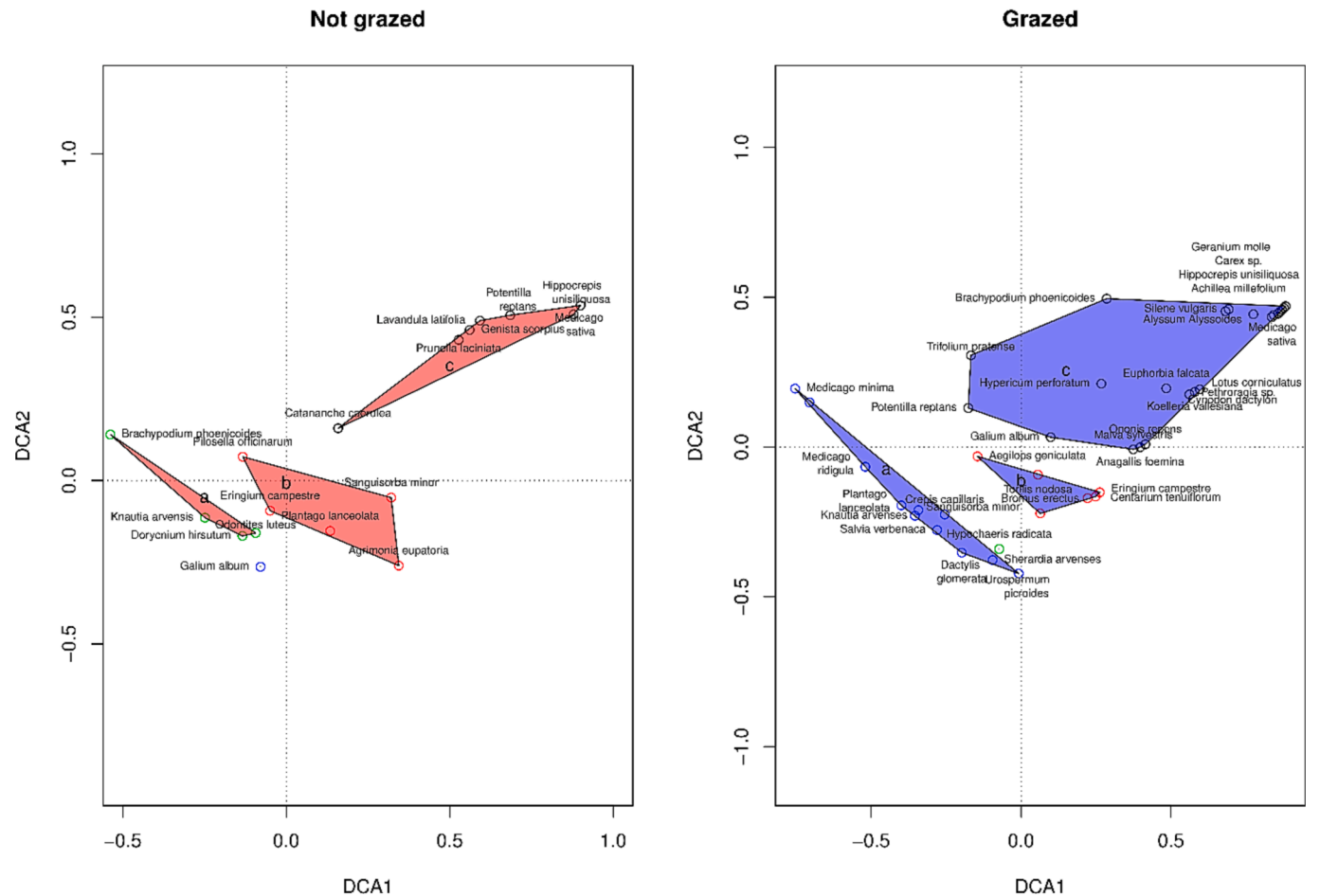
	No forest	Forest loss	Forest gain	Stable forest	UA [%]
No forest	19	2	0	0	90.5
Forest loss	2	9	0	1	75.0
Forest gain	0	0	59	5	92.2
Stable forest	0	0	8	45	84.9
PA [%]	90.5	81.8	88.1	88.2	OA = 88.0

*phoenicoides*, *Knautia arvensis*, *Galium* sp., *Dorycnium hirsutum* and *Odontites luteus* (Table 2). The IV of the first five species was higher in the abandoned pasture than in the donkey-grazed pasture. In the abandoned pasture, *Brachypodium phoenicoides* had the highest IV value (IV = 21.35 %), whereas in the donkey-grazed pasture *Medicago minima* had the highest (IV = 12.46 %). In all cases, the IV was different for the

equivalent species between the abandoned pasture and the donkey-grazed pasture. Moreover, the species with the highest IV values differed between the abandoned pasture and the donkey-grazed pasture.

Among the group of species with the lowest IV values, however, some species were equivalent between the pastures (*Potentilla reptans*, *Hippocrepis unisiliquosa*, and *Medicago sativa*). Nonetheless, in absolute numbers they were not the same. The low number of species found in the abandoned pasture and the high IV values among the most abundant species indicated the dominance of certain species. In contrast to the abandoned pasture in which *Brachypodium phoenicoides* dominated, in the donkey-grazed pasture this species had a VI of 1.22 %, which represents a 17-fold decrease relative to the abandoned pasture.

Regarding phytosociology, the Shannon-Wiener (H') and Pielou (J') indices in the donkey-grazed pasture were H' = 2.77 and J' = 0.76, compared to the values of H' = 2.21 and J' = 0.78 in the control pasture (Table 3). The results suggest that a higher floral diversity was found in the donkey-grazed pasture in comparison with the abandoned pasture. Moreover, the plant community was more heterogenous in the donkey-grazed pasture, with J' = 0.76, than in the abandoned pasture J' = 0.78 (Table 4). These indices indicated the presence of more species with less relative dominance values in the donkey-grazed pasture, which result in a better species diversity and a heterogenous vegetal community. From the 56 species in both pastures, only 9 were common to the two, which resulted in a 18.75 % of similarity under the Jaccard Similarity index (ISJa,b = 0.1875). In fact, 31 species are specific to the donkey-grazed pasture and 8 are specific to the abandoned pasture. In two pastures located so close together, one would expect a similar composition;



**Fig. 5.** Detrended correspondence analysis (DCA) ordination of plant communities in the grazed and non-grazed pasture. In both cases the DCA2 and DCA1 axes were selected to explain most of the variance. The color of the points indicates the grouping of species in each pasture, and the species name of each point is indicated with a label.

**Table 5**

Estimated area and confidence intervals of forest classes calculated according to Olofsson et al. (2014) from the map shown in Fig. 6 and the accuracy assessment obtained with the confusion matrix shown in Table 1.

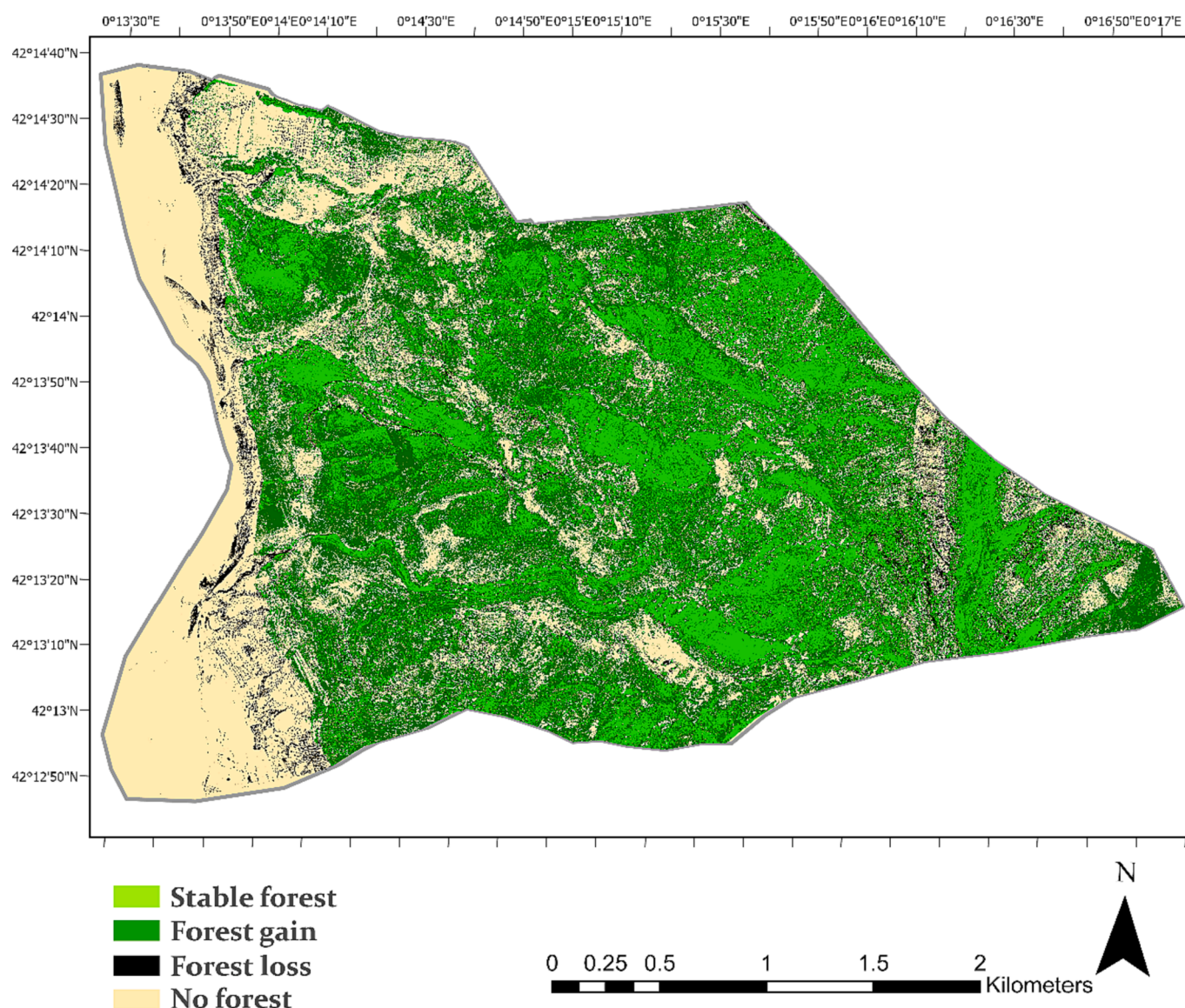
Class	Estimated area (ha)	Confidence intervals (ha)
No forest	188.4	33.6
Forest loss	78.1	32.9
Forest gain	348.3	17.0
Stable forest	242.4	53.6

nonetheless, it seems that the donkeys' grazing influenced the floral diversity.

In this sense, the DCA results in Fig. 5 show a clear separation among the compositions of three plant communities in each pasture. In the non-grazed DCA plot, the DCA1 axis explained 63 % and the DCA2 axis 36 % of the total variability in the species dataset. Meanwhile, in the grazed DCA plot, the DCA1 axis explained 59 % and the DCA2 axis 22 % of the total variability in the species dataset. In the grazed pasture, species with lower frequencies showed higher floristic variation than in the non-grazed pasture. Regarding the community formed by the most common species, the grazed pastures showed a greater diversity than the

non-grazed pasture (Fig. 5).

In the case of forest change, the classification we calculated achieved a relatively high overall accuracy of 88 % (Table 4). In general, all forest change classes have been correctly classified. The forest loss class was the least well classified, with some misclassification having occurred because the forest loss areas were minimal and mainly located along the former riverbank and some roads developed in recent times. Moreover, this class was susceptible to misclassifications due to shaded pixels, especially related to single trees. The estimated areas and confidence intervals calculated for each forest change class are shown in Table 5 and mapped in Fig. 6. The largest area estimated corresponds to forest gain ( $348.3 \pm 17.0$  ha), the second to stable forest area ( $242.4 \pm 53.6$  ha), the third to no forested area ( $188.4 \pm 33.6$  ha), and the fourth to forest loss ( $78.1 \pm 32.9$  ha). In terms of other land uses, agricultural land encompassed 114.25 ha in 1956, with most agricultural fields located in terraces (90.25 ha) and a minority (24 ha) located in lowlands (Fig. 7), mainly by the riverbank. In 2020 there were 31 ha of agricultural land, mainly comprised of olive trees followed by vineyards (Fig. 7), of which only 2.3 ha were not croplands in 1956. The reduction in agricultural land over six decades has been 73 %, with only 28.7 ha of stable cropland. Regarding the trails that connect different parts of the village



**Fig. 6.** Forest change map for the period 1956–2018, four classes are represented in the map: stable forest, forest gain, forest loss and no forest.

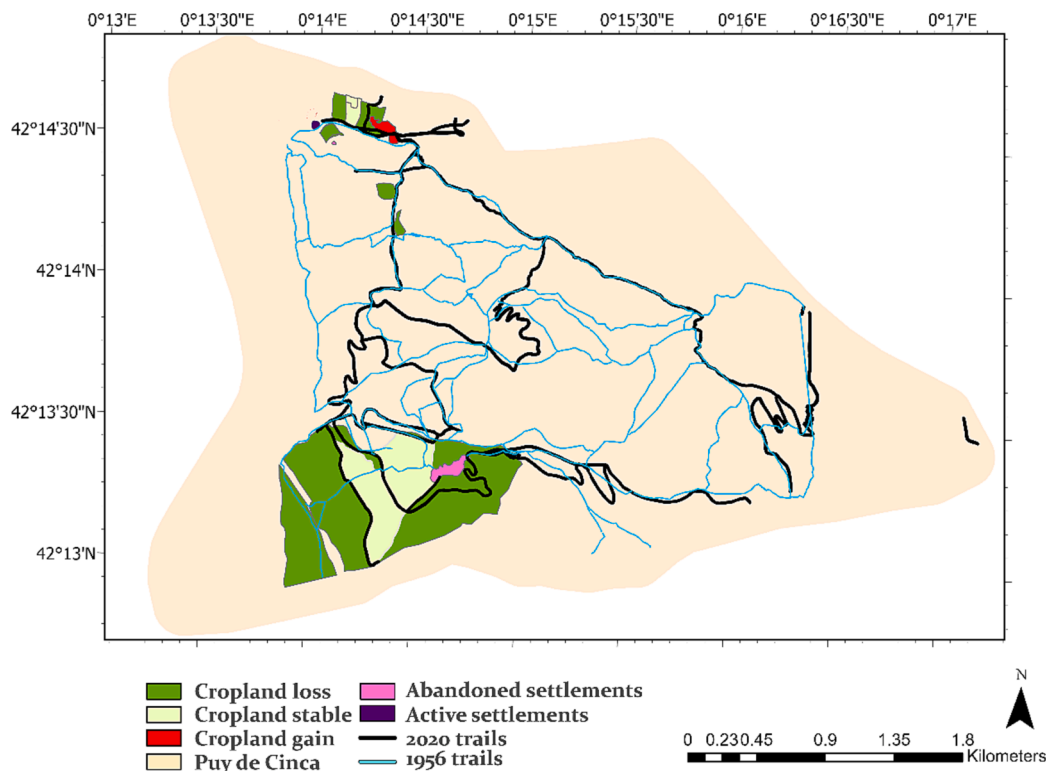


Fig. 7. Changes in cropland, settlements, and trails in Puy de Cinca from 1956 to 2018.

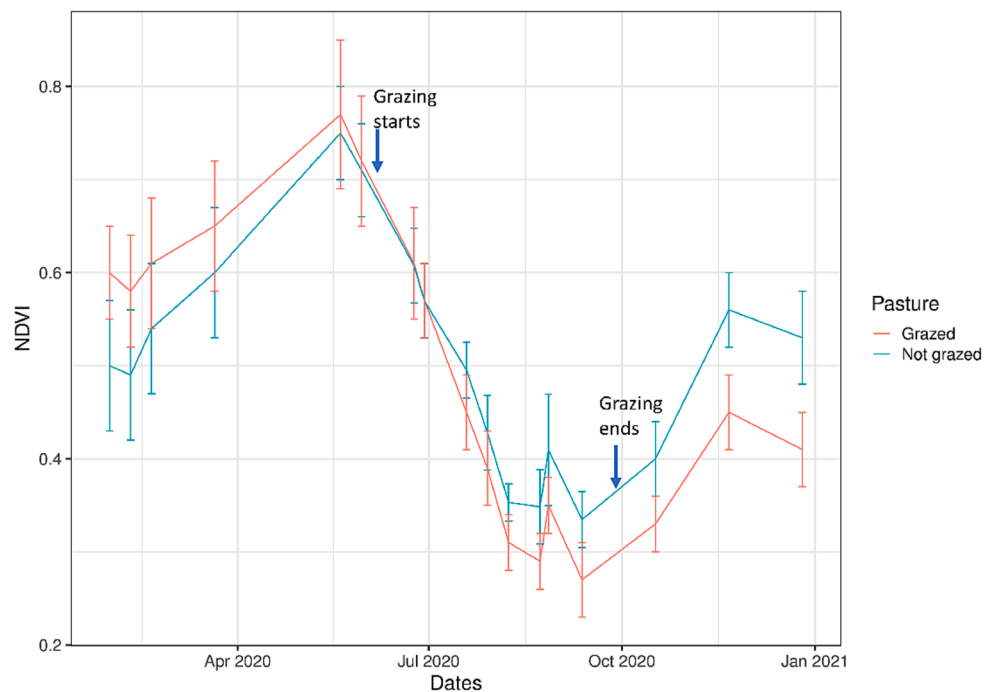


Fig. 8. Pattern of evolution of the NDVI from the control pasture (without grazing) and the donkey-grazed pasture.

territory, 88 km of trails were delineated in the aerial photographs from 1956, while the current (2022) trails only extend 33 km. This is a 62.5 % reduction in trail length over six decades (Fig. 7).

Before donkey grazing commenced, the NDVI values suggested that biomass and phytovolume in the grazed pasture was larger (Fig. 8). When the donkeys started grazing in June, the NDVI values for the 24th of June (2020) were similar in both the grazed and non-grazed (i.e.

control) pastures, with values of  $0.61 \pm 0.04$  and  $0.61 \pm 0.06$  respectively. On the 19th of July, the NDVI values suggested a change in the biomass of the pastures, likely due to the grazing activity, with a lower NDVI in the grazed, ( $0.45 \pm 0.02$ ), than in the non-grazed pasture ( $0.50 \pm 0.03$ ). This trend continued throughout the studied period, with the donkey-grazed pasture having lower NDVI values. In addition, environmental factors such as rainfall affect plant vigor, especially in



Mediterranean summers. Due to summer rain, an increase in biomass was observed between 23 August and 27 August. The NDVI value was nonetheless equally lower in the donkey-grazed pasture (Fig. 8). After grazing ended in the pasture the subsequent NDVI values were still lower in the previously grazed pasture than in the non-grazed pasture.

#### 4. Discussion

Regarding the first research question, “how much has the traditional landscape changed since the mid twentieth century (1956–2018) in Puy de Cinca?”, significant changes have been observed. This is, especially the case for forest changes, including an increase over that period of  $348.3 \pm 17.0$  ha in the forest cover, a 73 % reduction in agricultural land, and a 62.5 % reduction in trail length. These findings in Puy de Cinca align with a study on land use changes for a larger area elsewhere in Ribagorça county (the same county where Puy de Cinca is located), which showed a reduction of 50 % in agricultural land and a six-fold increase in forest area, including active reforestation (Sancho-Reinoso, 2013). In that study, the most extreme changes were observed in the mountainous areas rather than in the valleys of Ribagorça. The reduced forest cover that we observed from 1956 could be explained by several factors, such as the demographic pressure of the time and the need to obtain fuelwood and charcoal, which was also significant due to the abundance of oak forests (Solé i Sabarís, 1962). In that sense, the main drivers of forest expansion and agricultural decrease in Puy de Cinca might be linked to economic and water policies, plantation of fast-growing tree species, and rural depopulation. These changes are also observed with the number and extent of trails connecting different parts of the village with the former riverbanks, scrublands, forests, and agricultural lands that have disappeared. The trails, which were mainly used for human and animal transportation, have reverted to forest or scrublands due to the abandonment of agricultural fields and pastures. Among the several factors that have triggered the degradation of Puy de Cinca's traditional landscape, we argue that a key factor has been the destruction of the agricultural landscape along the riverbanks. The construction of the dam not only put these areas under water, but also changed the relationship to the main source driving the agrarian economy. Water management changed from a locally focused model agreed through communities of irrigators and their traditional ecological knowledge, to a more centralized management model (Bergua Amores, 2006).

Understanding landscape changes is a key factor to recover its functions. Indeed, traditional landscapes have been shaped by a prolonged human-environment interaction. Such a sustained relationship between people and the land, generates a profound knowledge of the territory (Oteros-Rozas et al., 2013; Reyes-García et al., 2014) regarding the distribution of species, the presence of pastures, the availability of water or the network of pathways for livestock and transportation. After the abandonment of villages and loss of bio-cultural knowledge, the reconstruction of these traditional landscapes is a challenge. In the case of Puy de Cinca, we have observed that remote sensing temporal images can contribute to rediscovering these land use changes. Furthermore, remotely sensed observations of changes in the landscape can guide policy implementation for the recovery of traditional landscapes and the withdrawal of damaging policies in the territories, as has been done elsewhere (González-Fernández et al., 2022; Osorno-Covarrubias et al., 2018). Capturing aerial or satellite scenes throughout time allows monitoring of temporal changes in Mediterranean landscapes (Lasanta and Vicente-Serrano, 2012; Zomeni et al., 2008). Moreover, remote sensing data allows monitoring of vegetation cover and understanding of agricultural lands and vegetation biophysical parameters such as biomass and other significant parameters for pasture management (von Keyserlingk et al., 2021).

Regarding the second research question, “does the grazing of Catalan donkeys reduce pasture biomass and phytovolume?”, differences in the NDVI, as a proxy for the photosynthetically active biomass and phytovolume of the pastures, were observed with Sentinel-2 satellite imagery

between grazed and non-grazed pastures. Indeed, the evolution of NDVI values throughout the grazing season suggested that the donkeys did impact the pasture's photosynthetically active biomass and phytovolume. However, NDVI values derived from Sentinel-2 spectral data showed some limitations when attempting to clearly distinguish between the grazed and non-grazed pastures. These limitations might be linked to the presence of trees and the lack of signal from the tree understory as well as the spatial resolution of the pixels (10 m), which might be limited to pick up the necessary fine details to accurately assess the effects of grazing. At the same time, the low-to-moderate-intensity of grazing might have made it challenging to sense changes in pasture biomass and phytovolume. The capacity of satellite imagery to monitor biophysical parameters related to biomass, such as the NDVI, has been effective at demonstrating the impact of large herbivores, other than donkeys, on pastures (Balata et al., 2022; Catorci et al., 2021; Navarro et al., 2020). Moreover, other articles have addressed donkey grazing as a forest management strategy and observed a reduction in flammable species in mountainous areas (Gartzia et al., 2009). Our results suggest the potential of Sentinel-2 data for monitoring low-to-moderate-intensity donkey grazing, with some limitations that could be improved by extending the number of pastures and seasons, as well as using broader spectral information, because high-resolution commercial satellites are not a current option for monitoring vegetation covers due to the existence of paywalls for accessing the images (Francini et al., 2020). The capacity to monitor the biomass with remote sensing data allows assessing the grazing impacts within the landscape. This entails significant implications in spatially controlling pastures biomass and flammable grasses for forest risk management in the case of Puy de Cinca and other equivalent sites.

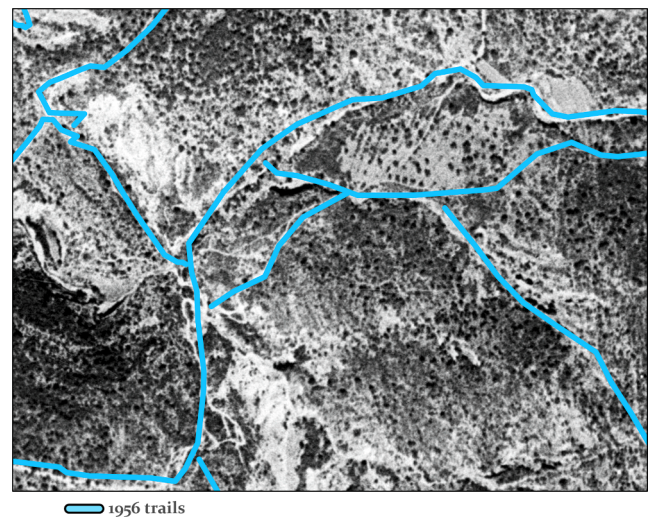
Therefore, monitoring pastures with satellite images, and particularly the exploitation of the temporal and spatial resolution of freely accessible Sentinel-2 imagery (Segarra et al., 2020), can assist with monitoring of biomass changes in donkey-grazed pastures, and also contribute to controlling the expansion or decline of forests and scrublands. At the same time, the recovery of trails within the traditional landscape should re-connect pastures and reestablish some agricultural land in terraces that require some animal traction due to the narrow nature of the fields. This has already been addressed in recent years by some farmers within the same province (Huesca) who are using horses in narrow vineyards and vegetable gardens with significant benefits regarding sustainable practices. Moreover, other alternatives for animal traction, such as the use of Catalan donkeys, could add value to the reestablishment of crops in these mountainous areas, and help continue this uniquely adapted breed.

Regarding the third research question, “is floral diversity in pastures with grazing donkeys greater than in abandoned pastures?”, an important result of the study is that the floristic composition of a donkey-grazed pasture showed greater diversity than an abandoned pasture without donkey activity. This aligns with other studies reporting higher plant diversity in low-intensity pastures with grazing by large herbivores such as horses (Köhler et al., 2016; Moineau et al., 2020), or a combination of cattle with occasional donkey grazing (Török et al., 2016), in contrast to non-grazed or intensively grazed pastures. Such studies, nonetheless, were carried out in relatively wet central European climates, which can be classified as possessing productive grasslands. Some authors argue that grazing by large herbivores enhances plant diversity, particularly in productive environments (Catorci et al., 2012). However, in poorer soils, grazing may negatively affect plant species richness, where soil resources limit the regrowth of plant species (Proulx and Mazumder, 1998). In the current study, our findings for a Mediterranean mountainous landscape in northeast Spain, suggest that donkey grazing can also contribute to enhance biodiversity by recovering plant diversity and improving forage value. Indeed, two of the three most abundant plants found in the grazed pasture are from the *Medicago* genus, which has high potential for seed and forage production (Porqueddu and González, 2006). Moreover, the second most abundant

species was *Cynodon dactylon*, which also has a high forage value and is adapted to Mediterranean environments. This species normally appears on well-formed soils where the impact of summer droughts can be counteracted by the development of a deeply penetrating root system reaching the water table (Galiano, 1985). In contrast, in the abandoned pasture, the most abundant species was *Brachypodium phoenicoides*, belonging to the Gramineae family, and it has a relatively low nutritional value compared to other species growing in Mediterranean pastures (Viano et al., 1996). The second most abundant species in the abandoned pasture was *Knautia arvensis*, which has also been described elsewhere in non-grazed pastures due to its ability to spread by clonal growth, resulting in higher competitive advantage in the plant community (Stránská, 2004). Comparing the presence of the *Plantago* genus between the abandoned and grazed pasture, the dominance of *Plantago lanceolata* was also greater in the grazed pasture, and it is known for being sensitive to shading, so a reduction in plant phytovolume following grazing by donkeys likely contributed to its development. Moreover, it produces nutritional forage for grazing animals (Stewart, 1996).

Donkey breeds are scarce in the Iberian Peninsula, but they exhibit more specialized distributions than other livestock and are adapted to a wide range of environmental heterogeneity (Alonso et al., 2022). For this reason, Catalan donkeys may contribute towards the recovery of traditional landscapes through their adaptation to the specific environment of Mediterranean mountains, increasing floral diversity in pastures, and reducing the tree understory while improving forest management with implications for fire risk reduction. Nevertheless, there are some limitations to this study, first regarding the number of Catalan donkeys available, and second, the spatial analysis that could be achieved of their effect on pastures and species diversity, which was confined to the northern part of the village and to a few pastures. Moreover, donkeys exhibit a preference for grassy species and do not graze significantly on scrubs, unlike other animals. Consequently, donkey grazing with other animals (such as goats) could function synergistically (Bartolomé et al., 2020) for the recovery of traditional mountain landscapes. Furthermore, combining several grazing animal species would increase the capacity of such environmental management strategies because the current donkey population is too small to cover the entirety of Puy de Cinca. Additionally, due to their preference for specific grass types and the resulting uneven grazing (Gartzia et al., 2009), the geolocation of donkeys and the use of satellite images to monitor their free grazing activity could help determine biomass accumulation and then coordinate management of other animals across the landscape.

Besides grazing, the findings of this study suggest that a general political change regarding mountainous rural areas could be addressed. The most important changes in these landscapes happened more than half a century ago in a centralized dictatorial Spain, which since then has continued to shape most rural areas with characteristic depopulation and poor income opportunities, as well as the loss of bio-cultural knowledge. Furthermore, in more recent decades, the few abandoned mountain villages recovered by people have continued to suffer political ambiguity and even adversities; in fact, after the restauration of democracy, Spanish politics has neglected, and even negated, mountain identities and ways of living. The main initiatives regarding the conservation of the Catalan donkey landrace and the recovery of abandoned villages such as Puy de Cinca is based on grass-roots activism from the few people remaining in the area who wish to preserve and recreate traditional approaches to living in the landscape and manage the mountain territory. In this sense, several workshops regarding the use of Catalan Donkey have been carried out in Puy de Cinca, and the use of satellite data to monitor and preserve the environment have been discussed with locals through a dialogue of voices. We see this as a critical application of remote sensing, as described by Bennett et al. (2022), which simultaneously exposes injustices, reflexively engages with situated local knowledge, and expands access to geospatial technologies.



**Fig. A1.** An example of the delineation of trails in a scene from 1956 is shown. Only long trails connecting different points of the studied areas were delineated, and shorter paths (as observed in the Figure) were not delineated as they did not represent the general network of connection in the landscape. Narrow long path-like shapes were considered trails and delineated visually.

The recovery of traditional mountain landscapes will benefit from a dialog of locals with remote sensing scholars.

## 5. Conclusion

We conclude that, at low-to-moderate-intensity grazing, Catalan donkeys help to reduce phytovolume and biomass and enhance plant diversity in pastures. The recovery of the endangered donkey landrace could play a relevant role in the environmental management of the abandoned mountainous Mediterranean landscape that we have studied in northeast Spain (Puy de Cinca). The landscape changes we quantified, including an increase in forest cover ( $348.3 \pm 17.0$  ha), a 73 % loss of croplands and a 62.5 % reduction in the trails interconnecting the territory, illustrates a loss of the landscape's original functions. We have proven, with some limitations, the suitability of Sentinel-2 images to detect reductions in biomass due to donkey grazing activity. The study may also be useful for other cases where the degradation of traditional landscapes in Mediterranean mountainous areas is significant and the recovery of animal landraces could help restore these landscapes and provide continuing environmental and ecosystem services.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

All data used is publicly available

## Acknowledgments

We acknowledge the support of the PID2019-106650RB-C21 project from the Ministerio de Ciencia e Innovación, Spain. J.S. is a recipient of a FPI doctoral fellowship from the same institution (grant: PRE2020-091907). J.L.A. acknowledges support from the Institutio Catalana de Recerca i Estudis Avançats (ICREA), Generalitat de Catalunya, Spain. J. F.M. acknowledges the support from Miguel Ángel Lapuyade and Belén Álvarez from the Laboratorio de Ruralización de la Aldea de Puy de Cinca, and the support and funds from the City Council of Secastilla and

from the Diputación de Huesca (329624M), Aragón, Spain.

## Appendix

See Fig. A1.

## References

- Albuquerque, U.P., de Lucena, R.F.P., Leal Alencar, N., 2010. Métodos e técnicas para coleta de dados etnobiológicos. In: Métodos e Técnicas Na Pesquisa Etnobiológica e Etnoecológica.
- Alonso, E.V., Morales, I., Antonio, C., Sal, G., 2022. The landscapes of livestock diversity: grazing local breeds as a proxy for domesticated species adaptation to the environment. *Landsc. Ecol.* 37, 1035–1048. <https://doi.org/10.1007/s10980-022-01429-5>.
- Álvarez-Martínez, J., Gómez-Villar, A., Lasanta, T., 2016. The use of goats grazing to restore pastures invaded by shrubs and avoid desertification: a preliminary case study in the Spanish Cantabrian Mountains. *L. Degrad. Dev.* 27, 3–13. <https://doi.org/10.1002/ldr.2230>.
- Azon-Pardo, R., 1998. Aldea Puy Cinca: recuperación de las razas de los Pirineos. *Mundo Ganad.* 18–20.
- Balata, D., Gama, I., Domingos, T., Proença, V., 2022. Using Satellite NDVI Time-Series to Monitor Grazing Effects on Vegetation Productivity and Phenology in Heterogeneous Mediterranean Forests. *Remote Sens.* 14, 2322. <https://doi.org/10.3390/rs14102322>.
- Bartolomé, J., Miró, J., Panadès, X., Broncano, M.J., Plaixats, J., Rigau, T., Milán, M.J., Baraza, E., 2020. Preference by donkeys and goats among five Mediterranean forest species: implications for reducing fire hazard. *Animals* 10, 1–9. <https://doi.org/10.3390/ani10081302>.
- Bennett, M.M., Chen, J.K., Alvarez León, L.F., Gleason, C.J., 2022. The politics of pixels: a review and agenda for critical remote sensing. *Prog. Hum. Geogr.* 46, 729–752. <https://doi.org/10.1177/03091325221074691>.
- Bergua Amores, J.A., 2006. Conflictos y liberación de energía Los movimientos antipantano y la reinención de los Pirineos. *Ripacurtia* 4, 115–140.
- Blondel, J., Aronson, J., 1995. Biodiversity and Ecosystem Function in the Mediterranean Basin: Human and Non-Human Determinants. pp. 43–119. doi: 10.1007/978-3-642-78881-9\_2.
- Buil, G.M., 1993. Perder el pueblo (Antropología aplicada y política hidráulica). *Revista de antropología social* 2, 185–238.
- Casasús, I., Bernués, A., Flores, N., Sanz, A., Valderrábano, J., Revilla, R., 2004. Livestock farming systems and conservation of Spanish Mediterranean mountain areas: the case of the “Sierra de Guara” Natural Park. 2. Effects of grazing on vegetation. *Cah. Options Méditerranéennes* 62, 199–202.
- Catorci, A., Gatti, R., Cesaretti, S., 2012. Effect of sheep and horse grazing on species and functional composition of sub-Mediterranean grasslands. *Appl. Veg. Sci.* 15, 459–469. <https://doi.org/10.1111/j.1654-109X.2012.01197.x>.
- Catorci, A., Lulli, R., Malatesta, L., Tavaloni, M., Tardella, F.M., 2021. How the interplay between management and interannual climatic variability influences the NDVI variation in a sub-Mediterranean pastoral system: Insight into sustainable grassland use under climate change. *Agr. Ecosyst. Environ.* 314, 107372. <https://doi.org/10.1016/j.agee.2021.107372>.
- Collantes, F., Pinilla, V., 2012. Peaceful Surrender: The Depopulation of Rural Spain in the Twentieth Century. Cambridge Scholars Publishing, Cambridge.
- Donald, P.F., Green, R.E., Heath, M.F., 2001. Agricultural intensification and the collapse of Europe's farmland bird populations. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 268, 25–29. <https://doi.org/10.1098/rspb.2000.1325>.
- Duguy, B., Alloza, J.A., Röder, A., Vallejo, R., Pastor, F., 2007. Modelling the effects of landscape fuel treatments on fire growth and behaviour in a Mediterranean landscape (eastern Spain). *Int. J. Wildl. Fire* 16, 619. <https://doi.org/10.1071/WF06101>.
- Francini, S., McRoberts, R.E., Giannetti, F., Mencucci, M., Marchetti, M., Scarascia Mugnozza, G., Chirici, G., 2020. Near-real time forest change detection using PlanetScope imagery. *Eur. J. Remote Sens.* 53, 233–244. <https://doi.org/10.1080/22797254.2020.1806734>.
- Galiano, E.F., 1985. The small-scale pattern of *Cynodon dactylon* in Mediterranean pastures. *Vegetatio* 63. <https://doi.org/10.1007/BF00044062>.
- García, M.B., Gómez, D., 2007. Flora del pirineo aragonés. Patrones espaciales de biodiversidad y su relevancia para la conservación. Pirineos.
- García-Ruiz, J.M., 1977. Grandes embalses y desorganización del espacio: El ejemplo del Alto Aragón. *Cuad. Investig. Geogr. e Hist.* 3, 31–46.
- García-Ruiz, J.M., Lasanta, T., Ruiz-Flano, P., Ortigosa, L., White, S., González, C., Martí, C., 1996. Land-use changes and sustainable development in mountain areas: a case study in the Spanish Pyrenees. *Landsc. Ecol.* 11, 267–277. <https://doi.org/10.1007/BF02059854>.
- Garrabou i Segura, R., 1985. Historia agraria de la España contemporánea. Crítica, Barcelona.
- Gartzia, M., Fillat, F., National, S., Aguirre, J., 2009. Utilización del ganado asnal para la prevención de incendios en masas forestales del Pirineo oscense. *Sist. Y Recur. Silvopastorales* 663–669.
- Geri, F., Amici, V., Rocchini, D., 2010. Human activity impact on the heterogeneity of a Mediterranean landscape. *Appl. Geogr.* 30, 370–379. <https://doi.org/10.1016/j.apgeog.2009.10.006>.
- Gomez Mendoza, J., Mata Olmo, R., 1992. Acciones forestales públicas desde 1940. *Agríc. y Soc.* 15–64.
- González-Fernández, A., Segarra, J., Sunny, A., Couturier, S., 2022. Forest cover loss in the Nevado de Toluca volcano protected area (Mexico) after the change to a less restrictive category in 2013. *Biodivers. Conserv.* 31. <https://doi.org/10.1007/s10531-022-02368-y>.
- Grasa, C.T., 2017. Consecuencias sociales de la acción repobladora del Patrimonio Forestal del Estado en la provincia de Huesca (1950-1971). In: La Restauración Forestal de España: 75 Años de Una Ilusión. Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente, pp. 282–311.
- Greider, T., Garkovich, L., 2010. Landscapes: The Social Construction of Nature and the Environment. *Rural. Sociol.* 59, 1–24. <https://doi.org/10.1111/j.1549-0831.1994.tb00519.x>.
- Guiot, J., Cramer, W., 2016. Climate change: The 2015 Paris Agreement thresholds and Mediterranean basin ecosystems. *Science* 354, 465–468. <https://doi.org/10.1126/science.aah5015>.
- Gulfás, J., Mairata, A., Frontera, J., Janer, I., Jaume, J., Cifre, J., 2016. Firebreak maintenance with equines in Serra de Tramuntana mountains (Mallorca, UNESCO world heritage). *Mt. Pastures Livest. Farming Facing Uncertain. Environ. Tech. Socio-Economic Challenges*. 336, 333–336.
- Gutierrez, J.P., Marmi, J., Goyache, F., Jordana, J., 2005. Pedigree information reveals moderate to high levels of inbreeding and a weak population structure in the endangered Catalanian donkey breed. *J. Anim. Breed. Genet.* 122, 378–386. <https://doi.org/10.1111/j.1439-0388.2005.00546.x>.
- Gutierrez-Peña, R., Mena, Y., Ruiz, F.A., Delgado-Pertinéz, M., 2016. Strengths and weaknesses of traditional feeding management of dairy goat farms in mountain areas. *Agroecol. Sustain. Food Syst.* 40, 736–756. <https://doi.org/10.1080/21683565.2016.1178202>.
- Herlin, A., Brunberg, E., Hultgren, J., Högborg, N., Rydberg, A., Skarin, A., 2021. Animal Welfare Implications of Digital Tools for Monitoring and Management of Cattle and Sheep on Pasture. *Animals* 11, 829. <https://doi.org/10.3390/ani11030829>.
- Hill, M.O., Gauch, H.G., 1980. Detrended Correspondence Analysis: An Improved Ordination Technique. In: Classification and Ordination. Springer Netherlands, Dordrecht, pp. 47–58. doi: 10.1007/978-94-009-9197-2\_7.
- Iriarte-Goñi, I., 2008. El consum de fusta en Espanya, (1900–2000). Aspectes econòmics i aproximació als efectes ambientals. *Recerques* 57, 49–70.
- Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. *Bull. la Société vaudoise des Sci. Nat.* 223–270.
- Jiménez-Olivencia, Y., Ibáñez-Jiménez, Á., Porcel-Rodríguez, L., Zimmerer, K., 2021. Land use change dynamics in Euro-mediterranean mountain regions: driving forces and consequences for the landscape. *Land Use Policy* 109. <https://doi.org/10.1016/j.landusepol.2021.105721>.
- Jordana i Vidal, J., Folch Lopez, P., 1998. La raza asnal Catalana programa de conservación y mejora de una población en peligro de extinción. *Arch. Zootec.* 47, 404–409.
- Jordana, J., 2010. Orígens i variabilitat de les races asinines segons les dades genètiques. El cas de la península Ibèrica i la recuperació del burro català. In: La Recuperació Del Burro Català. Aspectes Culturals i Biològics. Quaderns, pp. 57–75.
- Köhler, M., Hiller, G., Tischew, S., 2016. Year-round horse grazing supports typical vascular plant species, orchids and rare bird communities in a dry calcareous grassland. *Agr. Ecosyst. Environ.* 234, 48–57. <https://doi.org/10.1016/j.agee.2016.03.020>.
- Kugler, W., Grunfelder, H.P., Broxham, E., 2007. Donkey Breeds in Europe Inventory, Description, Need for Action, Conservation.
- Lasanta, T., Vicente-Serrano, S.M., 2012. Complex land cover change processes in semiarid Mediterranean regions: An approach using Landsat images in northeast Spain. *Remote Sens. Environ.* 124, 1–14. <https://doi.org/10.1016/j.rse.2012.04.023>.
- Malek, Z., Verbarg, P.H., Reijndersdorff, I., Bondeau, A., Cramer, W., 2018. Global change effects on land management in the Mediterranean region. *Glob. Environ. Chang.* 50, 238–254. doi: 10.1016/j.gloenvcha.2018.04.007.
- Maselli, F., Gilabert, M.A., Conese, C., 1998. Integration of High and Low Resolution NDVI Data for Monitoring Vegetation in Mediterranean Environments. *Remote Sens. Environ.* 63, 208–218. [https://doi.org/10.1016/S0034-4257\(97\)00131-4](https://doi.org/10.1016/S0034-4257(97)00131-4).
- Moinardeau, C., Mesléard, F., Ramone, H., Dutoit, T., 2020. Extensive horse grazing improves grassland vegetation diversity, seed bank and forage quality of artificial bankmounds (Rhône River - southern France). *J. Nat. Conserv.* 56, 125865. <https://doi.org/10.1016/j.jnc.2020.125865>.
- Navarro, L.M., Pereira, H.M., 2015. Rewilding Abandoned Landscapes in Europe. In: Rewilding European Landscapes. Springer International Publishing, Cham, pp. 3–23. doi: 10.1007/978-3-319-12039-3\_1.
- Navarro, C.J., Izquierdo, A.E., Araújo, E., Foguet, J., Grau, H.R., 2020. Rewilding of large herbivore communities in high elevation Puna: geographic segregation and no evidence of positive effects on peatland productivity. *Reg. Environ. Chang.* 20, 112. <https://doi.org/10.1007/s10113-020-01704-8>.
- Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Simpson, G., Solymos, P., Stevens, M., Wagner, H., 2015. Vegan: community ecology. R Package version 2.2-1.
- Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* 148, 42–57. <https://doi.org/10.1016/j.rse.2014.02.015>.
- Osorno-Covarrubias, F., Couturier, S., Piceno Hernández, M., 2018. Measuring from Space the Efficiency of Local Forest Management: The Successful Case of the Indigenous Community of Cherán, Mexico. *IOP Conf. Ser. Earth Environ. Sci.* 151, 012010. <https://doi.org/10.1088/1755-1315/151/1/012010>.
- Oteros-Rozas, E., Ontillera-Sánchez, R., Sanosa, P., Gómez-Baggethun, E., Reyes-García, V., González, J.A., 2013. Traditional ecological knowledge among



- transhumant pastoralists in Mediterranean Spain. *Ecol. Soc.* 18 <https://doi.org/10.5751/ES-05597-180333>.
- Papanastasis, V.P., Kyriakakis, S., Kazakis, G., 2002. Plant diversity in relation to overgrazing and burning in mountain mediterranean ecosystems. *J. Mediterr. Ecol.* 3, 53–63.
- Pielou, E.C., 1966. The measurement of diversity in different types of biological collections. *J. Theor. Biol.* 13, 131–144. [https://doi.org/10.1016/0022-5193\(66\)90013-0](https://doi.org/10.1016/0022-5193(66)90013-0).
- Porqueddu, C., González, F., 2006. Role and potential of annual pasture legumes in mediterranean farming systems. *Pastos* 36.
- Proulx, M., Mazumder, A., 1998. Reversal of grazing impact on plant species richness in nutrient-poor vs. nutrient-rich ecosystems. *Ecology* 79, 2581–2592.
- Reyes-García, V., Aceituno-Mata, L., Calvet-Mir, L., Garnatje, T., Gómez-Baggethun, E., Lastra, J.J., Ontillera, R., Parada, M., Rigat, M., Vallès, J., Vila, S., Pardo-de-Santayana, M., 2014. Resilience of traditional knowledge systems: The case of agricultural knowledge in home gardens of the Iberian Peninsula. *Glob. Environ. Chang.* 24, 223–231. <https://doi.org/10.1016/j.gloenvcha.2013.11.022>.
- Rouse Jr., J.W., Haas, R., Schell, J., Deering, D., 1974. Monitoring vegetation systems in the great plains with erts. *NASA Spec. Publ.* 351.
- Ruiz-Flaño, P., 1993. *Procesos de erosión en campos abandonados del Pirineo*. Geoforma Ediciones 191.
- Sancho-Reinoso, A., 2013. Land abandonment and the dynamics of agricultural landscapes in mediterranean mountain environments: the case of Ribagorza (Spanish pyrenees). *Erdkunde* 67, 289–308. <https://www.jstor.org/stable/23595373>.
- Schwabe, A., Süß, K., Storm, C., 2013. What are the long-term effects of livestock grazing insteppic sandy grassland with high conservation value? Results from a 12-year field study. *Tuexenia* 33, 189–212.
- Segarra, J., Buchailot, M.L., Araus, J.L., Kefauver, S.C., 2020. Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications. *Agronomy* 1–18.
- Shannon, C.E., Weaver, W., 1949. *The Mathematical Theory of Communication*. University of Illinois Press.
- Solé i Sabarís, L., 1962. *Geografia de Catalunya*. Aedos.
- Stewart, A.V., 1996. Plantain (*Plantago lanceolata*) - a potential pasture species. *Proc. New Zeal. Grassl. Assoc.* 77–86. [10.33584/jnzg.1996.58.2221](https://doi.org/10.33584/jnzg.1996.58.2221).
- Stránská, M., 2004. Successional dynamics of *Cynosurus* pasture after abandonment in Podkrkonosi. *Plant Soil Environ.* 50. [10.17221/4045-pse](https://doi.org/10.17221/4045-pse).
- Timmermann, A., Damgaard, C., Strandberg, M.T., Svenning, J.-C., 2015. Pervasive early 21st-century vegetation changes across Danish semi-natural ecosystems: more losers than winners and a shift towards competitive, tall-growing species. *J. Appl. Ecol.* 52, 21–30. <https://doi.org/10.1111/1365-2664.12374>.
- Török, P., Valkó, O., Deák, B., Kelemen, A., Tóth, E., Tóthmérész, B., 2016. Managing for species composition or diversity? Pastoral and free grazing systems in alkali steppes. *Agr. Ecosyst. Environ.* 234, 23–30. <https://doi.org/10.1016/j.agee.2016.01.010>.
- Verdú, J.R., Crespo, M.B., Galante, E., 2000. Conservation strategy of a nature reserve in Mediterranean ecosystems: The effects of protection from grazing on biodiversity. *Biodivers. Conserv.* 9, 1707–1721. <https://doi.org/10.1023/A:1026506725251>.
- Viano, J., Masotti, V., Gaydou, E.M., Giraud, M., Bourreil, P.J.L., Ghiglione, C., 1996. Composition of Liliifloreae from Mediterranean Pastures. *J. Agric. Food Chem.* 44 <https://doi.org/10.1021/jf960111a>.
- von Keyserlingk, J., de Hoop, M., Mayor, A.G., Dekker, S.C., Rietkerk, M., Foerster, S., 2021. Resilience of vegetation to drought: Studying the effect of grazing in a Mediterranean rangeland using satellite time series. *Remote Sens. Environ.* 255, 112270 <https://doi.org/10.1016/j.rse.2020.112270>.
- Zamora, J., Verdú, J.R., Galante, E., 2007. Species richness in Mediterranean agroecosystems: Spatial and temporal analysis for biodiversity conservation. *Biol. Conserv.* 134, 113–121. <https://doi.org/10.1016/j.biocon.2006.08.011>.
- Zomeni, M., Tzanopoulos, J., Pantis, J.D., 2008. Historical analysis of landscape change using remote sensing techniques: an explanatory tool for agricultural transformation in Greek rural areas. *Landsc. Urban Plan.* 86, 38–46. <https://doi.org/10.1016/j.landurbplan.2007.12.006>.







