

Verificació de pronòstics de temperatura i aplicació de models regressius localitzats

Treball Final del Màster en Meteorologia

Facultat de Física de la
Universitat de Barcelona

Alumne: Màxim Colls Xatart

Tutor: Joan Bech Rustullet

Juny del 2011

Agraïments

M'agradaria agrair al Servei Meteorològic de Catalunya per haver facilitat les sèries de dades utilitzades en aquest treball i al meu tutor Joan Bech per haver-me guiat i donat suport sempre que l'he necessitat.

També als companys de feina de l'LCFIB pel suport moral i tècnic que m'han donat en qualsevol problema que he pogut tenir.

Finalment als amics, companys de màster, Núria i música per haver-me acompanyat durant el temps que he dedicat a aquest treball. ☺

RESUM

Aquest treball pretén analitzar el pronòstic de temperatura en superfície (2 metres) dels models WRF i MM5 al territori de Catalunya i proposar una millora d'aquests pronòstics mitjançant models regressius localitzats.

Inicialment s'han caracteritzat les sèries de temperatura en superfície de 64 estacions meteorològiques. A més, s'ha realitzat un anàlisi més profund de 4 estacions de referència escollides de manera que es representés la varietat geogràfica i climàtica del territori estudiat.

S'ha iniciat un procés de validació dels pronòstics de temperatura per part dels models WRF i MM5. En el conjunt dels 64 punts repartits pel territori, s'han analitzat els resultats dels models i s'han comparat amb la persistència.

Finalment s'ha dissenyat un conjunt de models regressius localitzats per tal de millorar el pronòstic de la sortida dels models. Aquests models regressius aconseguixen una millora molt significativa sobretot en les estacions del Pirineu on la reducció de l'RMSE arriba al 60 per cent.

ABSTRACT

This work presents an analysis of air surface temperature (2 meters) forecasts from WRF and MM5 models in the area of Catalonia. Then, an improvement of those forecasts using local regression models is proposed.

At the beginning, 64 temperature series from automatic weather stations have been characterized. Furthermore, 4 representative series have been specially analyzed in order to study the characteristics of the territory.

A process of validation have been started to know the results of temperature forecasts from WRF and MM5 models. The results have been analyzed and compared with the persistence.

Finally, localized regression models have been designed to improve the numerical weather predictions outputs. Those local models obtain remarkable results, especially in the area of the Pyrenees where the RMSE has been reduced about 60 percent.

ÍNDIX

RESUM	5
ABSTRACT	5
1 INTRODUCCIÓ	9
1.1 Objectius	9
1.2 Motivacions	9
1.3 Treball relacionat	10
1.4 Eines utilitzades	10
1.5 Models de predicció numèrica (MPN)	11
1.5.1 MM5	11
1.5.2 WRF	11
1.6 Xarxa d'estacions meteorològiques automàtiques (XEMA)	12
1.7 Característiques climàtiques del territori	12
1.8 Estacions Meteorològiques de referència	13
1.8.1 Estació de Salòria (ZB)	13
1.8.2 Estació del Raval (X4)	14
1.8.3 Estació de Seròs (VL)	14
1.8.4 Estació de Sant Pere Pescador (U2)	14
2 MINERIA DE DADES DE LES SÈRIES TEMPORALS DE TEMPERATURA	15
2.1 Descripció de les sèries	15
2.1.1 Sèries temporals de les EMA (Observacions)	15
2.1.2 Sèries temporals dels models (Pronòstics)	15
2.2 Caracterització de les sèries d'EMA	15
2.2.1 Anàlisi Territorial de les sèries	16
2.2.2 Tests d'autocorrelació	19
3 VERIFICACIÓ DELS PRONÒSTICS DELS MODELS	22
3.1 Indicadors de validesa dels pronòstics	22
3.1.1 Error Mig (ME)	22
3.1.2 Error Mig Absolut (MAE)	22

3.1.3	Error Quadràtic Mig (RMSE)	22
3.1.4	Desviació estàndard (SD, σ)	23
3.1.5	Correlació (CORR, ρ)	23
3.2	Qualitat dels pronòstics dels models	23
3.3	Errors en l'estimació de l'altura de les estacions	26
3.4	Anàlisi de la persistència	28
4	APLICACIÓ DE MODELS REGRESSIUS LOCALITZATS	31
4.1	Definició de model regressiu	31
4.1.1	Model autoregressiu	31
4.2	Disseny del model lineal regressiu	31
4.2.1	Correcció de l'error en l'estimació de l'altura de les EMA	32
4.2.2	Model lineal utilitzant els MPN i la persistència	33
4.3	Avaluació global dels resultats	33
5	CONCLUSIONS	37
5.1	Resultats obtinguts	37
5.2	Treball futur	37
	REFERÈNCIES	38
	TAULA DE CONTINGUTS	40
	ANNEX I – TAULA D'EMA	42
	ANNEX II – DIAGRAMES DE TAYLOR	44
	ANNEX III – VALOR DELS COEFICIENTS BETA	47

1 INTRODUCCIÓ

1.1 Objectius

Aquest treball està dividit en tres fases cada una de les quals té objectius individuals a assolir:

- Per una banda, es disposa de 64 sèries de dades observacionals i dels models WRF i MM5 repartides en el territori de Catalunya. Cadascuna de les sèries es correspon a la posició d'una estació meteorològica automàtica (EMA). L'objectiu principal de la primera fase és el de realitzar un procés de mineria de dades del camp de temperatura en superfície (a 2 metres d'alçada) de les sèries observacionals.
- La característica més important de les sèries utilitzades en aquest treball és que les sortides dels models de predicció numèrica (MPN) WRF i MM5 estan interpolades a les posicions de les EMA. Aquesta característica permet realitzar un procés de verificació molt acurat per cadascun dels models. A partir de diversos mètodes de verificació (Wilks, 2006; EUMETCAL, 2011; Taylor, 2001) es qualificaran els resultats i s'analitzaran les característiques dels pronòstics per tot el territori.
- Una vegada coneguda la qualitat dels pronòstics de temperatura dels models WRF i MM5, l'objectiu següent és el de millorar aquests pronòstics mitjançant la correcció d'errors sistemàtics i l'aplicació de models regressius localitzats. Un cop dissenyats els models, s'avaluaran les millores obtingudes respecte els pronòstics directes dels MPN.

1.2 Motivacions

El pronòstic de temperatura en superfície (2 metres) té aplicacions directes en molts caps diferents com poden ser l'agricultura, la seguretat viària o l'oci. En l'agricultura, la temperatura en superfície és una mesura clau per estimar els fluxos de calor i humitat superfície - atmosfera utilitzats en la simulació i en el control dels cultius. A més a més, la previsió de temperatures extremes és molt important per tal de poder actuar preventivament i evitar situacions que provocarien greus pèrdues en les collites.

Per altra banda, el camp de temperatura en superfície serveix d'entrada en molts altres models com ara els de predicció de la temperatura superficial de les carreteres. Aquests són utilitzats per detectar situacions de congelament i així evitar situacions de perill greus (Bogren et al, 2002). Un altre cas és el dels models que pronostiquen el tipus de precipitació a partir de camps superficials (Ortiz, 2009; Gjertsen et al, 2005). A més d'utilitzar-se per evitar situacions de risc, també poden ajudar en la optimització de la producció de neu artificial en les estacions d'esquí.

1.3 Treball relacionat

Les sèries de dades utilitzades en aquest treball provenen d'un projecte del Servei Meteorològic de Catalunya (SMC) que tenia com a objectiu analitzar i valorar els pronòstics dels MPN en el territori de Catalunya. Aquestes sèries s'han utilitzat per aprofundir en l'anàlisi del pronòstic de la temperatura en superfície (2 metres).

La verificació del pronòstic de temperatura no és un camp de recerca tant extens com ho pot ser el del pronòstic de vent o precipitació, no obstant, s'han trobat alguns articles amb objectius similars com ara Valeriano (2001). El treball citat inicia un procés d'avaluació de les previsions de temperatura en superfície del model MM5 a l'àrea del Perú. Alguns dels problemes que s'esmenten a l'article estan relacionats amb la poca resolució dels models digitals de terreny emprats a l'hora de modelar les zones muntanyoses com és el cas dels Andes a l'àrea del Perú. Aquest problema també s'ha tractat en aquest treball en el cas de l'àrea dels Pirineus.

La utilització de models estadístics per millorar la sortida dels models de predicció numèrica (MPN) ha estat una tècnica emprada des dels inicis de la utilització operativa d'aquests models (William et al, 1974). Especialment, aquesta tècnica s'ha utilitzat per localitzar els pronòstics (*downscaling*). Per altra banda, els models autoregressius permeten la millora dels pronòstics mitjançant un altre tipus d'enfocament: l'estudi de l'evolució de la pròpia sèrie. Un exemple d'aquesta tècnica es pot trobar a l'article de Ríos-Moreno (2007) on mitjançant diversos dissenys de models autoregressius s'aconsegueix una millora del pronòstic de temperatura a curt termini.

1.4 Eines utilitzades

Per l'explotació de les dades i la visualització dels resultats s'ha utilitzat el paquet estadístic R.

R és un llenguatge de programació i un entorn de desenvolupament de software per a l'obtenció de càlculs i gràfics estadístics. El llenguatge R ha esdevingut un estàndard entre els estadístics per l'anàlisi de dades. El codi font de l'R està disponible de forma gratuïta sota la Llicència Pública General GNU. R compta amb una gran comunitat a la xarxa que fan que hi hagi una gran quantitat d'informació i ajuda a l'abast. L'entorn de desenvolupadors en R publica manuals de suport, guies amb exemples a més d'*scripts* (rutines escrites en R per realitzar tasques determinades) d'usuaris que decideixen compartir el seu treball seguint d'aquesta manera la filosofia de l'entorn.

Alguns dels manuals que s'han utilitzat es llisten a les referències (Zucchini et al, 2006; Teetor, 2011). El software es pot descarregar des de la web del projecte (R-Project, 1993-2011 <http://www.r-project.org/>).



Figura 1 - Logotip del paquet estadístic R

1.5 Models de predicció numèrica (MPN)

1.5.1 MM5

MM5 és la quinta generació del model de mesoscala desenvolupat per la *Pennsylvania State University* (PSU) juntament amb el *National Center for Atmospheric Research* (NCAR).

És un model d'àrea limitada dissenyat per simular circulacions atmosfèriques de mesoscala. Aquesta inclou fenòmens que tenen lloc a escales espacials que van des d'uns quants quilòmetres fins a uns quants centenars (com ara tempestes, brises o sistemes frontals). També permet aniuaments per tal de refinar l'àrea de predicció. Utilitza coordenades sigma i es pot configurar com a model hidrostàtic o no-hidrostàtic.

El model i altres eines relacionades es poden descarregar des de la web de la comunitat (<http://www.mmm.ucar.edu/mm5/>).

1.5.2 WRF

El model WRF (*Weather Research and Forecasting*) és un model de mesoscala de nova generació dissenyat conjuntament per la *National Oceanic and Atmospheric Administration* (NOAA), pel *National Center for Atmospheric Research* i més de 150 entitats i universitats col·laboradores.

Va ser dissenyat com el successor del model MM5 i inclou totes les característiques d'aquest. La possibilitat d'utilitzar-se tant de forma operativa com en estudis de recerca atmosfèrica el fan un model molt versàtil i amb una comunitat d'usuaris creixent arreu del món.

El model es pot descarregar des del web de la organització (<http://www.wrf-model.org>).

1.6 Xarxa d'estacions meteorològiques automàtiques (XEMA)

La XEMA és la Xarxa d'Estacions Meteorològiques Automàtiques que gestiona el Servei Meteorològic de Catalunya (SMC). Aquesta xarxa es compon d'un total de 165 Estacions Meteorològiques Automàtiques (EMA) que transmeten la informació recollida al SMC.

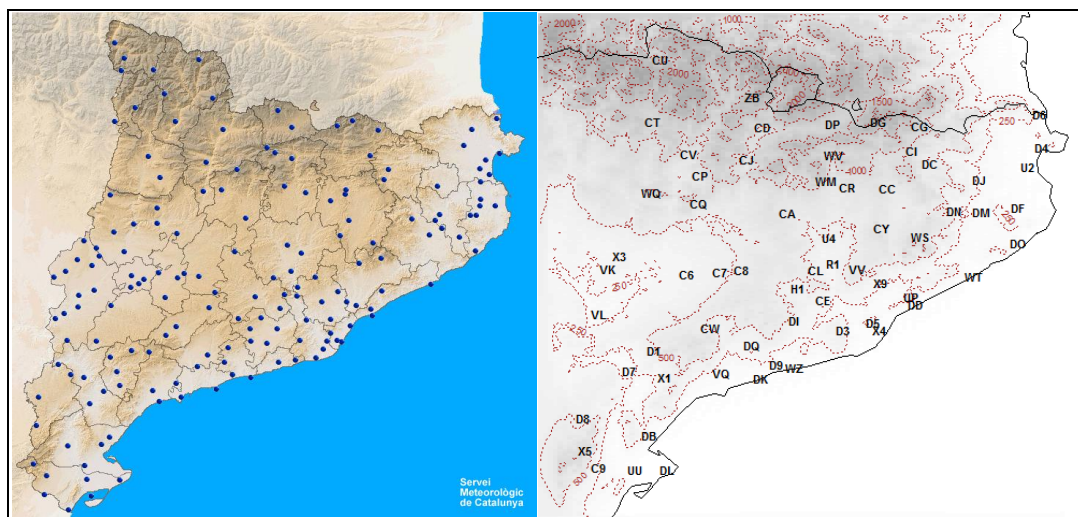


Figura 2 – A l'esquerra, mapa amb la situació de totes les estacions de la XEMA. A la dreta, mapa amb els codis de les estacions utilitzades

Cada estació recull informació de la temperatura, precipitació, velocitat i direcció del vent, humitat relativa, irradiància, pressió atmosfèrica i gruix de neu al terra. La freqüència i les característiques d'aquestes dades es poden consultar a la pàgina web del SMC (<http://www.meteo.cat>).

Les dades que s'utilitzen en aquest treball per realitzar la verificació dels models i la caracterització de les sèries temporals procedeixen d'aquesta xarxa d'estacions. Concretament es disposa de les dades recollides durant un any d'un subconjunt de 64 EMA. La llista d'aquestes 64 estacions es troba a l'ANNEX I. El conjunt de les dades han superat un control de qualitat aplicat pel mateix SMC. Més informació sobre el control de qualitat es pot consultar al mateix web del SMC.

1.7 Característiques climàtiques del territori

La situació latitudinal de Catalunya, entre les zones climàtiques temperada i tropical, la posició entre dos mars i dos continents i l'enorme varietat geogràfica del territori, fan que sigui un territori amb un singular mosaic de climes i amb una gran complexitat meteorològica.

Poques àrees amb l'extensió de Catalunya reuneixen la seva diversitat climàtica i afronten el repte d'una prognosi meteorològica tan complicada. Aquesta varietat de fenòmens meteorològics pot fer que la qualitat de les prediccions d'un MPN variï molt

en funció de l'àrea geogràfica del territori. Hi ha fenòmens que un MPN pot predir millor que altres, per això una verificació en diferents àrees geogràfiques pot ser d'especial interès a l'hora de determinar la validesa de les dades que proporciona un MPN.

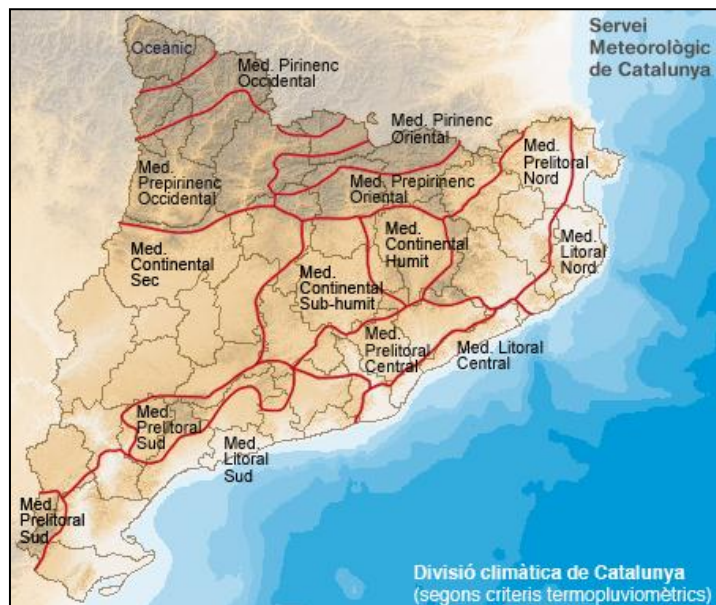


Figura 3 - Distribució climàtica de Catalunya. Font de la imatge: Servei Meteorològic de Catalunya.

1.8 Estacions Meteorològiques de referència

Per tal de fer un estudi territorial més profund de les sèries temporals s'han seleccionat algunes estacions de referència que representen diverses particularitats geogràfiques del territori.

1.8.1 Estació de Salòria (ZB)

L'estació de Salòria es troba dins el municipi d'Alins a la comarca del Pallars Sobrià. És una estació d'alta muntanya situada a 2451 metres sobre el nivell del mar i està situada prop de la carena que porta al cim de Salòria (2790 m).

Aquesta estació representa les característiques d'una estació de muntanya i possiblement destacarà la debilitat en la precisió geogràfica per part dels models de predicció numèrica. Per altra banda, la nuvolositat en les àrees de muntanya és més freqüent i irregular. Aquesta característica també pot dificultar la predicció de temperatura, sobre tot pel que fa a l'encert en l'amplitud tèrmica diària.



Figura 4 – Estació de Salòria

1.8.2 Estació del Raval (X4)



Figura 5 – Estació del Raval

L'estació del Raval es troba al centre de la ciutat de Barcelona a una altura de 33 metres sobre el nivell del mar.

Aquesta és una estació urbana i es pot veure influenciada per l'efecte de l'illa de calor urbana. No obstant, a l'estar situada al terrat d'un edifici, la influència de l'illa de calor pot minvar.

Per altra banda, és una estació situada prop del mar i per tant, en quant a efectes sobre la temperatura, es veurà afectada per la inèrcia tèrmica del mar.

1.8.3 Estació de Seròs (VL)

L'estació de Seròs es troba al municipi de Seròs a la comarca del Segrià a 100 metres d'altura sobre el nivell del mar. Es situa a les afores d'una petita àrea urbana.

Geogràficament, l'estació es troba en una petita vall prop del riu Ebre. Aquesta situació pot propiciar fenòmens com la boira que a la vegada afecten directament al pronòstic de la temperatura en superfície.

Tot i que la distància amb el mar no és gran, es pot considerar una estació d'interior degut a les característiques del relleu d'aquesta àrea.



Figura 6 – Estació de Seròs

1.8.4 Estació de Sant Pere Pescador (U2)

L'estació de Sant Pere Pescador es troba al municipi que porta al mateix nom a la comarca de l'Alt Empordà. Es situa a només 4 metres sobre el nivell del mar.

Aquesta estació es troba molt a la vora de la costa a més d'estar el mig d'una zona d'Aiguamolls i camps de conreu. Tot i això, els forts vents que hi bufen molt sovint contrasten amb l'ambient d'humitat el qual està fortament lligat degut a la seva situació.



Figura 7 – Estació de Sant Pere Pescador

2 MINERIA DE DADES DE LES SÈRIES TEMPORALS DE TEMPERATURA

2.1 Descripció de les sèries

En aquesta secció es descriurà el conjunt de dades utilitzat per realitzar la caracterització i la posterior validació de les sèries temporals que corresponen a la evolució de la temperatura en superfície en diferents punts de la geografia de Catalunya.

Es disposa de dades d'EMA en 64 punts diferents del territori (descrits a la secció 1.6) i valors de sortides dels models numèrics WRF i MM5 interpolades a cada un dels punts on s'hi localitza una EMA. El període que comprenen les dades és de l'any 2008.

2.1.1 Sèries temporals de les EMA (Observacions)

Les sèries temporals de les EMA són de freqüència horària. Aquestes enregistren informació de la temperatura, humitat, precipitació, pressió atmosfèrica, radiació solar, velocitat i direcció del vent. Aquestes dades s'utilitzaran per caracteritzar les sèries, verificar les sortides dels models i construir un model regressiu que millori la qualitat de les prediccions dels models.

2.1.2 Sèries temporals dels models (Pronòstics)

Les sèries temporals de les sortides dels models WRF i MM5 són de freqüència tri-horària. Les sèries contenen la informació de la temperatura, humitat pressió atmosfèrica velocitat i direcció del vent.

Les sèries dels models es componen de pronòstics inicialitzats a les 00Z per a intervals de cada 3 hores fins a 48 hores després de la inicialització (3h, 6h, 9h, ..., 48h). En aquest cas les sortides de les 00Z són dades analitzades on s'ha usat com a *first guess* una predicció anterior del model ECMWF.

El valor més important de les dades dels models és que aquestes estan interpolades als punts on s'hi troba una EMA. D'aquesta manera la verificació dels pronòstics es pot realitzar de forma molt acurada. Per la interpolació de les dades no només s'ha hagut de tenir en compte la latitud i la longitud a on es troba l'EMA sinó també l'alçada. L'error que s'hagi pogut cometre a l'estimar l'alçada del punt interpolat pot ser d'especial importància ja que pot provocar errors sistemàtics en els pronòstics.

2.2 Caracterització de les sèries d'EMA

En aquesta secció es realitzarà una caracterització de les sèries temporals de temperatura en superfície de les EMA.

2.2.1 Anàlisi Territorial de les sèries

En un territori divers com aquest, s'hi troben diferències molt significatives en les sèries temporals de la temperatura en superfície. L'efecte de l'altitud o la proximitat amb el mar són exemples clars de factors que influeixen en la temperatura.

Un dels efectes territorials més evidents és sobre la temperatura mitjana de la sèrie ja que aquesta depèn de la latitud i especialment de l'altitud. La distribució geogràfica

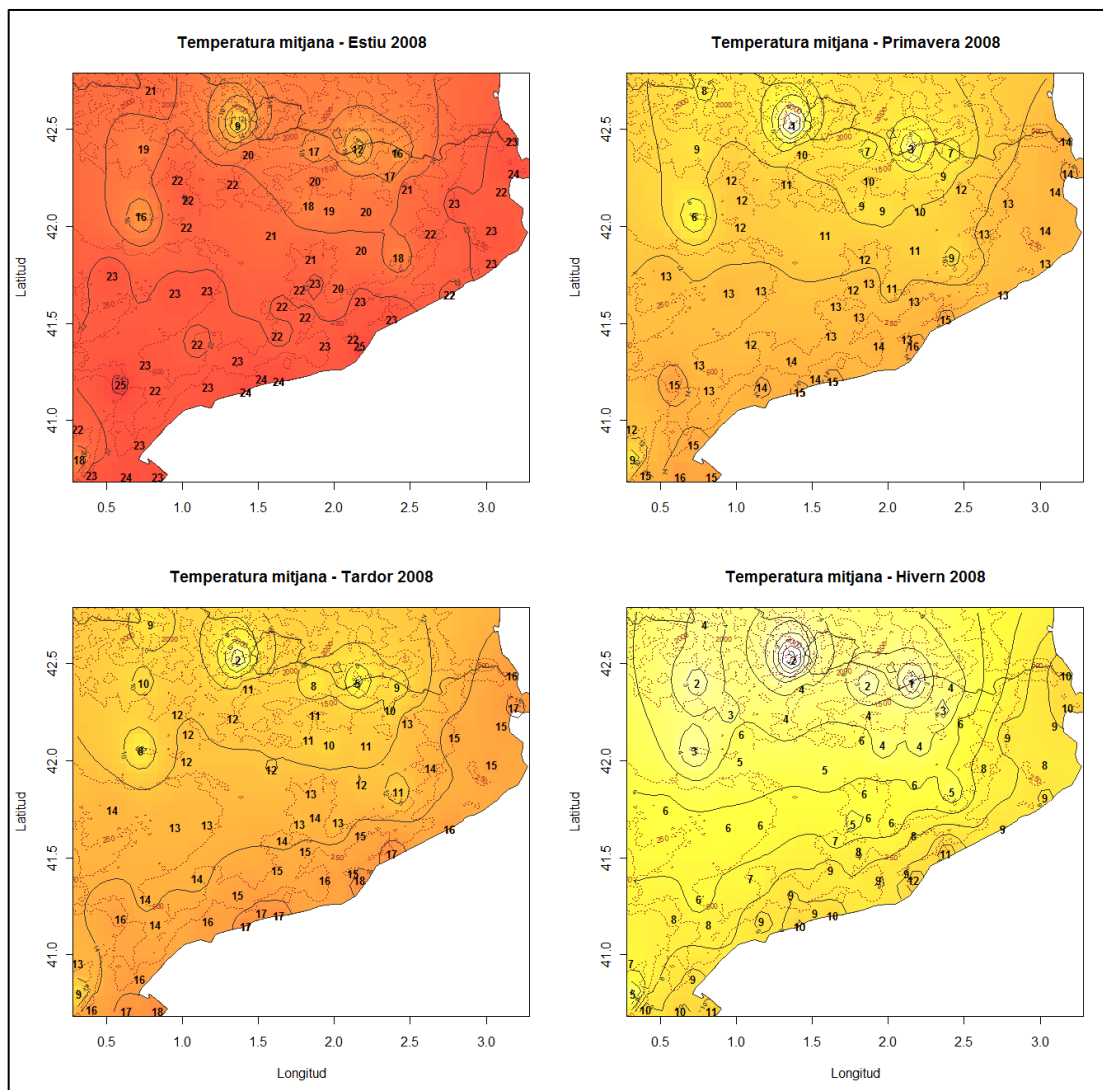


Figura 8 - Distribució territorial de la temperatura mitjana estacional

de la temperatura mitjana per a cada estació de l'any s'il·lustra a la **Figura 8** on també s'han dibuixat les corbes de nivell per tal de tenir informació del relleu. La correspondència de les estacions és la següent: Desembre, Gener i Febrer per l'hivern; Març, Abril i Maig per la primavera; Juny, Juliol i Agost per l'estiu i Setembre, Octubre i Novembre per la tardor. Tots els mesos corresponen al 2008.

A la figura es veu clarament l'efecte de l'altitud sobre la temperatura mitjana. En qualsevol període de l'any hi ha una diferència d'uns 10°C entre la temperatura mitjana d'una de les estacions de l'alt Pirineu (~2000m) i una de la costa.

És també interessant veure la distribució dels valors de la sèrie. El valor mitjà no té perquè ser el valor més freqüent si la distribució no és *normal*. A la **Figura 9** es representen les distribucions de les sèries de cada una de les estacions de referència. L'eix de les y és el tant per u de mostres que cauen dins el rang i l'eix de les x és la temperatura. La línia vertical discontinua de color vermell indica on es troba la mitjana de la sèrie.

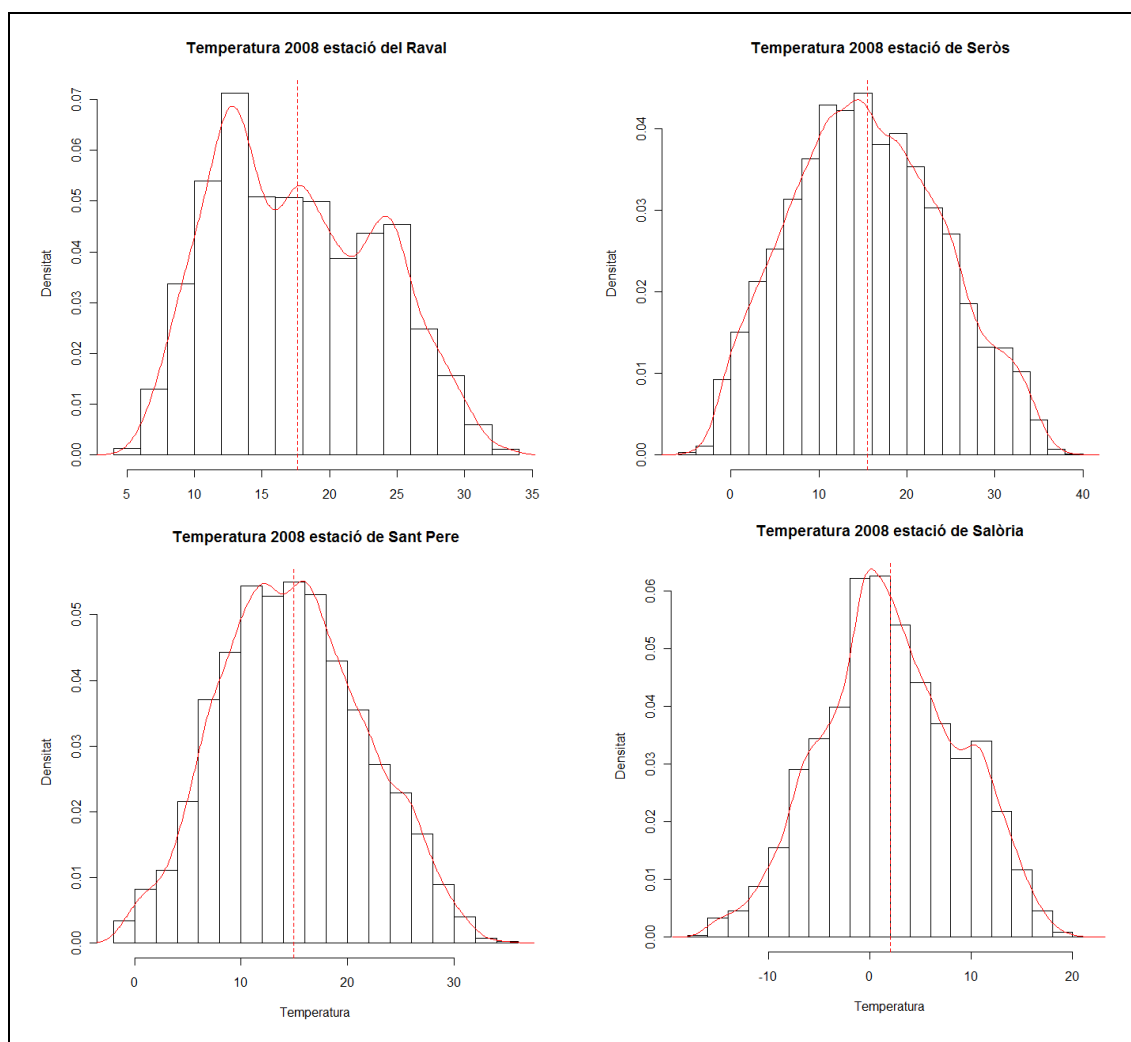


Figura 9 - Distribucions de les sèries de temperatura de cada estació de referència

L'estació que difereix més d'una distribució *normal* és la estació urbana del Raval. En aquest cas, els valors més freqüents de la sèrie són d'entre 13 i 14°C mentre que la mitjana és al voltant dels 17.5°C. A més, a l'histograma hi ha un altre màxim sobre els 24°C que reflecteix els llargs i calorosos estius de la zona urbana de Barcelona.

L'estació del Raval, a diferència de les altres, presenta dues situacions meteorològiques molt diferents que es donen de manera comuna.

Les altres estacions presenten *modes* pròximes als valors mitjans i per tant distribucions més properes a la *normal*. L'estació que sembla tenir una distribució més normal és l'estació costanera de Sant Pere Pescador. Un tret a destacar és el màxim secundari que s'enregistra a l'estació de Salòria (~11°C), semblant al que es dona a l'estació del Raval i no a les altres dues.

Per altra banda, un altre efecte de la geografia sobre l'evolució de les sèries pot ser sobre l'amplitud tèrmica diària d'aquestes. En aquest cas l'efecte del mar és d'especial importància, la inèrcia tèrmica de l'aigua modera les temperatures del territoris més propers de manera que l'amplitud tèrmica és menor. A la **Figura 10** es representen els valors mitjans de l'amplitud tèrmica diària per cada estació a tot el territori. Com era de

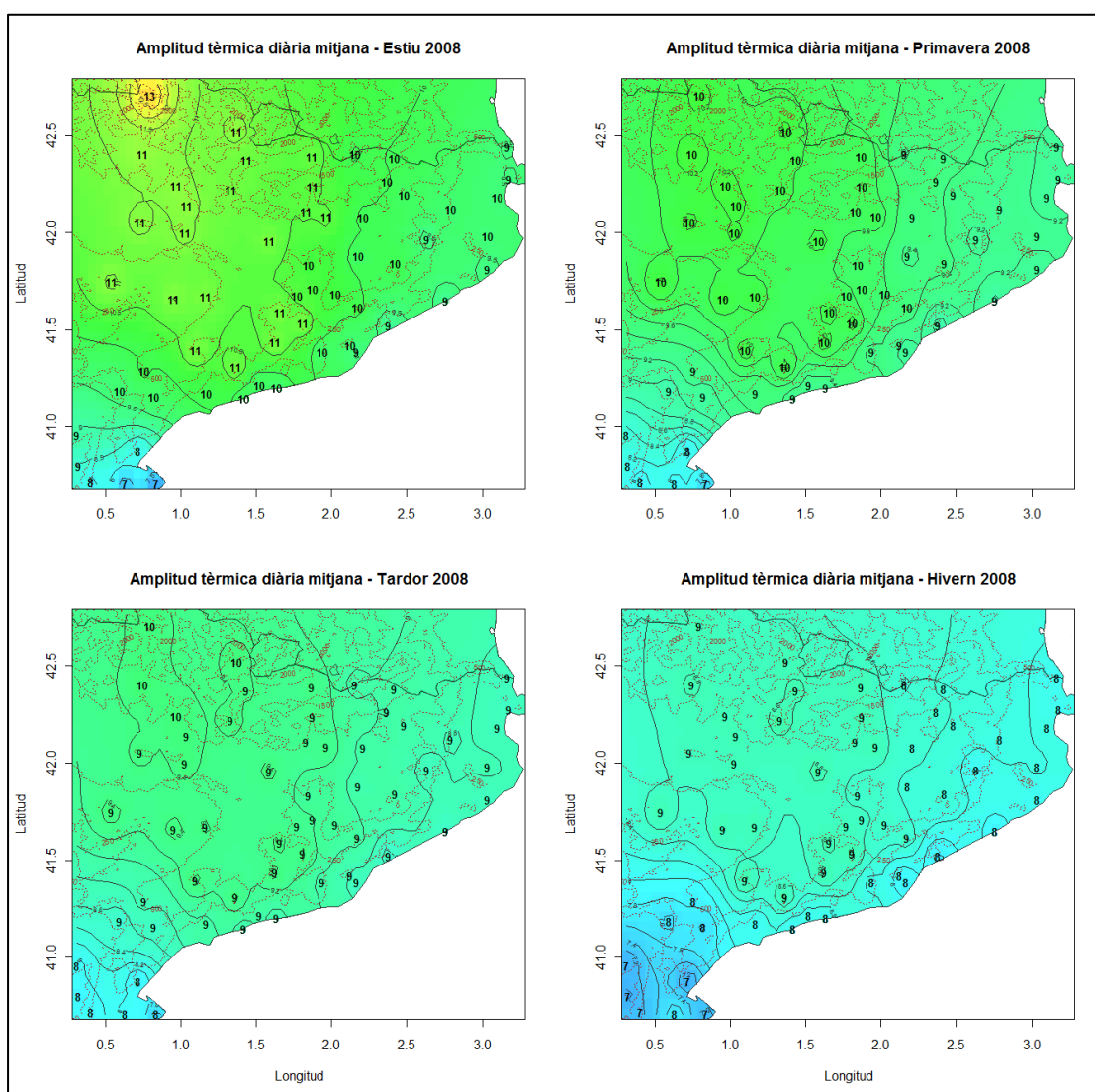


Figura 10 - Distribució territorial de l'amplitud tèrmica diària mitjana estacional

suposar, l'amplitud tèrmica augmenta a mida que ens endinsem a l'interior i es suavitza prop del mar.

L'amplitud diària mitjana és màxima a l'estiu i a l'interior mentre que els valors més baixos s'enregistren a l'hivern a l'àrea de les terres de l'Ebre. Tot i fer la mitjana estacional es veuen diferències notables entre els diferents punts del territori. L'amplitud tèrmica és una mesura molt variable al llarg dels dies i a grans trets depèn de la insolació i la nuvolositat. A l'estiu, quan hi ha més insolació, la temperatura màxima diària augmenta i així també, l'amplitud tèrmica. A la nit, el refredament és més pronunciat a l'interior i per això la diferència entre l'amplitud a la costa i l'interior augmenta.

2.2.2 Tests d'autocorrelació

Els tests d'autocorrelació per les sèries temporals ens indiquen la correlació entre els propis elements de la sèrie. El mètode per calcular la correlació entre dues sèries X i Y és la següent:

$$\text{corr}(X, Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

On E és l'operador de valor esperat, μ_X és la mitjana de la sèrie X i σ_X la desviació estàndard.

En el cas dels tests d'autocorrelació no es tindran dues sèries sinó que s'utilitzen sèries derivades de la pròpia sèrie original. D'aquesta manera, es compara la sèrie $(x_i, x_{i+1}, \dots, x_n)$ amb la $(x_{i-1}, x_i, \dots, x_{n-1})$ i totes les sèries resultants de desplaçar tants elements com interressi.

Aquest doncs, és també un test que avalua la persistència de la sèrie. A la **Figura 11** es mostren els resultats del test d'autocorrelació per a cada estació de referència. A l'eix y hi ha el valor de la correlació i a l'eix x el retard (*lag*). La correlació amb un retard de 0 elements sempre valdrà 1 (valor màxim en l'índex de correlació) ja que significa la correlació entre la sèrie original i aquesta mateixa sense desplaçament. L'autocorrelació amb un retard d'1 element indicarà la correlació entre el valor actual i el valor de fa una hora. Anàlogament, l'autocorrelació amb un retard de 24, 48 i 72 elements indicarà la correlació entre el valor actual i el de fa 1, 2 i 3 dies respectivament.

És interessant veure la gran diferència entre test de cada una de les estacions. En primer lloc, l'estació on la persistència a 24 hores aconsegueix una correlació més alta és la del Raval. De fet, en aquesta estació la correlació amb la persistència a 48 hores és del 0.9, un valor relativament elevat comparant-ho amb les altres estacions. Pel contrari,

l'estació d'alta muntanya de Salòria és la que la correlació amb la persistència a 24 hores té un valor més baix.

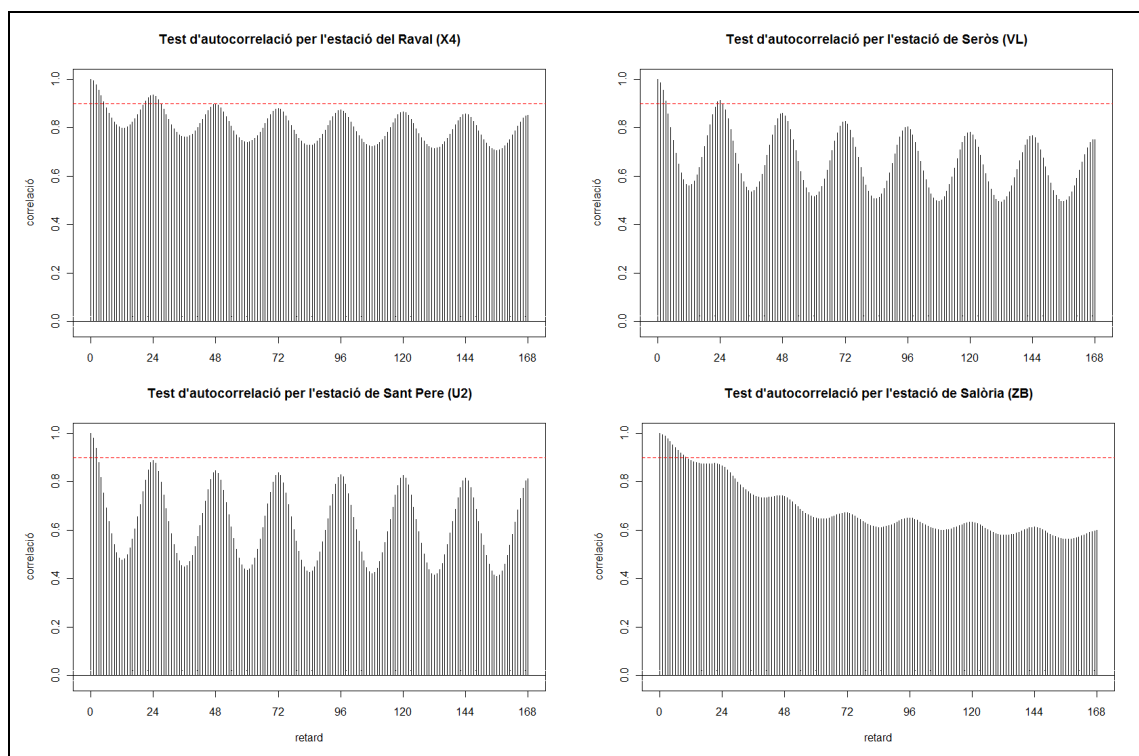


Figura 11 - Tests d'autocorrelació per cada una de les estacions de referència

La variabilitat intradiària¹ és també un tret diferencial entre les estacions de referència. El comportament en totes les estacions és similar i lògic, la correlació oscil·la al llarg del dia i tendeix a minvar quan més augmenta el retard. No obstant, aquests dos factors canvien substancialment en cada cas:

- L'amplitud de l'oscil·lació intradiària és la diferència entre el valor de correlació màxim i mínim en un mateix dia. Aquesta amplitud reflecteix la variabilitat de la temperatura al llarg del dia i indirectament pot tenir una relació amb la nuvolositat mitjana ja que un dels efectes de la nuvolositat és reduir l'amplitud tèrmica diària. La hipòtesis és consistent amb les dades de radiació solar que apareixen a la **Taula 1**. Entre els casos d'estudi, l'estació costera de Sant Pere Pescador és la que té una amplitud més gran mentre que la que enregistra una amplitud més petita és l'estació de Salòria.
- La tendència general de la correlació a descendir al llarg dels dies és una característica lògica i és un indicador de la inèrcia del territori. Com era

¹ Dins d'un mateix dia

d'esperar, les dues estacions d'interior (Seròs i Salòria) tenen una tendència més pronunciada a descendir que les estacions de la costa (Raval i Sant Pere).

Estació		Radiació Solar Mitjana Any 2008
Codi	Nom	W/m²
U2	Sant Pere Pescador	222.24
VL	Seròs	196.84
X4	Raval	196.77
ZB	Salòria	184.75

Taula 1 - Radiació solar diària mitjana per l'any 2008

3 VERIFICACIÓ DELS PRONÒSTICS DELS MODELS

En aquesta secció es realitzarà un procés de validació dels pronòstics dels models WRF i MM5. Després d'una breu introducció sobre els índexs utilitzats, s'analitzarà el comportament i les característiques de la qualitat dels pronòstics.

3.1 Indicadors de validesa dels pronòstics

3.1.1 Error Mig (ME)

L'índex ME (*Mean Error*) mesura l'error mig dels pronòstics. Al no considerar l'error absolut, aquest índex no serveix per mesurar la precisió ja els errors es podrien compensar si els errors són de signe diferent. No obstant, sí que serà vàlid per mesurar el *biaix* de les prediccions.

$$ME = \frac{\sum_{i=1}^n (o_i - f_i)}{n}$$

3.1.2 Error Mig Absolut (MAE)

L'índex MAE (*Mean absolute error*) mesura l'error absolut mitjà dels pronòstics sense considerar-ne la direcció. Avalua la precisió de variables contínues. El MAE és un índex lineal que mesura els errors dels pronòstics igualment ponderats.

$$MAE = \frac{\sum_{i=1}^n |o_i - f_i|}{n}$$

3.1.3 Error Quadràtic Mig (RMSE)

L'índex RMSE (*Root mean squared error*) és l'error quadràtic mig d'un conjunt de prediccions. Avalua la precisió de variables contínues. Com que l'error és elevat al quadrat abans de mitjanar-los, l'RMSE dóna relativament més pes als errors grans. Això significa que aquest índex és més útil quan els errors més grans són particularment no-desitjats.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (o_i - f_i)^2}{n}}$$

Els índexs MAE i RMSE es poden utilitzar conjuntament per diagnosticar la variació dels errors d'un conjunt de prediccions. L'RMSE sempre serà més gran o igual que l'MSE. Quant més gran sigui la diferència entre els dos índexs, més gran serà la variància dels errors individuals de la mostra. Si els dos són igual, significa que tots els errors són de la mateixa magnitud.

3.1.4 Desviació estàndard (SD, σ)

La desviació estàndard d'una sèrie és una mesura de dispersió.

$$\sigma = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$$

Per mesurar la validesa d'un pronòstic serà necessari comparar la desviació estàndard de la sèrie de les observacions i les prediccions. Un tret comú de molts models és la incapacitat de predir valors màxims o mínims, és a dir, valors llunyans a la mitjana de la sèrie. Aquesta característica s'anomena agudesesa (*sharpness*) i és un tret molt valuós a tenir en compte. La desviació estàndard d'una sèrie de pronòstics ens pot donar una idea sobre l'agudesesa de les prediccions. Una desviació estàndard més baixa que la sèrie d'observacions pot ser indicatiu de poca agudesesa en els pronòstics. El diagrama de Taylor, utilitzat més endavant en aquest treball, il·lustra la correspondència de la desviació estàndard entre les sèries.

3.1.5 Correlació (CORR, ρ)

La correlació indica la força i la direcció d'una relació lineal entre dues sèries temporals:

$$\rho_{X,Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

Els valors que pot valdre la correlació van de 0 a 1 essent 1 el valor que indica una correlació més alta entre les dues sèries.

3.2 Qualitat dels pronòstics dels models

Es disposen de pronòstics de temperatura en superfície a fins a 48 hores després de la inicialització. En aquest cas, totes les inicialitzacions són a les 00Z de cada dia de l'any 2008.

Un primer anàlisi dels pronòstics consisteix en veure l'evolució de la qualitat dels pronòstics a mida que passa el temps, des de les 00Z fins a 48 hores després. Al disposar únicament de inicialitzacions a les 00Z, els pronòstics a 12, 24 o 36 hores sempre correspondran a les 12Z, 00Z i 12Z (+1 dia) respectivament. Aquesta característica de les dades implica que no només es reflecteixi la tendència de l'error al llarg del temps sinó també l'error implícit a cada instant del dia.

A la **Figura 12** es mostren quatre gràfiques corresponents a l'evolució dels índexs RMSE i ME per a cada un dels models. A cada gràfica s'hi representen les quatre estacions de referència, cada una amb un color de línia diferent.

Per a les gràfiques corresponents a l'RMSE, exceptuant el valor de les 0 hores que conté un error implícit degut al *first guess* utilitzat, hi ha una lleu tendència creixent en l'error comès pel pronòstic, fet totalment coherent. A part de la tendència general, la característica més destacable és l'error implícit a l'hora del dia. Pràcticament en tots els casos, l'error més gran en el pronòstic de temperatura en superfície es comet a les 12Z,

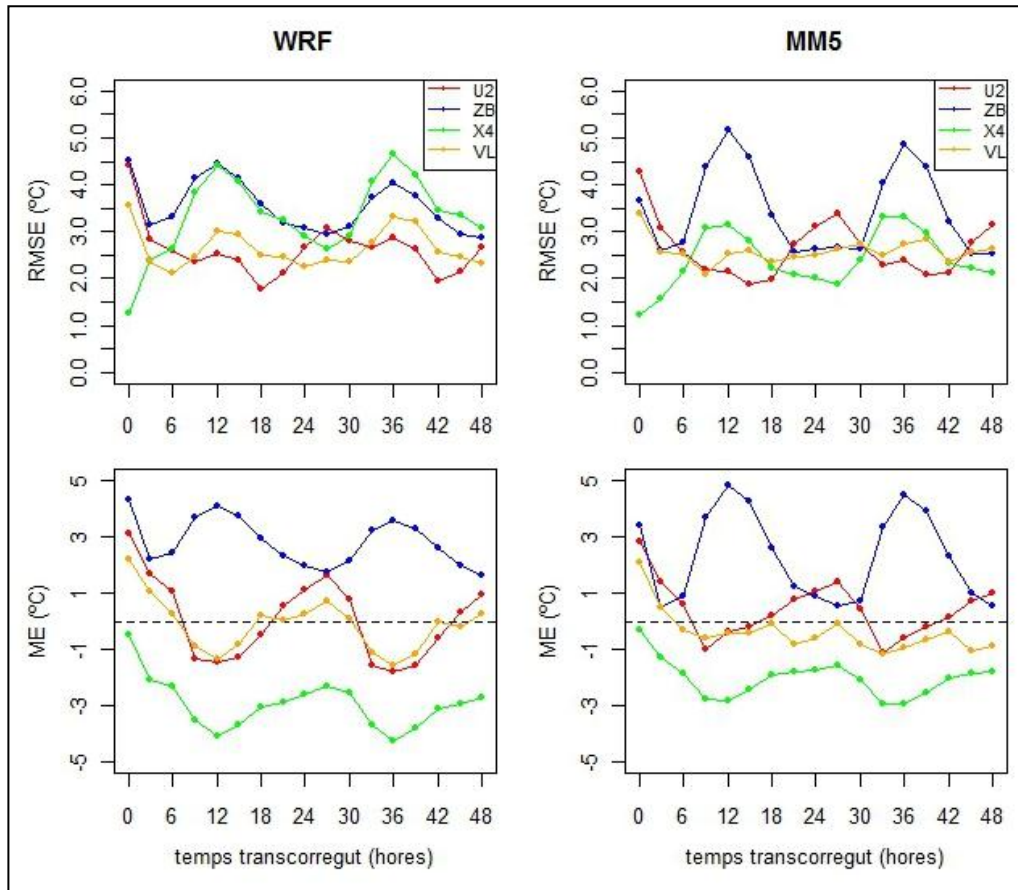


Figura 12 - Evolució dels índexs RMSE i ME en el temps per a cada estació de referència. A l'esquerra el model WRF i a la dreta l'MM5

hora en què l'efecte de l'escalfament solar és màxim. L'hora del dia on es comet el mínim error no és tant clara tot i que correspondria en un instant on l'activitat solar és mínima.

A les gràfiques de l'índex ME es pot veure el biaix de l'estació i el tipus d'error mig comès a cada hora del dia. En els dos models, tant el comportament per a cada estació com el biaix és molt similar.

Per una banda hi ha l'estació d'alta muntanya de Salòria que la tendència és de sobreestimar la temperatura sobretot a les hores centrals del dia. El cas totalment oposat a aquest és el de l'estació urbana del Raval. La tendència és de subestimar la temperatura especialment quan la insolació és màxima. En el cas de l'estació del Raval

aquesta tendència pot ser deguda a que els models no tenen en compte l'efecte de l'illa de calor en les estacions que es troben dins un nucli urbà important.

Els altres dos casos tenen un comportament similar. Tant a l'estació de Sant Pere Pescador com la de Seròs, els models tendeixen a subestimar la temperatura durant el dia i a sobreestimar-la durant la nit.

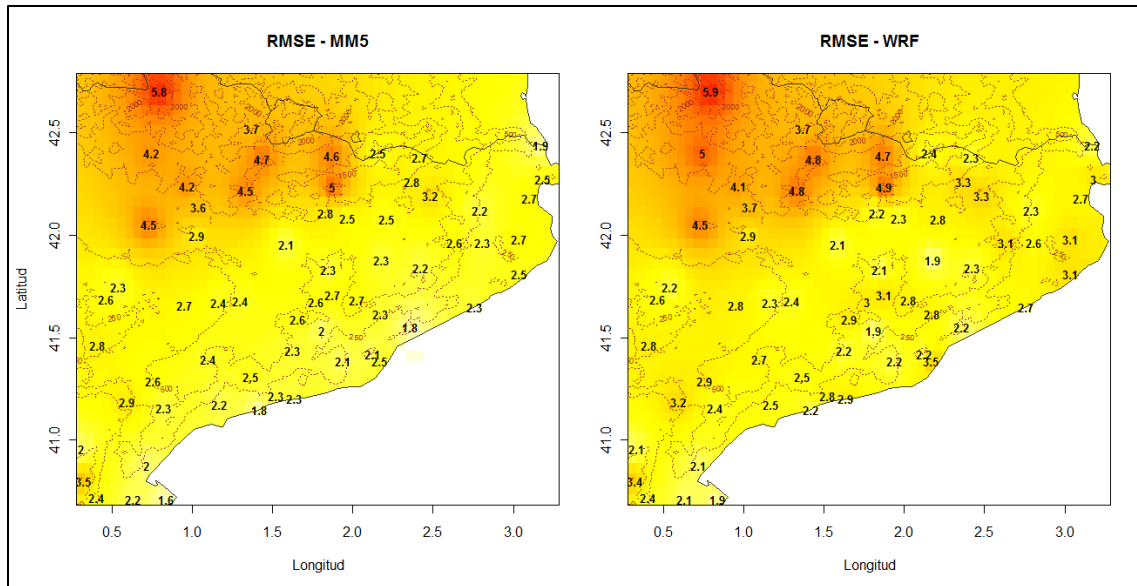


Figura 14 - Distribució territorial dels índexs RMSE per a cada un dels models

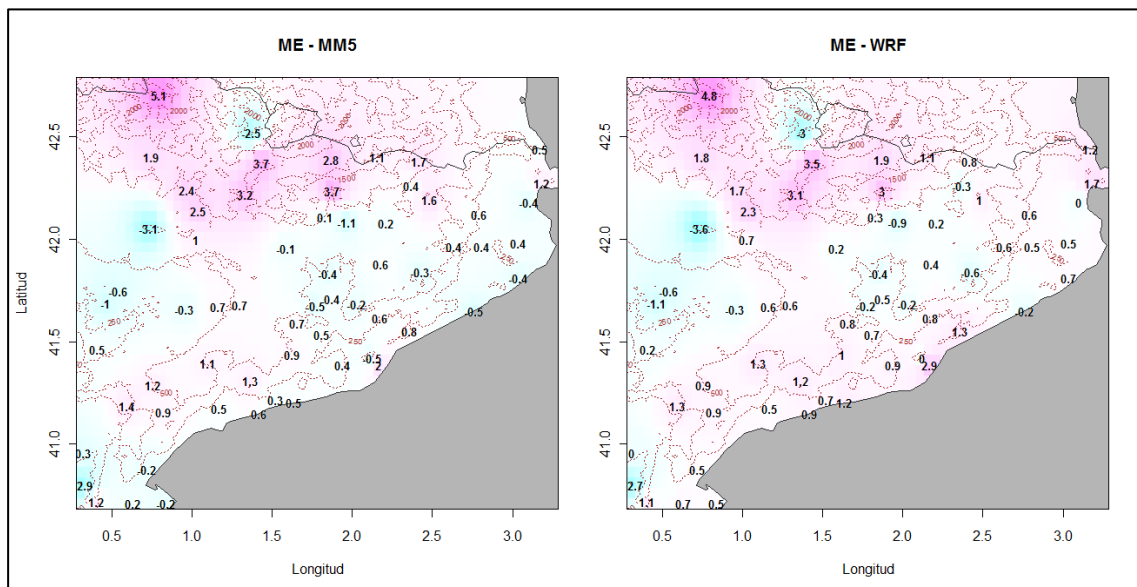


Figura 13 – Distribució territorial de l'índex ME per a cada un dels models

A més de veure l'evolució dels índexs en el temps també és interessant veure'n la distribució en el territori. Així es representa a les **Figura** per a totes les estacions i per a cada model.

Es pot veure clarament com els errors més grans en el pronòstic de temperatura es cometen a les estacions que es troben a la zona del Pirineu i Prepirineu. Aquest error, tal i com es pot veure en els mapes de l'índex ME, no té el mateix signe en tots els casos. Varia dels 3°C positius als -5°C. Al següent apartat es veurà com part d'aquest error és degut a l'error en l'estimació de l'altura de l'estació.

3.3 Errors en l'estimació de l'altura de les estacions

Una de les característiques més importants de les sortides dels models és que aquestes s'han interpolat als punts geogràfics on s'hi troba una EMA. Aquesta interpolació, a més de calcular la sortida a partir de la latitud i longitud dels punts de malla que utilitza el model, també calcula l'altura que li correspon al punt. Com que aquesta altura s'ha calculat a partir d'un model digital de terreny simplificat (amb poca resolució), les altures interpolades tampoc correspondran amb l'altura real del punt interpolat (punt on es troba l'EMA).

A la **Figura 15** es mostren els errors comesos a l'interpolat l'altura per a cada estació. El signe positiu o negatiu de l'error indica si s'ha sobreestimat o subestimat l'altura. Es pot veure que els errors més importants es cometen a les àrees muntanyoses i que aquests varien de signe segons si l'estació es troba a una vall o a una cima. L'error

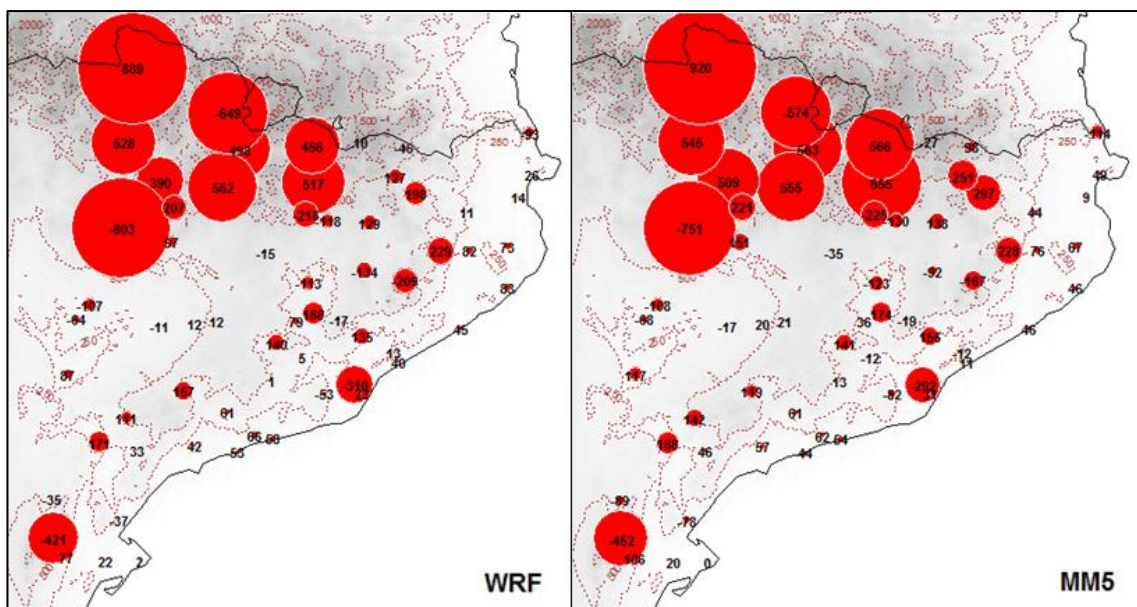


Figura 15 - Diferències en l'estimació de l'altura en metres en els punts de les estacions. Estimació menys altura real del model WRF a l'esquerra i MM5 a la dreta

és diferent en cada un dels models degut a la utilització de configuracions diferents a l'hora d'inicialitzar el model.

Per tal d'avaluar l'efecte que pot produir aquest error s'ha relacionat aquest error amb els índexs RMSE i ME. Aquesta relació és la que s'il·lustra a la **Figura 16**; en els primer dos casos, es relaciona l'error absolut en l'estimació de l'altura amb l'RMSE. Es pot veure que la relació s'ajusta molt bé a una relació lineal.

Més interessant que la relació de l'error absolut amb l'índex RMSE és la relació entre l'error (amb signe) i l'error mig. El resultat d'aquesta relació indicarà el biaix que introdueix l'error de l'estimació de l'altura en les prediccions de temperatura.

Tal com es veu a les dues gràfiques que relacionen l'error en l'altura i l'ME, la relació s'aproxima molt bé a una relació lineal. En els dos models, la relació és aproximadament: $ME = -0.35 - 0.0046 \cdot error$

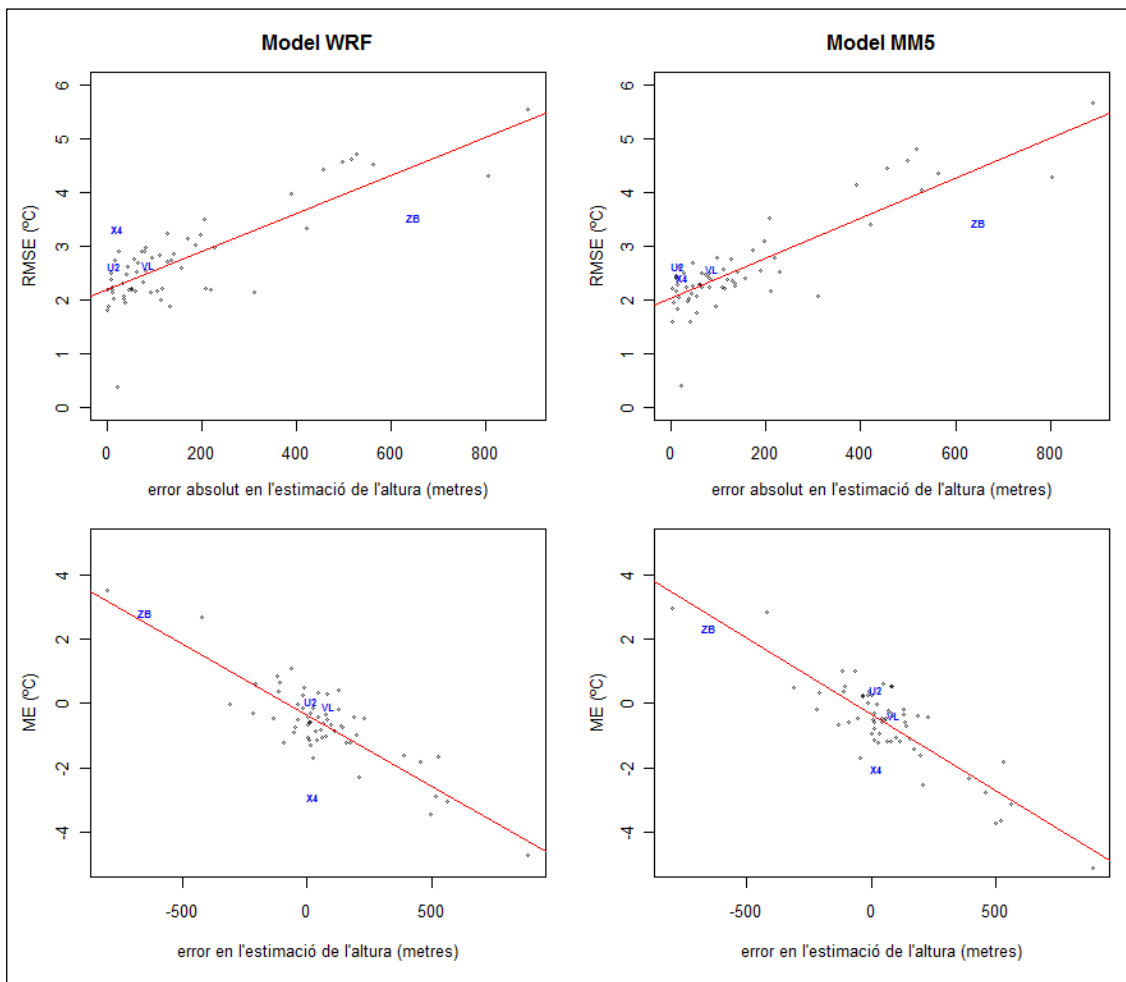


Figura 16 – Variació de l'RMSE i l'ME en funció de l'error comès a l'estimar l'altura de l'estació automàtica. La línia vermella indica la tendència de l'error.

3.4 Anàlisi de la persistència

La qualitat de les prediccions, sovint, es compara amb la persistència, el model més simple i a la vegada bastant eficaç per pronosticar la temperatura i altres camps de la meteorologia. La persistència a 24 hores és el model que considera que la temperatura persistirà les pròximes 24 hores.

Un tipus de gràfic molt útil per resumir estadísticament la similitud entre dues sèries és l'anomenat *diagrama de Taylor* (Taylor, 2001). Aquest diagrama permet representar tres indicadors estadístics (desviació estàndard, RMSE i correlació) en un espai de dues dimensions gràcies a la relació intrínseca que hi ha entre aquests tres valors. No obstant la representació dels valors no és exacte sinó que s'ajusta de manera que s'aproximin de

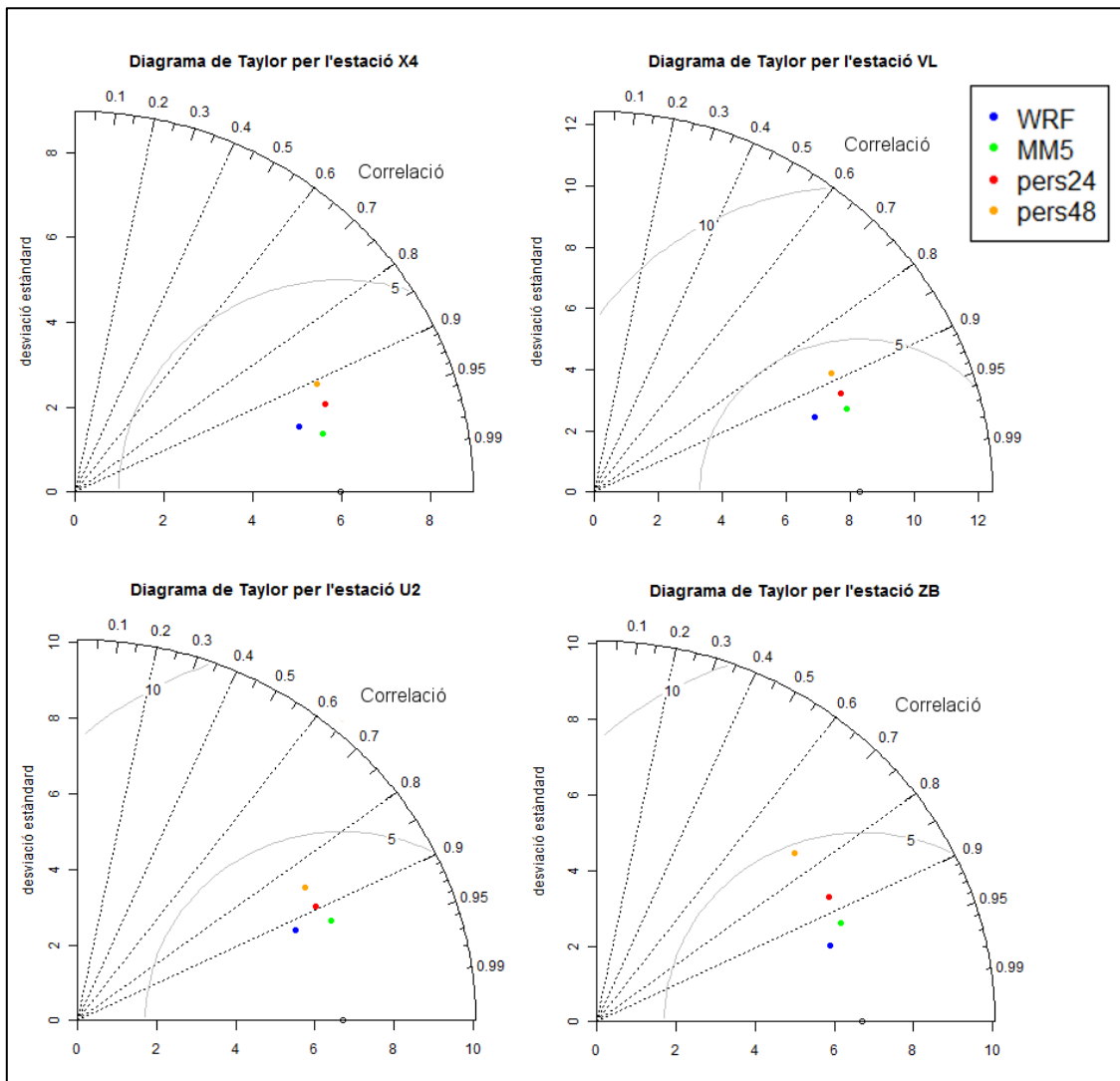


Figura 17 - Diagrames de Taylor per a cada estació de referència comparant els pronòstics dels models WRF i MM5 fins a 48 hores vista i la persistència de 24 i 48 hores. Als diagrames s'il·lustra la desviació estàndard de les sèries, la correlació i l'RMSE amb les observacions.

la millor manera els tres índexs; per aquesta raó, juntament amb el diagrama de Taylor és important adjuntar el valor numèric individual de cada un dels indicadors (**Taula 2**).

En els quatre diagrames de la **Figura 17** es comparen els pronòstics dels models WRF i MM5 i la persistència a 24 i 48 hores al llarg de tot l'any 2008. En tots els casos s'han avaluat els pronòstics dels models fins a 48 hores vista.

Cada sèrie està representada amb un punt d'un color diferent mentre que la sèrie corresponent a les observacions està representada amb un punt de color blanc sobre l'eix de les x. En els quatre casos, segons el diagrama de Taylor, els models superen a la persistència i, tal i com indicaven els tests de autocorrelació, la qualitat de la persistència és major en les estacions de la costa que en les de l'interior. No obstant, el diagrama de Taylor dona més importància a la correlació que a l'índex RMSE. Si s'observen els valors individuals a la **Taula 2** es pot veure com l'RMSE dels models no és sempre superior al de la persistència.

Al ser el diagrama de Taylor una il·lustració valuosa pel que fa a la quantitat d'informació que conté, s'ha realitzat el diagrama per a cada una de les estacions de què es disposava. El conjunt dels 60 diagrames restants es troben a l'**Annex II** i a la **Figura 18** hi ha un resum global dels resultats d'aquests diagrames.

Estació	Model	Desviació estàndard	RMSE	Correlació
X4	WRF	5.3491	3.4508	0.9562
	MM5	5.7106	2.4955	0.9706
	P24	5.9898	2.0991	0.9385
VL	WRF	7.3692	2.8203	0.9428
	MM5	8.2645	2.8063	0.9448
	P24	8.3314	3.2693	0.9227
U2	WRF	6.0547	2.6745	0.9174
	MM5	6.8146	2.6875	0.9244
	P24	6.7240	3.1016	0.8934
ZB	WRF	6.2059	3.6638	0.9463
	MM5	6.6272	3.6518	0.9200
	P24	6.7153	3.3950	0.8720

Taula 2 - Resum dels valors de la desviació estàndard, RMSE i correlació per a cada estació i model. En el cas dels models, s'han utilitzat els pronòstics de fins a 48 hores vista.

Els conjunt de resultats segueixen el mateix patró en bona part de les estacions: El model MM5 aconsegueix relativament millors resultats que el model WRF i ambdós models aconsegueixen batre la persistència. L'àrea que difereix més d'aquest patró és la de les estacions del Pirineu i Prepirineu. En alguns d'aquests casos, la persistència aconsegueix millors resultats que els models i també el model WRF resulta més eficaç que l'MM5.

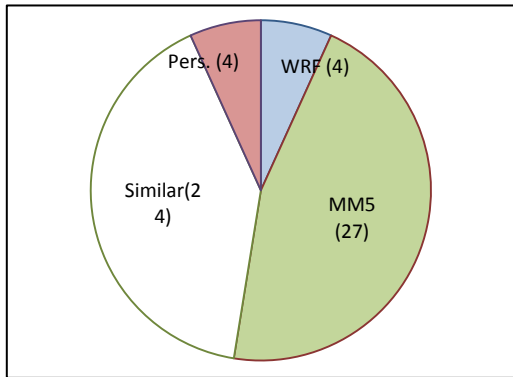


Figura 18 - Categorització de les estacions segons el model que ha obtingut millors resultats. La categoria Similar correspon a les estacions on els models WRF i MM5 obtenen resultats similars i la categoria Pers. correspon a aquelles estacions on la persistència obté millors resultats que algun dels models. El valor entre parèntesis correspon al nombre d'estacions corresponents a la categoria.

4 APLICACIÓ DE MODELS REGRESSIUS LOCALITZATS

4.1 Definició de model regressiu

Els models regressius s'utilitzen per l'anàlisi estadístic multivariant (o multivariable) i tenen com a objectiu establir una relació entre una variable dependent (Y), les variables independents (X_i) i un terme aleatori ε .

Matemàticament, donades les sèries $(Y_i, X_{i1}, X_{i2}, \dots, X_{ip})$ on $i = (1, 2, \dots, n)$, la relació entre les observacions Y_i i les variables independents X_{ij} ve donada per:

$$Y_i = \beta_0 + \beta_1 \varphi_1(X_{i1}) + \dots + \beta_p \varphi_p(X_{ip}) + \varepsilon_i$$

on β_1, \dots, β_p són els coeficients de la regressió, $\varphi_1, \dots, \varphi_p$ poden ser funcions no-lineals i ε_i és l'error de la relació

4.1.1 Model autoregressiu

Un model autoregressiu és un cas concret de model lineal regressiu. És un tipus de fórmula de predicció lineal que té com a finalitat predir uns valors de sortida a partir dels valors de sortides anteriors de la pròpia sèrie. Matemàticament, un model autoregressiu d'ordre p , $AR(p)$, es defineix com:

$$X_t = c + \sum_{i=1}^p \beta_i X_{t-i} + \varepsilon_t$$

on β_1, \dots, β_p són els paràmetres del model, c és una constant i ε_t és soroll blanc.

4.2 Disseny del model lineal regressiu

La finalitat del disseny d'un model regressiu és millorar les prediccions de la temperatura en superfície per part dels models operatius MM5 i WRF. Per tal que les

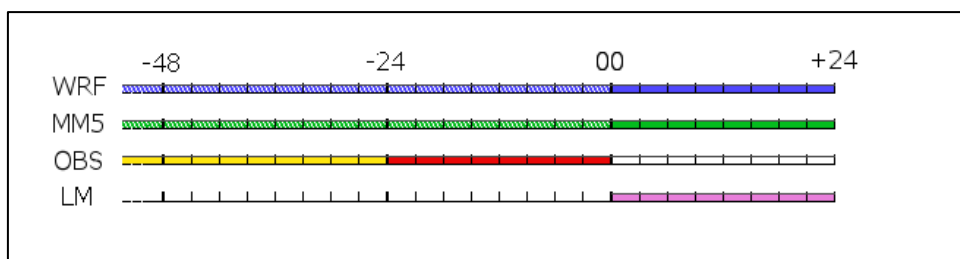


Figura 19 - Il·lustració de l'estat de les sèries en un instant de temps qualsevol per a l'aplicació operativa del model lineal.

prediccions segueixin sent operatives, s'ha de tenir en compte de quines dades es disposa a l'instant que es vol fer el pronòstic. A la **Figura 19** s'il·lustren les sèries temporals a un instant de temps concret: WRF, MM5 i Observacions. La sèrie LM correspon a les prediccions que realitzarà el model regressiu.

Com a model inicial, en aquest treball es proposa millorar les prediccions de temperatura fins a 24 hores i deixar la millora de les prediccions fins a 48 hores per a un treball futur. Tenint en compte aquesta característica, es podrà utilitzar la persistència a 24 hores per millorar la predicció del model.

4.2.1 Correcció de l'error en l'estimació de l'altura de les EMA

Un ajust que pot millorar significativament el resultat en algunes estacions és un model el qual corregeix l'error causat per la diferència entre l'altura real i del model.

Tal com s'ha calculat a la secció 3.3, l'error en el pronòstic de la temperatura provocat per l'error en l'estimació és lineal. D'aquesta manera, un primer model per millorar la sortida dels models es definiria com:

$$X'_t = X_t + err_{altura} \cdot 0.0046$$

Aquesta correcció de la sortida del model és particular per a cada estació meteorològica ja que depèn de la posició geogràfica de l'estació. A la **Taula 3** es mostren els valors de l'índex RMSE abans i després de la correcció per les estacions automàtiques de referència. En aquests quatre casos, es veu clarament com les millores més importants s'obtenen en les estacions amb un error més gran.

Estació	Model	Error (m)	RMSE (°C)		
			X _t	X' _t	
X4	WRF	22.89	3.3074	3.2205	3%
	MM5	31.04	2.3694	2.2582	5%
VL	WRF	87.07	2.8220	2.8499	-1%
	MM5	116.95	2.7614	2.7901	-1%
U2	WRF	14.11	2.7261	2.7325	0%
	MM5	9.02	2.7079	2.7165	0%
ZB	WRF	-649.03	3.8957	2.1093	46%
	MM5	-573.81	3.7930	2.6092	31%

Taula 3 - Resultats abans i després de la correcció de l'error en l'altura.

Amb aquest ajust s'aconsegueixen millores de prop del 40% en les estacions del Pirineu occidental i del 15% en altres estacions de muntanya del Prepirineu i Prelitoral.

4.2.2 Model lineal utilitzant els MPN i la persistència

A continuació es planteja un model regressiu que combina la sortida dels dos models i la persistència a 24 hores. També s'ha tingut en compte la correcció de l'error descrit a la secció anterior. Matemàticament, la sèrie proposada es descriu com:

$$X'_t = (0.0046 \cdot zerr) + (\beta_0 + \beta_1 \cdot MM5_t^{24} + \beta_2 \cdot WRF_t^{24} + \beta_3 \cdot X_{t-24} + \beta_4 \cdot err_{t-24})$$

on $MM5_t^{24}$ i WRF_t^{24} són les prediccions fins a 24 hores dels models, X_{t-24} és la persistència a 24 hores i err_{t-24} és l'error mig comès pels models 24 hores abans de la predicció.

Els paràmetres beta variaran segons cada estació automàtica i es calcularan a partir d'una mostra de la sèrie d'observacions. Per mantenir la independència dels paràmetres beta, la sèrie utilitzada per ajustar-los i la utilitzada per validar els resultats del model no pot ser la mateixa. Per aquest motiu, s'ha dividit la sèrie de temperatura del 2008 en dues parts separant els primers sis mesos de la segona part de l'any.

La sèrie que resulta de la regressió lineal consta de dues parts: la primera meitat s'ha generat emprant la fórmula parametritzada amb la segona part de l'any i la segona meitat que s'ha generat amb la fórmula parametritzada amb els primers sis mesos de l'any. D'aquesta manera es conserva la independència dels paràmetres beta i les observacions a més de tenir una sèrie completa d'un any sencer. A l'**Annex III** es mostren el valor d'aquests paràmetres beta per a cada període i estació automàtica.

4.3 Avaluació global dels resultats

Per avaluar l'eficàcia del model regressiu es compararan els resultats de la mateixa manera que s'ha realitzat la fase de validació dels models WRF i MM5. Per començar, s'han realitzat els diagrames de Taylor per les prediccions fins a 24 hores dels dos MPN, persistència a 24 i 48 hores i a més, els resultats del model regressiu.

A la **Figura 20** es mostren els diagrames de Taylor de cada una de les estacions automàtiques de referència. En tots els casos s'obté una millora significativa respecte el model que obtenia millors resultats.

El diagrama de Taylor il·lustra tres índexs estadístics en un espai de dues dimensions, Degut a aquesta limitació, els resultats que es mostren no sempre coincideixen amb els resultats reals sinó que en alguns casos es realitza un ajustament per tal que els tres valors s'aproximin el més possible als valors reals. A l'ajustar els punts al diagrama, l'índex que té més pes és la correlació mentre que el que l'RMSE en

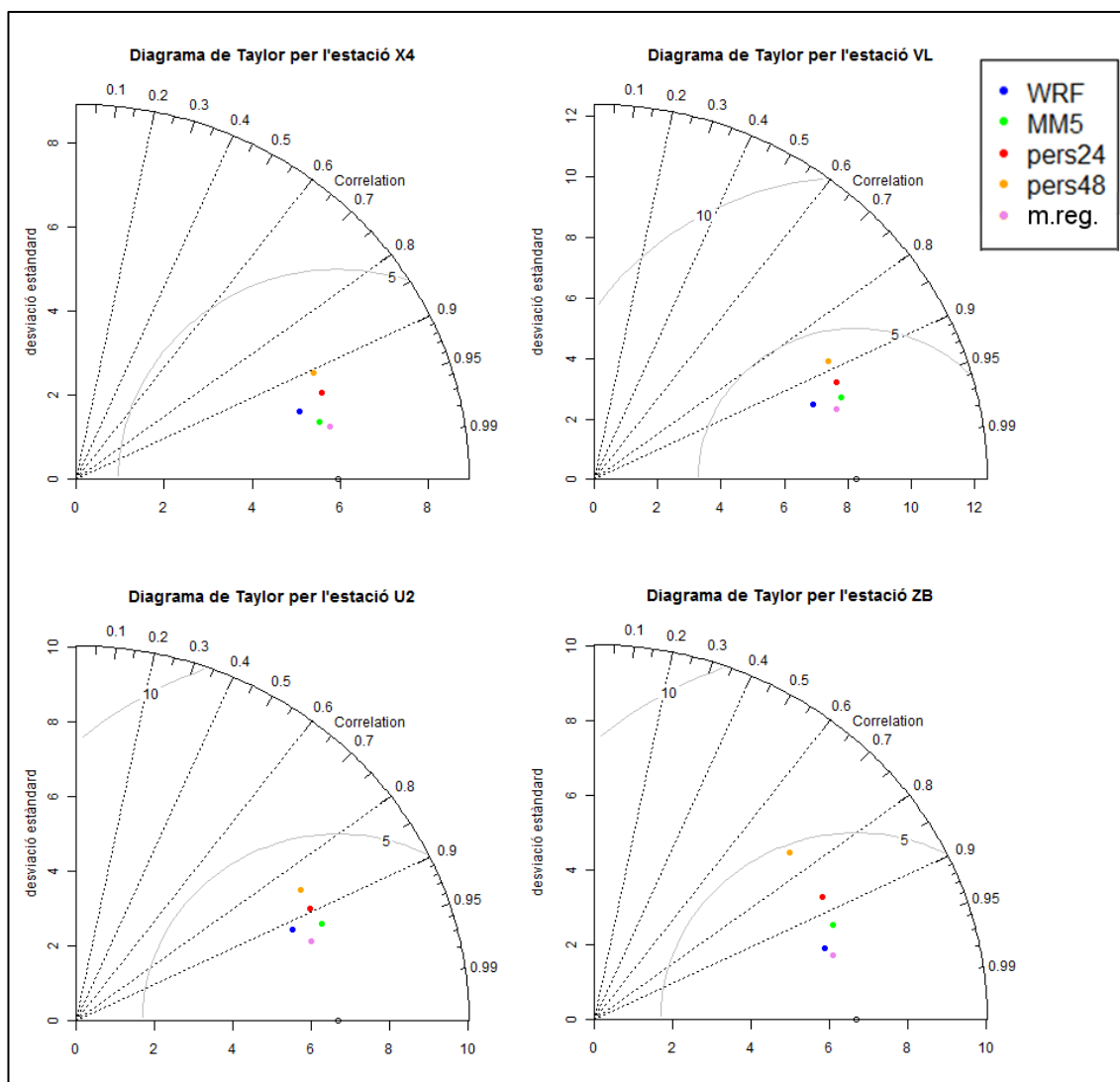


Figura 20 – Diagrames de Taylor comparant la sortida dels MPN, les persistències i la sortida del model regressiu.

molts casos no queda ben representat. Al ser l'RMSE un índex molt important pel que a la predicció de temperatura, s'ha realitzat un anàlisi més detallat d'aquest índex.

A la **Taula 4** es compara l'índex RMSE obtingut pels models WRF, MM5 i persistència a 24 hores amb l'obtingut pel model regressiu. A l'última columna es quantifica la reducció de l'error per a cada un dels models. Es pot veure també com en alguns casos l'RMSE obtingut per la persistència és millor que el cas dels MPN tot i que als diagrames de Taylor no sembla així; com s'ha dit, això és degut a que el diagrama de Taylor dona més importància a la correlació que a l'RMSE.

Per tenir una visió global de l'RMSE obtingut en cada una de les 64 estacions s'ha il·lustrat territorialment aquest índex utilitzant els mateixos colors que a la **Figura** .

Estació	Model	RMSE (°C)		
		X_t	X'_t	
X4	WRF	3.295	1.288	61%
	MM5	2.367		46%
	P24	2.101		39%
VL	WRF	2.826	2.407	15%
	MM5	2.764		13%
	P24	3.276		27%
U2	WRF	2.724	2.247	18%
	MM5	2.691		17%
	P24	3.104		28%
ZB	WRF	3.888	1.831	53%
	MM5	3.788		52%
	P24	3.402		46%

Taula 4 – Resultats de l'aplicació del model lineal regressiu. Es mostra l'RMSE obtingut i la millora respecte els models WRF, MM5 i persistència 24 hores.

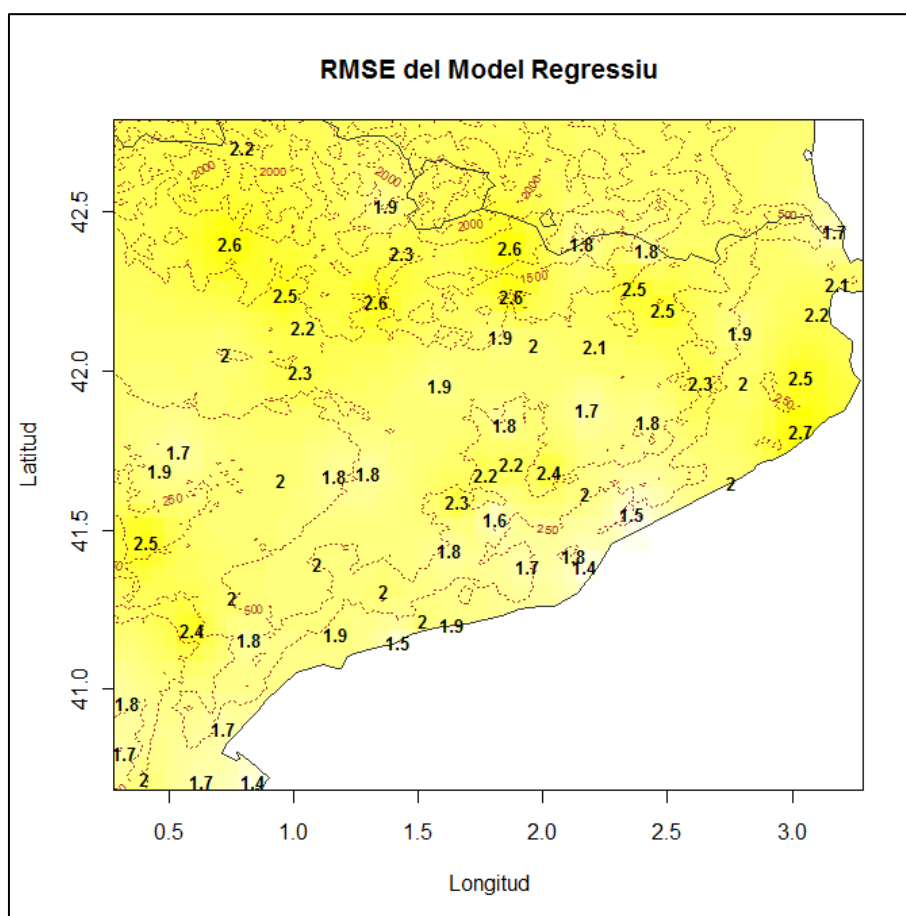


Figura 21 - Distribució de l'índex RMSE per la sortida del model lineal regressiu.

Utilitzant el model regressiu s'ha eliminat l'error que es localitzava sobretot a l'àrea del Pirineu occidental. A la **Figura 21** es pot veure com l'RMSE s'ha homogeneïtzat en tot el territori eliminant l'efecte de l'altura i lligant la quantitat d'error a fenòmens localitzats.

Per tal de quantificar la millora que suposa la utilització d'aquest model regressiu s'ha calculat la reducció mitjana de l'RMSE en totes les EMA del territori respecte els MPN. A la **Figura 22** es pot veure com les millores més importants es focalitzen a l'àrea del Pirineu occidental exceptuant el cas de l'estació dels Ports i del Raval que també superen el 50% de reducció de l'RMSE. A la resta del territori les reduccions de l'RMSE varien del 10 al 30%.

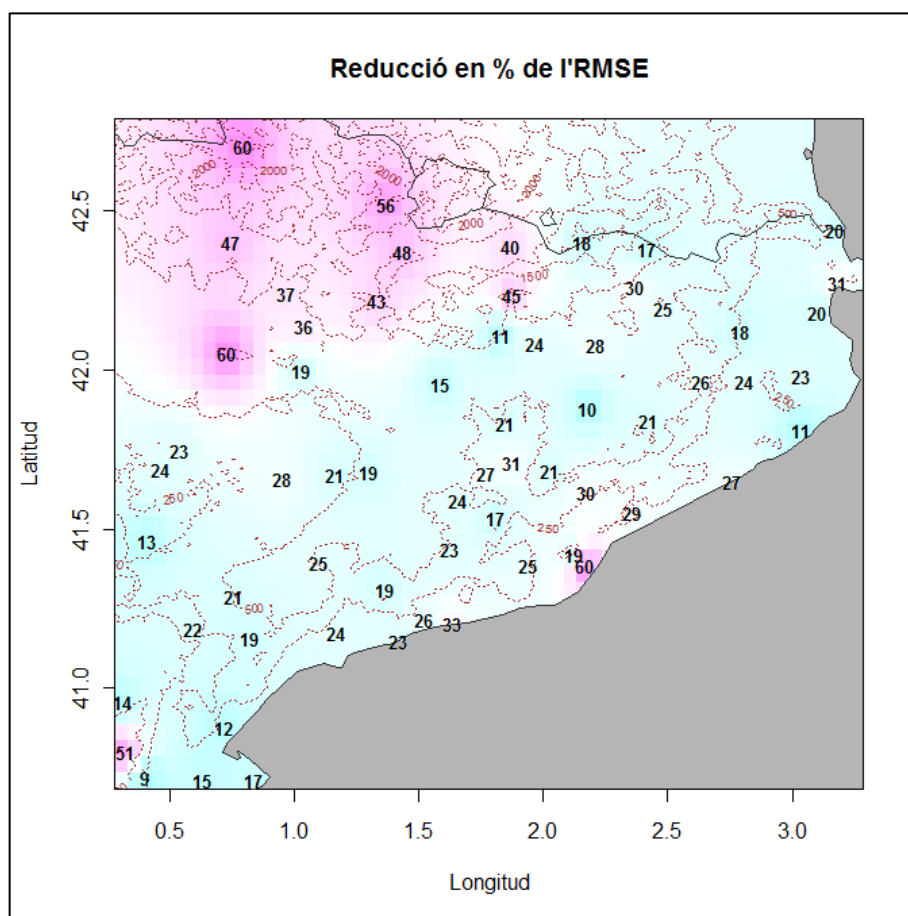


Figura 22 - Distribució del percentatge de reducció de l'error RMSE mig del model regressiu.

5 CONCLUSIONS

5.1 Resultats obtinguts

En aquest treball s'han caracteritzat les sèries de temperatura en superfície (2 metres) registrades per 64 estacions meteorològiques automàtiques repartides pel territori de Catalunya. A més, s'ha realitzat un anàlisi més profund de 4 estacions de referència escollides de manera que es representés la varietat geogràfica i climàtica del territori estudiat.

En una segona fase s'ha iniciat un procés de validació dels pronòstics de temperatura en superfície per part dels models WRF i MM5. En el conjunt de les 64 estacions meteorològiques automàtiques emprades per la verificació dels resultats dels models, el model MM5 obté resultats lleugerament superiors al model WRF en bona part dels punts. Comparant amb la persistència a 24 hores, els models l'aconsegueixen batre excepte en algunes estacions de muntanya.

Pel que fa a l'aplicació de models regressius localitzats a cada un dels 64 punts del territori tractats, s'ha reduït l'RMSE dels pronòstics de temperatura entre el 10 i el 60%. Aquesta millora ha estat especialment important en les àrees del Pirineu Occidental on els errors en l'estimació de l'altura de les estacions era més gran.

5.2 Treball futur

Un procés important a realitzar és el d'anàlisi de temperatures extremes de les sèries i la verificació dels pronòstics per part dels models de predicció numèrica. La predicció dels màxims i mínims de temperatura són especialment valuosos i a la vegada complicats de pronosticar. Caldria veure a quin tipus de millora s'atribueix al percentatge de reducció de l'RMSE que obtenen els models regressius.

Les millores en els pronòstics de temperatura en superfície obtingudes a partir dels models regressius poden ser utilitzats per millorar el funcionament d'altres models que utilitzin com a entrada aquests valors. Un dels pronòstics que es vol millorar és el de predicció del tipus de precipitació per tal de complementar la informació del radar meteorològic i millorar potencialment la predicció a curt termini d'aquest fenomen (Ortiz, 2009).

Pel que fa als models regressius, seria bo calcular els coeficients beta utilitzant la sèrie sencera d'un any i validar-los utilitzant una altra sèrie d'un any sencer. A més, els models regressius es podrien sofisticar incorporant camps diferents als de temperatura. Aquests podrien ser per exemple el de radiació solar, temperatura en altura o humitat relativa.

REFERÈNCIES

- Bech J., Codina B., Lorente J. (2007). Forecasting weather radar propagation conditions. *Meteorology and atmospheric physics*(96), 229-243.
- Bogren, J. & Gustavsson, T. (2002). Site specific road surface temperature forecast improvements by use of radiation measurements. *Proceedings of the 11th SIRWEC conference*. Sapporo.
- EUMETCAL. (2011). *EUMETCAL*. (Finish Meteorological Institute) Retrieved 2011, from <http://www.eumetcal.org/resources/ukmeteocal/verification/>
- Gjertsen, U., Odegaard, V. (2005). The water phase of precipitation: a comparison between observed, estimated and predicted values. *Atmospheric Research*(77), 218-231.
- Jolliffe I. T., Stephenson D. B. . (2003). *Forecast verification: a practitioner's guide in atmospheric science*. Wiley Ed.
- Koistinen, J., Saltikoff, E. (1998). Experience of customer products of accumulated snow, sleet and rain. *Advanced Weather Radar System, COST 75*, 397-406.
- MM5 Comunity Model. (2003). Retrieved from <http://www.mmm.ucar.edu/mm5/>
- Murphy, A.H., B.G. Brown, and Y.-S. Chen. (1989). Diagnostic verification of temperature forecasts. *Weather and Frecasting*(4), 485-501.
- Ortiz, J. A. (2009). Anàlisi d'observacions de superfície en relació al diagnòstic de tipus de precipitació observat amb radar. *Dipòsit UB*.
- Ríos-Moreno, G. J. (2007). Modelling temperature in intelligent buildings by means of autoregressive models. *Automation in construction*(17), 713-722.
- R-Project. (1993-2011). Retrieved from <http://www.r-project.org>
- Shumway R.H., S. D. (2006). *Time series analisys and its applications: R Examples*. Retrieved from <http://www.stat.pitt.edu/stoffer/tsa2>
- Smith, Goodchild, Longley. (2006 - 2011). *Geospatial Analysis - a comprehensive guide*. <http://www.spatialanalysisonline.com>.
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*
- Teetor, P. (2011). *R CookBook*. O'Reilly Ed.

Valeriano, E. M. (2001). *Evaluacion de los pronosticos de tempratura del aire obtenidas con el modelo MM5 para algunas ciudades de la costa y sierra del Perú*. Instituto Geofísico del Perú.

Wilks, D. S. (2006). *Statistical methods in atmospheric sciences*. Elsevier Ed.

William K., Harry G. (1974). Forecasting local weather by means of model output statistics. *American Meteorological Society*.

WRF Model Web Site. (2006). Retrieved from <http://www.wrf-model.org/index.php>

Zucchini W., Nenadic O. (n.d.). *Time series analisys with R*.

TAULA DE CONTINGUTS

Figura 1 - Logotip del paquet estadístic R.....	10
Figura 2 – A l'esquerra, mapa amb la situació de totes les estacions de la XEMA. A la dreta, mapa amb els codis de les estacions utilitzades	12
Figura 3 - Distribució climàtica de Catalunya.....	13
Figura 4 – Estació de Salòria.....	13
Figura 5 – Estació del Raval.....	14
Figura 6 – Estació de Seròs	14
Figura 7 – Estació de Sant Pere Pescador	14
Figura 8 - Distribució territorial de la temperatura mitjana estacional	16
Figura 9 - Distribucions de les sèries de temperatura de cada estació de referència.....	17
Figura 10 - Distribució territorial de l'amplitud tèrmica diària mitjana estacional.....	18
Figura 11 - Tests d'autocorrelació per cada una de les estacions de referència	20
Figura 12 - Evolució dels índexs RMSE i ME en el temps per a cada estació de referència. A l'esquerra el model WRF i a la dreta l'MM5.....	24
Figura 13 – Distribució territorial de l'índex ME per a cada un dels models	25
Figura 14 - Distribució territorial dels índexs RMSE per a cada un dels models	25
Figura 15 - Diferències en l'estimació de l'altura en metres en els punts de les estacions. Estimació menys altura real del model WRF a l'esquerra i MM5 a la dreta	26
Figura 16 – Variació de l'RMSE i l'ME en funció de l'error comès a l'estimar l'altura de l'estació automàtica. La línia vermella indica la tendència de l'error.....	27
Figura 17 - Diagrames de Taylor per a cada estació de referència comparant els pronòstics dels models WRF i MM5 fins a 48 hores vista i la persistència de 24 i 48 hores. Als diagrames s'il·lustra la desviació estàndard de les sèries, la correlació i l'RMSE amb les observacions.	28
Figura 18 - Categorització de les estacions segons el model que ha obtingut millors resultats. La categoria Similar correspon a les estacions on els models WRF i MM5 obtenen resultats similars i la categoria Pers. correspon a aquelles estacions on la persistència obté millors resultats que algun dels models. El valor entre parèntesis correspon al nombre d'estacions corresponents a la categoria.....	30

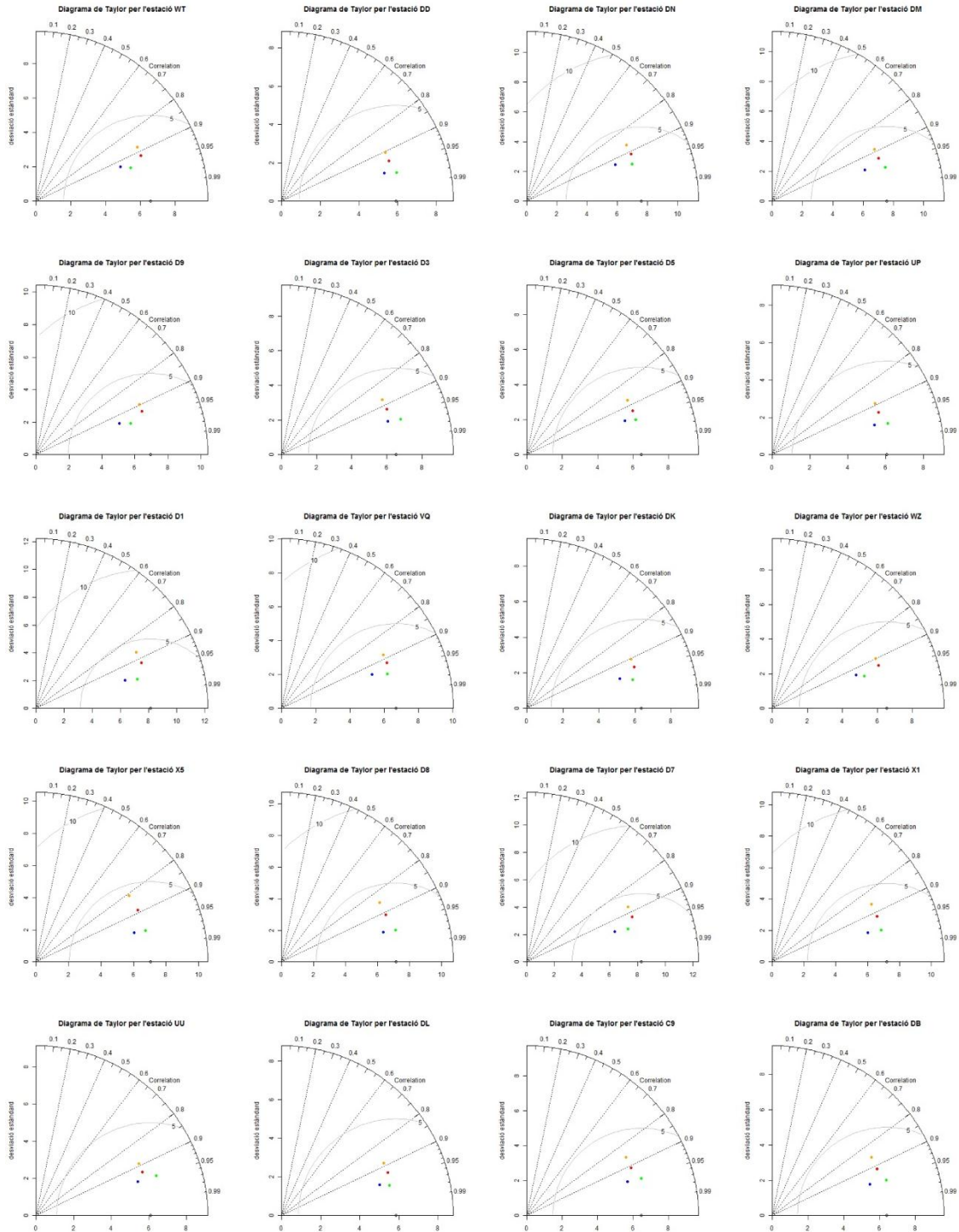
Figura 19 - Il·lustració de l'estat de les sèries en un instant de temps qualsevol per a l'aplicació operativa del model lineal.	31
Figura 20 – Diagrames de Taylor comparant la sortida dels MPN, les persistències i la sortida del model regressiu.	34
Figura 21 - Distribució de l'índex RMSE per la sortida del model lineal regressiu.....	35
Figura 22 - Distribució del percentatge de reducció de l'error RMSE mig del model regressiu.....	36
Taula 1 - Radiació solar diària mitjana per l'any 2008	21
Taula 2 - Resum dels valors de la desviació estàndard, RMSE i correlació per a cada estació i model. En el cas dels models, s'han utilitzat els pronòstics de fins a 48 hores vista.....	29
Taula 3 - Resultats abans i després de la correcció de l'error en l'altura.	32
Taula 4 – Resultats de l'aplicació del model lineal regressiu. Es mostra l'RMSE obtingut i la millora respecte els models WRF, MM5 i persistència 24 hores.....	35

ANNEX I – Taula d’EMA

Code	Latitud (°)	Longitud (°)	Altura real (m)	Altura MM5 (m)	Altura WRF (m)
Amposta [UU]	40.70961	0.63314	3	23	25
Illa de Buda [DL]	40.70920	0.83531	0	0	2
Mas de Barberans - Abocador [C9]	40.72018	0.40087	240	346	317
el Perelló [DB]	40.87497	0.71670	179	101	142
PN dels Ports [X5]	40.79883	0.31926	1055	603	634
Horta de Sant Joan [D8]	40.95619	0.31014	515	426	480
Vinebre [D7]	41.18682	0.59479	53	241	224
Falset [X1]	41.15561	0.82057	359	405	392
Margalef [D1]	41.28705	0.75488	404	546	515
Constantí [VQ]	41.17312	1.16881	112	169	154
Torredembarra [DK]	41.14863	1.41952	2	46	55
Cunit [WZ]	41.20371	1.63451	17	71	75
el Vendrell [D9]	41.21724	1.52214	59	121	125
Vallirana [D3]	41.38378	1.93675	252	170	199
Barcelona - Observatori Fabra [D5]	41.42043	2.12500	411	119	101
Barcelona - el Raval [X4]	41.38566	2.16877	33	64	56
Cabrils [UP]	41.54958	2.35810	81	69	94
Malgrat de Mar [WT]	41.64895	2.75766	2	48	47
Vilassar de Mar [DD]	41.51354	2.38785	44	55	84
Anglès [DN]	41.96276	2.63221	150	378	379
Girona - Bombers [DM]	41.96196	2.80805	72	148	154
Castell d'Aro [DO]	41.81042	3.03352	14	60	97
la Bisbal d'Empordà [DF]	41.97927	3.03649	28	95	101
Portbou [D6]	42.43690	3.16757	196	83	103
Roses [D4]	42.27244	3.18286	24	73	50
Sant Pere Pescador [U2]	42.17907	3.09788	4	13	18
Banyoles [DJ]	42.11841	2.79085	176	220	187
Olot [DC]	42.19323	2.48375	421	718	619
Muntanyola [CY]	41.88002	2.17980	816	724	682
Orís [CC]	42.07580	2.20976	626	764	755
Viladrau [WS]	41.84191	2.41989	953	786	744
Molló [CG]	42.37920	2.41597	1405	1503	1359
Núria (1.971 m) [DG]	42.40047	2.15632	1971	1944	1981
Sant Pau de Segúries [CI]	42.26014	2.36536	852	1103	979
Caldes de Montbui [X9]	41.61446	2.16938	176	332	311

Sant Llorenç Savall [VV]	41.68320	2.02770	528	509	511
Castellnou de Bages [U4]	41.83278	1.84804	507	384	394
el Pont de Vilomara [R1]	41.70862	1.87339	210	384	398
Sant Salvador de Guardiola [CL]	41.67585	1.76900	349	385	428
Guardiola de Berguedà [WV]	42.23597	1.87593	788	1443	1305
la Quar [CR]	42.08216	1.96355	873	743	756
Santuari de Queralt [WM]	42.10907	1.82857	1167	942	949
Das [DP]	42.38791	1.86750	1097	1663	1553
la Seu d'Urgell [CD]	42.37267	1.43386	849	1412	1347
Organyà [CJ]	42.21816	1.33241	566	1121	1128
Clariana de Cardener [CA]	41.95554	1.58616	693	658	678
els Hostalets de Pierola [CE]	41.53290	1.80924	316	304	321
Òdena [H1]	41.58824	1.65417	333	474	473
Font-rubí [DI]	41.43475	1.62493	415	428	416
Vila-rodona [DQ]	41.30911	1.36369	287	348	348
l'Espluga de Francolí [CW]	41.39427	1.10001	446	565	603
Cervera [C8]	41.67758	1.29689	554	575	566
Tàrraga [C7]	41.66880	1.16335	427	447	439
Castellnou de Seana [C6]	41.65845	0.95275	264	247	253
Alguaire [X3]	41.74463	0.53679	370	262	263
Raimat [VK]	41.68531	0.46178	286	219	222
Seròs [VL]	41.46229	0.40818	100	217	187
Vilanova de Meià [CQ]	41.99731	1.02678	594	745	691
la Pobla de Segur [CV]	42.24111	0.96552	513	1022	903
Montsec d'Ares (1.572 m) [WQ]	42.05316	0.73078	1572	821	769
Sant Romà d'Abella [CP]	42.13924	1.03895	690	911	897
el Pont de Suert [CT]	42.39957	0.74482	823	1368	1351
Salòria (2.451 m) [ZB]	42.52122	1.36704	2451	1877	1802
Vielha e Mijaran [CU]	42.70045	0.79499	1002	1922	1891

ANNEX II – Diagrames de Taylor



ANNEX III – Valor dels coeficients beta

Valors dels paràmetres beta pel primer període (Gener - Juny 2008):

CODI ESTACIÓ	RMSE (°C)	β_0 (°C)	β_1	β_2	β_3	β_4
UU	1.6970	0.6231	0.3925	0.3391	0.2409	-0.3417
DL	1.4499	-0.0936	0.6339	0.2498	0.1161	-0.3995
C9	1.9957	1.3488	0.5666	0.2385	0.1615	-0.3413
DB	1.7483	-0.1365	0.6465	0.2704	0.0648	-0.4645
X5	1.7400	-3.7820	0.7318	0.3583	-0.0976	-0.4996
D8	1.7641	-0.1491	0.6302	0.3209	0.0276	-0.4273
D7	2.4455	0.6508	0.6753	0.2386	0.1234	-0.3454
X1	1.7895	0.1978	0.6957	0.3091	0.0042	-0.5068
D1	2.0416	0.3605	0.7843	0.2387	0.0203	-0.4424
VQ	1.9314	0.0028	0.6373	0.2916	0.0846	-0.3941
DK	1.5478	0.1843	0.7131	0.2452	0.0642	-0.3944
WZ	1.8961	-0.2442	0.8139	0.2198	0.0036	-0.5996
D9	2.0125	0.0035	0.7816	0.1365	0.1096	-0.4382
D3	1.6947	0.3681	0.5249	0.4226	0.0224	-0.4073
D5	1.7836	-1.8223	0.6729	0.3228	0.0034	-0.5098
X4	1.3565	1.1508	0.8225	0.1683	0.0026	-0.4706
UP	1.4896	0.7699	0.6123	0.3279	0.0389	-0.3583
WT	1.9861	-0.8264	0.7929	0.1894	0.0581	-0.4986
DD	1.3836	0.8423	0.6083	0.3228	0.0511	-0.2914
DN	2.2810	1.1600	0.7308	0.1134	0.1451	-0.4304
DM	2.0417	0.5088	0.5889	0.2563	0.1494	-0.4191
DO	2.7246	-3.5813	0.6613	0.1231	0.3791	-0.2114
DF	2.4955	0.5014	0.6419	0.2389	0.1133	-0.4112
D6	1.6602	0.3482	0.5947	0.4344	-0.0448	-0.2152
D4	2.0506	0.7791	0.5394	0.3152	0.1546	-0.2548
U2	2.1860	0.1898	0.3919	0.3879	0.1986	-0.2869
DJ	1.8542	0.5164	0.5751	0.3055	0.0973	-0.4169
DC	2.4575	1.1644	0.6551	0.2000	0.1638	-0.4578
CY	1.7045	-0.6315	0.3604	0.5723	0.0754	-0.4052
CC	2.1017	0.1347	0.5129	0.2994	0.2189	-0.3719
WS	1.8328	-1.4060	0.5587	0.4361	0.0158	-0.5565
CG	1.8286	0.4874	0.5477	0.4436	-0.0037	-0.4717
DG	1.8244	0.5950	0.3291	0.5792	0.0887	-0.3361
CI	2.4608	0.4877	0.6475	0.1361	0.2287	-0.3679
X9	1.9541	0.2895	0.6567	0.2806	0.0914	-0.4932
VV	2.3618	-0.0794	0.6529	0.2908	0.0415	-0.5606
U4	1.7754	-0.8701	0.4556	0.4244	0.1206	-0.3993
R1	2.2096	0.4375	0.6666	0.2130	0.1492	-0.4438
CL	2.2345	-0.7443	0.6693	0.2307	0.1429	-0.4217
WV	2.5966	3.4040	0.6880	-0.0070	0.3668	-0.2422
CR	1.9789	-1.4341	0.3818	0.5057	0.1325	-0.3106

WM	1.8780	-0.7340	0.2481	0.7511	-0.0144	-0.5026
DP	2.6338	2.7647	0.7373	0.0006	0.3138	-0.3208
CD	2.2769	3.0469	0.4558	0.3289	0.2746	-0.3452
CJ	2.5678	3.0446	0.5271	0.2442	0.3097	-0.2252
CA	1.8954	-0.4027	0.4481	0.5056	0.0675	-0.3670
CE	1.5661	0.3639	0.5249	0.3772	0.0838	-0.3679
H1	2.2758	0.4871	0.6601	0.2613	0.1031	-0.4659
DI	1.7597	0.7914	0.4870	0.4170	0.0679	-0.3780
DQ	2.0320	0.7888	0.5313	0.2932	0.1722	-0.3105
CW	1.9626	1.3327	0.6682	0.2046	0.1076	-0.3922
C8	1.7691	-0.1746	0.5415	0.4702	0.0116	-0.5753
C7	1.7838	-0.1629	0.4969	0.5168	0.0051	-0.5970
C6	1.9974	-0.7448	0.3656	0.3487	0.3132	-0.3052
X3	1.7384	-1.1200	0.5537	0.4255	0.0275	-0.5231
VK	1.8844	-1.2538	0.5199	0.4241	0.0603	-0.4411
VL	2.5104	-0.1063	0.4519	0.3172	0.2502	-0.1859
CQ	2.2539	0.0467	0.5797	0.3445	0.1208	-0.4173
CV	2.5198	1.3857	0.5099	0.2609	0.3158	-0.3444
WQ	1.9650	-4.5847	0.2174	0.7550	-0.0061	-0.6508
CP	2.2246	0.9791	0.6141	0.3359	0.1220	-0.4210
CT	2.6295	2.7175	0.8841	-0.2014	0.3447	-0.2913
ZB	1.8601	-4.7219	0.0881	0.8971	0.0359	-0.4469
CU	2.2355	6.1845	0.7458	0.0497	0.2094	-0.3396

Valors dels paràmetres beta pel segon període (Juny - Desembre 2008):

CODI ESTACIÓ	RMSE	β_0	β_1	β_2	β_3	β_4
UU	1.6581	0.9908	0.4324	0.3601	0.1641	-0.1945
DL	1.4802	-0.2568	0.6186	0.2614	0.1311	-0.1434
C9	1.7636	1.5017	0.3291	0.5272	0.0982	-0.2015
DB	1.6659	0.3378	0.5323	0.3641	0.0692	-0.2649
X5	1.6846	-3.7275	0.5595	0.4540	-0.0107	-0.4164
D8	1.7783	-0.1156	0.4844	0.3881	0.1056	-0.2613
D7	2.1822	0.5623	0.4671	0.4404	0.1537	-0.3038
X1	1.7806	0.5517	0.4926	0.4488	0.0603	-0.4010
D1	2.0879	0.5256	0.6298	0.3328	0.0682	-0.3607
VQ	1.8157	0.3276	0.5538	0.2937	0.1583	-0.3615
DK	1.5401	0.5201	0.6209	0.2148	0.1549	-0.2468
WZ	1.7394	-0.5539	0.6936	0.1951	0.1822	-0.3618
D9	1.8338	-0.7781	0.7295	0.1966	0.1429	-0.3369
D3	1.5419	0.5795	0.4556	0.4085	0.0940	-0.3787
D5	1.6009	-1.4471	0.4721	0.3969	0.1201	-0.3814
X4	1.2200	0.6195	0.7720	0.1613	0.1039	-0.3482
UP	1.3870	0.7760	0.5198	0.3289	0.1353	-0.2635

WT	1.9676	-0.8263	0.7228	0.0950	0.2321	-0.2321
DD	1.2509	0.6168	0.5112	0.3046	0.1758	-0.2423
DN	2.2296	1.2406	0.5869	0.2206	0.1898	-0.3238
DM	1.8405	0.4390	0.4564	0.3373	0.2065	-0.2970
DO	1.8611	-0.8119	0.6965	0.0956	0.2774	-0.1833
DF	2.2093	0.4428	0.5492	0.1655	0.2885	-0.1603
D6	1.4937	0.2131	0.4202	0.5308	0.0444	-0.2621
D4	1.9420	0.0199	0.7064	0.1921	0.1627	-0.2051
U2	2.3077	-0.0787	0.5467	0.2914	0.1596	-0.2862
DJ	1.7490	0.5956	0.4441	0.3779	0.1526	-0.2678
DC	2.1757	1.1833	0.5074	0.3242	0.1860	-0.3598
CY	1.5557	-0.3268	0.2445	0.6347	0.0981	-0.3210
CC	1.9914	0.2692	0.4748	0.3170	0.2315	-0.3483
WS	1.8492	-1.3892	0.3726	0.5354	0.0873	-0.3384
CG	1.7633	0.2828	0.4705	0.5297	-0.0020	-0.4203
DG	1.8661	0.4957	0.5021	0.4794	-0.0136	-0.4550
CI	2.2326	0.6317	0.5782	0.1409	0.2839	-0.3036
X9	1.8652	0.2235	0.5094	0.3131	0.2232	-0.3709
VV	2.1971	-0.2620	0.5596	0.3332	0.1009	-0.4140
U4	1.7324	-0.4102	0.3737	0.4698	0.1170	-0.3609
R1	2.1117	0.5673	0.5692	0.2625	0.1964	-0.4313
CL	2.1109	-0.1603	0.5492	0.2728	0.2017	-0.3805
WV	2.4090	3.4241	0.6333	0.0565	0.3407	-0.3093
CR	1.7471	-1.3816	0.3181	0.5750	0.1089	-0.3079
WM	1.8082	-0.7118	0.1555	0.7568	0.0427	-0.3858
DP	2.8339	2.9296	0.6724	0.0593	0.2869	-0.3939
CD	2.2433	3.2878	0.4923	0.2777	0.2468	-0.4017
CJ	2.4942	3.2644	0.5253	0.2443	0.2677	-0.3685
CA	1.6736	-0.3130	0.3736	0.5248	0.1070	-0.3195
CE	1.5132	0.5787	0.3628	0.4983	0.1090	-0.3706
H1	2.0713	0.5428	0.5549	0.2968	0.1726	-0.3837
DI	1.5798	1.3096	0.3368	0.4229	0.1666	-0.2800
DQ	1.7795	0.6042	0.3085	0.4896	0.2280	-0.2145
CW	1.8007	1.4553	0.5795	0.3122	0.0832	-0.3848
C8	1.8860	0.1934	0.3395	0.5886	0.0766	-0.4928
C7	1.6849	0.1510	0.3351	0.6013	0.0715	-0.5185
C6	2.2085	-0.3593	0.4371	0.3702	0.2033	-0.4185
X3	1.6954	-0.8147	0.3548	0.5273	0.1164	-0.4364
VK	2.0225	-0.7913	0.4077	0.4605	0.1353	-0.3779
VL	2.3028	0.8821	0.4165	0.2926	0.2801	-0.1568
CQ	2.2110	0.5757	0.4379	0.4729	0.1058	-0.4239
CV	2.3143	1.7747	0.4330	0.3513	0.2734	-0.4337
WQ	1.8947	-4.7137	0.1835	0.7563	0.0066	-0.6292
CP	2.0498	1.4989	0.4978	0.4184	0.1263	-0.4241
CT	2.6163	2.8459	0.8104	-0.0543	0.2546	-0.4473
ZB	1.8016	-4.7153	0.2579	0.7890	-0.0601	-0.5754
CU	2.3041	6.8570	0.9089	-0.0999	0.1708	-0.3134

