Theoretical and Applied Climatology (2024) 155:1777–1792
https://doi.org/10.1007/s00704-023-04725-5

ORIGINAL PAPER

Comparison of linear, generalized additive models and machine learning algorithms for spatial climate interpolation

Josep Bonsoms[1] · **Miquel Ninyerola**[2]

Received: 3 February 2022 / Accepted: 23 October 2023 / Published online: 8 November 2023

Abstract
Geospatial atmospheric data is the input variable of a wide range of hydrological and ecological spatial models, many of which are oriented towards improving the socioeconomic and environmental sustainability. Here, we provide an evaluation of machine learning (ML) methods for the spatial interpolation of annual precipitation, minimum and maximum temperatures for a mountain range, in this case, the Pyrenees. To this end, this work compares the performance and accuracy of multiple linear regressions (MLR) and generalized additive models (GAM) against five ML methods (K-Nearest Neighbors, Supported Vector Machines, Neural Networks, Stochastic Gradient Boosting and Random Forest). The ML algorithms outperformed the MLR and GAM independently of the predictor variables used, the geographical sector analyzed or the elevation range. Overall, the differences between ML algorithms are negligible. Random Forest shows a slightly higher than average accuracy for the spatial interpolation of precipitation ($R^2 = 0.93$; MAE = 70.44 mm), whereas Stochastic Gradient Boosting is the best ML method for the spatial interpolation of the mean maximum annual temperature ($R^2 = 0.96$, MAE = 0.43 ºC). Stochastic Gradient Boosting, Neural Networks and Random Forest have similar performances for the spatial interpolation of the mean minimum annual temperature ($R^2 = 0.98$, MAE = 0.19 ºC). Results presented here can be valuable for the past and future climate spatial analysis, environmental niche modelling, hydrological projections, and water management.

1 Introduction

Spatial atmospheric data of mountain areas is essential for meteorological prediction and natural hazard identification (Pozdnoukhov et al. 2009). Over long timescales, spatial climate data are necessary for evaluating the climate spatiotemporal patterns and trends (e.g., Beniston et al. 2010). From a hydro-climatological perspective, spatial atmospheric data are the input variable of many hydrological models (e.g., Pelliciotti et al. 2008) that are necessary for glacier modelling (e.g., Huss and Fischer 2016) and estimating future mountain water resources (e.g., López-Moreno al., 2009), which represent significant water resources for the downstream areas (Viviroli et al. 2007). Furthermore, accurate spatial climate data are required for a wide range of environmental disciplines, such as determining ecological niches and evaluating forest areas under risk of fire (e.g., Turco et al. 2018).

The accuracy of the spatial interpolation of climate variables in mountainous zones is variable depending on the temporal set defined or the month of the year included (e.g., Ninyerola et al. 2000; 2005; 2007a; b). The complex mountain topography generates cold pools and non-linear oscillations of air temperature (e.g., Brunetti et al. 2006; Frei 2013). In addition, the different ways mountains are exposed to the main air masses causes rain shadow effects (e.g., Bonsoms et al. 2021a). The geographical factors pointed out, together with the lack of meteorological stations at high elevations, decrease the accuracy of the spatial climate data in short distances (Daly et al. 2008). Hence, a detailed evaluation of the spatial interpolation methods in mountain zones is essential.

Climate variables, namely minimum and maximum annual temperature (Tmin and Tmax, respectively) and precipitation (PP), are usually spatially interpolated using the regression methodology proposed by Agnew and Palutikof (2000) at the beginning of the twenty-first century. In their study, they introduced geographical data (elevation, distance to the sea, latitude

✉ Josep Bonsoms
josepbonsoms5@gmail.com

[1] Department of Geography, Universitat de Barcelona, Barcelona, Spain

[2] Department of Animal Biology, Plant Biology and Ecology, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Catalunya, Spain

and longitude) in a Geographical Information System (GIS) and performed a multiple lineal regression (MLR) to successfully produce climate maps of the Mediterranean Basin. The good results obtained promoted the diffusion of the method. Generally, the MLR model is combined with global, local, geostatistical and hybrid methods for mapping purposes (Burrough and McDonnell 1998b, a; Vicente-Serrano et al. 2003; Peña-Angulo et al. 2016). For instance, Ninyerola et al. (2000; 2007a; b) generated the Digital Climate Atlas of the Iberian Peninsula based on the spatial interpolation of the MLR residuals. Similarly, Peña-Angulo et al. (2016) performed an MLR with a Kriging interpolation of Tmin and Tmax for the entire Iberian Peninsula. In addition, the Digital Climate Atlas of Andorra was obtained with the spatial interpolation of the MLR residuals, using the Inverse Distance Weighting (IDW) for mapping temperature and the Splines for PP (Batalla et al. 2016). The MLR-based method (statistical model and independent variables) has been modified over the years and applied to different predicted variables. For example, in the Pyrenees, López-Moreno and Nogués-Bravo (2005) used the Generalized Additive Models (GAM) for modelling the spatial distribution of snow depth. Later, a comparison of the performances of GAM and MLR showed better results with the first method (López-Moreno and Nogués-Bravo 2006). The same conclusions were detected with other climate variables. The GAM and the Generalized Linear Models (GLMs) provided better results than the linear methods for modeling the spatial distribution of temperature and precipitation (López-Moreno et al. 2007), evapotranspiration (Vicente-Serrano et al. 2007) and fog (Vicente-Serrano et al. 2010). Some advances have also been made regarding the independent variables used. For instance, the combination of remote sensing and vegetation indexes was introduced in the MLR equations (e.g., Cristobal et al. 2008; Mira et al. 2017). The best results in terms of accuracy (RMSE) and performance (R^2) for temperature and PP are usually found at a daily scale, when the geographical data and the circulation weather types (CTs) are combined. This was observed for the spatial interpolation of temperature and PP, based on an MLR (Esteban et al. 2009; Lemus-Canovas et al. 2018). Subsequently, the comparison of MLR with GAM and GLMs showed that better results were obtained with the last two methods (Lemus-Canovas et al. 2019).

Comparisons between mapping methods, namely IDW and Kriging (Ninyerola et al. 2000) and IDW and Splines (Ninyerola et al. 2007a; Ninyerola et al. 2007b; Esteban et al. 2009; Batalla et al. 2016; Lemus-Canovas et al. 2018) revealed negligible differences. Hence, further efforts for decreasing the spatial climate data interpolation error should be focused on improving the regression model. In this sense, the linear-based methods have some limitations. Linear models can underestimate climate data because some climate variables, such as PP, do not show linearity with elevation (e.g., Henn et al. 2019). In the last decades, machine learning (ML) methods have been introduced in the climatological science, showing promising results. However, no study to date has compared ML methods with MLR or GAM models for the spatial interpolation of climatological (i.e., > 30 years) data (Tmin, Tmax and PP) for an entire mountain range, or analysed the accuracy by climate sectors of a mountain range. Therefore, in this study we aim to address this knowledge gap by evaluating five tuned ML algorithms, and comparing the results with the ones obtained with MLR and GAM in the Pyrenees Mountain range as a study case. The ML methods included in this work are classified into three groups: (i) ML techniques based on non-linear approaches, such as the K-Nearest-Neighbours (KNN) and the Supported Vector Machines with radial basis kernel (SVM); (ii) ML based on Neural-Network (NN) techniques and (iii) ML based on ensemble regression trees, using boosting techniques, such as the Stochastic Gradient Boosting (GMB) or bagging techniques, such as Random Forest (RF). Topographically complex sectors have different climate patterns, and the accuracy of each method can depend on the spatial Pyrenean sector (Vicente-Serrano et al. 2003). Hence, we determined which method is the best for the entire mountain range with climate clusters from the mountain range (CL) and for different elevation ranges. This study addresses the following objectives:

(i) To compare the performance of seven different methods for the spatial interpolation of annual Tmin, Tmax and PP in a mountain range, in this case the Pyrenees, for the 1981–2015 period.
(ii) To analyse the accuracy and performance of the spatial interpolation methods as a function of CL and elevation range.

The article is structured as follows. In Sect. 2 the study area is presented. Section 3 provides a description of the data and the methods used. The results and discussion are presented in Sect. 4. Finally, Sect. 5 summarises the main conclusions.

2 Study area

The Pyrenees are located in the south-west of Europe, between the north of Spain and the south of France, at latitudes ranging from 42 to 43 ºN and longitudes between 2ºW and 3ºE. The highest elevations are found in the central area of the mountain system (Fig. 1), with mountain peaks exceeding 3000 m (e.g., Aneto, 3404 m.a.s.l).

The Pyrenees mountain range encompasses different types of climates (Bonsoms et al. 2021b), supplies most of the annual runoff of the Ebro River Basin (López-Moreno et al. 2009) and shows marked geoecological differences (Oliva et al. 2016). The climate is governed by a prevailing westerly circulation during the cold half of the year and a relative disconnection of the general circulation moving into the eastern Iberian Peninsula, due

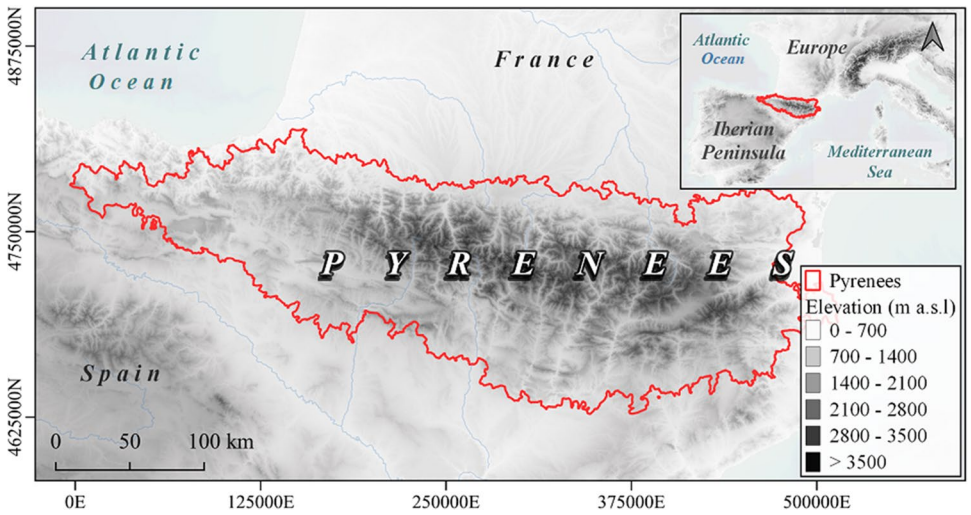

Fig. 1 Map of the study area analyzed in this work. The limits of the Pyrenees are defined according to the CLIMPY project (https://www.opcc-ctp.org/en/climpy). The coordinate reference system is ETRS89 (UTM 30N)

to the topographical effects and the west–east alignment of the mountain zones (e.g., Martín-Vide and López-Bustins 2006). In the western and central zone of the southern slopes, PP from December to March is governed by negative phases of the North Atlantic Oscillation (NAO) and westerly flows (Buisán et al. 2016). The influence of the NAO decreases towards the northern and southern slopes of the eastern sector of the mountain range (Alonso-González et al. 2020a, b). In the eastern area, negative phases of the Western eastern Mediterranean Oscillation (WeMO), related to south-eastand east Mediterranean advections, govern the PP timing (Martín-Vide and López-Bustins 2006; López-Bustins and Lemus-Cánovas, 2018). The different low-frequency climate patterns that govern the study zone are in turn translated into different PP regimes and timings. Between December and March, PP is ca. > 600 mm on the southern slopes of the western area; being < 400 mm in the central area and < 200 mm in the eastern strip (Buisán et al. 2016). At a local scale, small scale variations of PP and temperature are observed. Likewise, on the valleys, basin scale-inversions during winter with stable weather conditions are frequent (e.g., Pepin and Kidd 2006). The above-mentioned climatological contrasts at regional scale suggest that it is required to analyze each spatial interpolation method depending on elevation and CL.

3 Data and methods

3.1 Data

3.1.1 Observed dataset

To avoid large uncertainties related to the climate data quality and homogenization, in this work we used the CLIMPY (Characterization of the evolution of climate and provision of information for adaptation in the Pyrenees) gridded dataset (available at https://zenodo.org/record/3611127#.YLszfkzY2x). CLIMPY is a transnational climate project that analyzed the evolution of the Pyrenean climate during the last half of the century, providing a grid (1 km × 1 km; over more than 50,000 km^2, Fig. 1) of PP, Tmin and Tmax for the whole mountain range. The dataset was created following a reconstruction, gap filling and quality control process detailed in Serrano-Notivoli et al. (2017). The temporal period analyzed encompasses all days from 1981 to 2015. The Meteorological records are managed by the Spanish National Agency (AEMET), Météo-France (MF) and the Meteorological Service of Catalonia (SMC).

3.1.2 Predictor variables

The independent variables used in previous works in this area are generally based on geographical data (e.g., Ninyerola et al. 2000; Vicente-Serrano et al. 2003; Esteban et al. 2009; Batalla et al. 2009; Lemus-Canovas et al. 2018 and 2019). Other variables, such as the topographic wetness index or the potential solar radiation could be included. However, no significant contribution for the spatial interpolation of climate variables has been observed (see, for example, Mira et al. 2017), and therefore they have been excluded from this work.

The initial variables used in this article were the normalized values of elevation, distance to the coast (continentality), latitude and longitude (e.g., Hofstra et al. 2008). All of the independent variables are based on a Digital Elevation Model (DEM) with 90-m of spatial resolution, downloaded

from the Shuttle Radar Topography Mission (SRTM). Continentality was generated using the Euclidean distance to the coast line following Ninyerola et al. (2000). A Pearson correlation test (r) was performed between the predictor variables to avoid multicollinearity. Then, continentality was excluded due to the high and statistically significant correlation with longitude ($r = -0.9$, $p < 0.05$ and $r = 0.9$, $p < 0.05$, respectively). The other independent variables do not show a relevant correlation ($r < 0.4$, $p < 0.05$). Therefore, the predictor variables are elevation, latitude and longitude.

3.2 Methods

The PP, Tmin and Tmax values (daily resolution) were averaged for each specific grid site and year. Subsequently, the data were randomly split into two groups: the training and the testing dataset. Cristobal et al. (2008) used 60% of the dataset for training the methods and the remaining for testing. On the other hand, Feng et al. (2019) split the data into 80% for training the ML methods, and the other 20% for testing. Meyer et al. (2016a, b) compared four ML methods using 33% of the data for training the algorithms and the remaining for validation purposes. We divided the data into the same proportions as this last study since we had enough climatological data to do so (Table 2).

3.2.1 Statistical and Machine Learning methods

This section presents an overview of each method. A detailed methodological description can be consulted in the supplementary material.

The MLR assumes a linear regression between the independent variables and the dependent variable. GAM is a non-parametric extension of the generalized linear models with smoothing functions to fit non-linear responses of the independent variables (Hastie and Tibshirani 1987). We included five ML algorithms: (i) the KNN (Bishop, 1995), based on the Minkowski distance of the training data points and assuming a rectangular kernel; (ii) the NN, a ML based on a distributed system of neurons that creates a wide range of non-linear functions with more than one interconnected layer (Haykin 1998). We also included (iii) the SVM, based on an exponential radial bias function (Vapnik, 1998). Regarding the regression trees, we included (iv) GBM (Friedman 2001), a boosting method that generates an additive model that reduces the loss function. Finally, (v) we added the RF, a ML algorithm based on decision trees and bootstrap aggregation (Breiman 2001).

Previous literature showed that tuning the hyperparameters of the ML algorithms provides better results than the default ML parameters (e.g., Tripathi et al. 2006). Then, we parameterized the GAM and tuned the ML algorithms (Table 1), according to a tenfold cross-validation based on a grid selection, using the caret package (Kuhn 2020) of R (R Core Team 2018). Subsequently, the methods were applied to an independent test dataset to check the accuracy and performance metrics. The detailed error metrics of the different hyperparameters tested in the training dataset can be consulted in the supplementary materials. No significant differences were found between the training and testing datasets, which evidences that there was no ML overfitting.

The distribution of the data by elevation range is presented in Table 2 and discussed in the results and discussion section.

3.2.2 Evaluation of the methods

The performance of the methods was evaluated using five types of accuracy metrics: the Residuals (Eq. 1), Mean Absolute Error (MAE; Eq. 2), Root Mean Square Error (RMSE; Eq. 3), agreement index (Willmott's D; Eq. 4) and coefficient of determination (R^2; Eq. 5). The MAE and RMSE summarize the mean differences between the predicted and observed values, showing low values when the accuracy is high. On the contrary, high R^2 and Willmott's D are related to high levels of performance, and the latter is less sensitive to outliers than the other accuracy parameters (Willmott 1982).

$$\text{Residual} = Oi - Ei \tag{1}$$

$$\text{MAE} = \frac{1}{\text{N}} \sum_{i=1}^{\text{N}} |Ei - Oi| \tag{2}$$

$$\text{RMSE} = \sqrt{\frac{1}{\text{N}} \sum_{i=1}^{\text{N}} (Ei - Oi)^2} \tag{3}$$

$$\text{Willmott's } D = 1 - \sqrt{\frac{\sum_{i=1}^{\text{N}} (Ei - Oi)^2}{\sum_{i=1}^{\text{N}} (|E'i| - |O'i|)^2}} \tag{4}$$

$$R^2 = \left[\frac{\sum_{i=1}^{\text{N}} (Oi - \overline{O}) \sum_{i=1}^{\text{N}} (Ei - \overline{E})}{\sqrt{\sum_{i=1}^{\text{N}} (Oi - \overline{O})^2} \sqrt{\sum_{i=1}^{\text{N}} (Ei - \overline{E})^2}} \right]^2 \tag{5}$$

where N is the number of samples, Ei is the predicted value and Oi is the observed values. $\overline{O}$ is the average of the observed values and $\overline{Ei}$ represents the mean of the predicted values. $E'i$ is the difference between Ei and $\overline{O}$, whilst $O'i$ is the difference between Oi and $\overline{O}$.

3.2.3 K-means unsupervised classification

The spatial regionalization of the Pyrenees was performed using the objective K-means method. The K-means is a non-hierarchical cluster analysis method introduced by Forgy

Table 1 Description of the methods used in this work and ML hyperparameters selected

Methods		Hyperparameters tested	Tmin	Tmax	PP	Time*
Statistical	Multiple Lineal Regression (MLR)	X	X	X	X	0.02''
	Generalized Additive Models (GAM)	X	X	X	X	31''
ML	*K*-Nearest Neighbors (KNN)	Kmax From 5 to 43, with increments of 2	21	11	5	6′25''
	Support vector machine Radial Basis Function (SVM)	Cost 0.25, 0.5, 1	1	1	1	20′67''
		Epsilon	0.1	0.1	0.1	
	Neural Networks (NN)	Size From 10 to 30, with increments of 5	30	30	30	2h57'
		Decay From 0.001 to 0.205, with increments of 0.05	0.005	0.005	0.205	
	Stochastic Gradient Boosting (GBM)	Nº Trees 100, 300 and 500	500	500	500	51′81''
		Variable interaction depth 5, 10 and 15	15	15	15	
		Shrinkage of tree expansion From 0.1 to 1 with increments of 0.2	0.1	0.1	0.1	
		Sampled predictors at node 5, 10 and 15	15	5	10	
	Random Forest (RF)	Nº trees 100, 300 and 500	500	500	500	9h43'
		Ntry 2 to 10 with increments of 1	2	2	2	

* Computing time. RAM = 16 GB

Table 2 Distribution of the training and testing datasets by elevation range

Elevation (m)	Training Points (% of the training data)	Testing Points (% of the testing data)
< = 500	3921 (24.94)	7706 (24.54)
500–1000	6216 (39.53)	12,427 (39.58)
1000–1500	2749 (17.48)	5650 (17.99)
1500–2000	1621 (10.31)	3164 (10.07)
2000–2500	1020 (6.48)	2054 (6.54)
2500–3000	191 (1.21)	390 (1.24)
> = 3000	3 (0.01)	6 (0.01)

(1965) and widely used for determining hydrological and climate zones (e.g., Carvalho et al. 2016; Alonso-González et al. 2020a, b, among other works). The CLs are defined by applying the algorithm to the whole spatial climate dataset. The K-means is based on three steps (e.g., Hartigan and Wong 1979): (i) The K centroids (t_1, …,t_k) are randomly initialized inside the feature area. Subsequently, (ii) the distance of the data objects to the centroids ($d(x_1, tk)$) $k \in [\![1, K]\!], i \in [\![1, N]\!]$ is calculated, and each point is assigned to the nearest centre: $C(i) = arg\ \min_k \{d(t_1, t_k)\}$. Finally, (iii) the centroids are established as the arithmetic average point of the cluster:$t_k = \frac{\sum_{i=1}^{N} 1C(i)=k}{\sum_{i=1}^{N} 1C(i)=k}$. The algorithm was repeated 100 times after the convergence. The spatial clustering allowed us to define different CLs ranging from $k = 2$ to $k = 8$. The optimal CL numbers were selected using the mean of the squared distances between the CL centers (BETWENSS) and the total sum of squares (TOTSS). The BETWENSS and the TOTSS were divided and multiplied by 100, obtaining the explained variance (EV). The number of CL was selected according to the slope change of the scree test (Cattell 1966). For Tmin and Tmax we retained four CL, explaining 88% of the EV. For PP, we retained five CL which explain 93% of the EV (Table 3).

4 Results and discussion

First we present a detailed characterization of the predicted and observed annual values of Tmin, Tmax and PP in the Pyrenees. We provide a fair comparison between the methods included in this work, given that all of them have been analysed with the same initial conditions. To this end, we tested the accuracy of the methods using an independent and randomly selected dataset, which constitutes 66% of the database of the total grid points. We also analyzed the spatial

Table 3 EV of each number of CL. The selected EVs are expressed in bold

Variables	CL2	CL3	CL4	CL5	CL6	CL7	CL8
Tmin and Tmax	**70.2**	**83.2**	**88.2**	90.8	92.2	93.6	94.6
PP	**70.2**	**83.7**	**90.6**	**93.6**	95.3	96.7	97.0

distribution of the accuracy and performance metrics by CLs and by elevation ranges (steps of 500 m), providing a robust evaluation of the interpolation methods.

4.1 Tmin, Tmax and PP analyses on the independent dataset

The basic statistics of Tmin, Tmax and PP are summarized at Table 4. The average observed Tmin is 4.70 ºC. The Standard Deviation (SD) of Tmin is 2.21 ºC, with an annual amplitude higher than 10 ºC. The maximum observed Tmin reaches 9.79 ºC, whereas the minimum Tmin is -1.86 ºC. The MLR and the NN overestimated the minimum Tmin (-3.02 ºC and -2.02 ºC, respectively). The MLR (GAM) underestimated (overestimated) the maximum Tmin values 8.18 ºC (10.07 ºC). The remaining methods reproduced the observed values. RF shows the lowest bias, both for minimum (-1.65 ºC) and maximum Tmin (9.74 ºC).

The average observed and estimated Tmax values are the same: 12.63 ºC. The observed SD is 3.07 ºC, with a minimum Tmax of 2.62 ºC and a maximum of 17.83 ºC. The MLR estimated the lowest minimum Tmax values (2.03 ºC). The contrary is observed with the GAM (3.58 ºC). The ML methods overestimated the minimum Tmax (3.50 ºC, on average). However, all the ML show a good performance in the maximum Tmax estimation (17.83 ºC vs 17.30 ºC, observed and estimated values, respectively).

The average observed and estimated annual PP for the entire Pyrenees is 1060 mm. The SD of all the interpolation methods is approximately 430 mm, except for the MLR and GAM (330 mm and 388 mm, respectively). The geographical sectors of the range where the maximum PP is found (3444 mm) almost double the values estimated by the MLR (1831 mm).

4.2 Evaluation of the accuracy and performance of the methods

Overall, the methods show good accuracy and performance, since they were trained with a large sample size (Table 2). The ML methods outperformed MLR and GAM in the spatial prediction of Tmin, Tmax and PP. For Tmin, the RMSE ranged from 0.28 ºC (KNN, NN, GMB and RF) to 0.62 ºC (MLR). The KNN, NN, GMB and RF show the lowest MAE (0.19 ºC). The same is observed with D; except for GAM and MLR, all the algorithms reach the optimal value (1). The worst R^2 is measured with MLR (0.92) and GAM (0.96). The ML methods show the same performance ($R^2 = 0.98$). The R^2 is slightly lower for Tmin than for Tmax (Table 5). The GMB scores the lowest Tmax error (RMSE = 0.64 ºC, MAE = 0.43 ºC). The worst spatial interpolation method

Table 4 Descriptive statistics of the observed (OBS) and predicted annual Tmin, Tmax and PP values

Variable		OBS	MLR	GAM	KNN	SVM	NN	GMB	RF
Tmin (º C)	Min	-1.86	-3.02	-1.58	-1.27	-1.41	-2.02	-1.64	-1.65
	Avg	4.70	4.70	4.70	4.70	4.70	4.70	4.70	4.70
	Max	9.79	8.18	10.07	9.46	9.72	9.68	9.65	9.74
	SD	2.21	2.12	2.17	2.19	2.19	2.19	2.19	2.19
	Lapse rate (ºC/100 m)	-0.31	-0.35	-0.32	-0.31	-0.30	-0.31	-0.30	-0.29
Tmax (º C)	Min	2.62	2.03	3.58	3.56	3.67	2.26	3.08	3.16
	Avg	12.63	12.63	12.63	12.63	12.62	12.63	12.63	12.63
	Max	17.83	18.26	17.10	17.21	17.25	17.18	17.28	17.21
	SD	3.07	2.94	2.96	3.00	3.01	2.99	3.00	2.99
	Lapse rate (ºC/100 m)	-0.40	-0.47	-0.43	-0.37	-0.38	-0.38	-0.38	-0.38
PP (mm)	Min	339.84	232.13	257.11	356.67	423.25	390.86	202.80	353.71
	Avg	1062.62	1059.57	1063.42	1060.00	1045.11	1060.68	1062.71	1062.27
	Max	3564.03	1831.26	2553.76	3270.38	3126.74	2874.71	3349.97	3257.32
	SD	455.54	330.01	388.31	437.54	414.88	420.33	436.82	435.72
	Lapse rate (mm/100 m)	41.59	30.34	30.34	42.59	45.09	20.64	32.79	32.31

is MLR (RMSE = 0.92 ºC, MAE = 0.70 ºC) and GAM (RMSE = 0.79 ºC, MAE = 0.57 ºC) followed by SVM (RMSE = 0.71 ºC, MAE = 0.48 ºC).

For PP, the RMSE of MLR (315.21 mm) and GAM (239.04 mm) is almost double that measured with the ML algorithms (*ca*. 120 mm; Fig. 2). Three ML methods show the best spatial prediction of PP: the RF (RMSE = 118.89 mm; MAE = 70.44 mm) followed by GMB (RMSE = 128.83 mm; MAE = 82.98 mm) and KNN (RMSE = 136.77 mm; MAE = 80.26 mm). For temperature, the ML methods explained only *ca.* 6% more of the variance than MLR. For PP, there are large differences between the linear methods and ML (Fig. 3). MLR and GAM are not able to predict the highest values of PP, reaching only 52% and 72%, respectively, of the explained variance. However, the ML methods are able to predict *ca.* 40% (20%) more of the variance than MLR (GAM). NN and SVM show good results ($R^2 = 0.85$ and $R^2 = 0.86$, respectively), but worse than those obtained by KNN and GMB ($R^2 = 0.91$ and 0.92, respectively). RF has the best performance ($R^2 = 0.93$).

Our results are in accordance with Duhan and Pandey (2015), who showed that ML methods (i.e., the SVM) are able to better estimate monthly Tmin and Tmax than MLR-based methods for climate downscaling. The ML techniques, such as RF in combination with IDW (Fig. 4), have been successfully applied for environmental modelling (e.g., Li et al. 2013). Regression trees were found to be one of the most precise ML types for the spatial interpolation of temperature. This was observed for the interpolation of ground-based temperature data (Appelhans et al. 2015) as well as for remote sensing products (e.g., Noi et al. 2017; Meyer et al. 2016a, b). The RF model could be improved for the spatial prediction of temperature and PP by incorporating the nearest geographical data as a predictor variable (Hengl et al. 2018), obtaining better results than other geostatistical methods alone, such as IDW or Kriging (Sekulić et al. 2020). The computation training of the ML algorithms requires more time than linear approaches (nearly 10 h for all the variables analyzed, Table 1). However, ML techniques, such as RF, do not rely on linear effects or data distribution (Rodiguez-Galiano et al. 2012) and are robust to outliers (Breiman 2001). Moreover, RF parametrization depends on only two parameters (nº trees and mtry). The reconstruction and the gridding of climate data are based on the same methods (Serrano-Notivoli and Tejedor 2021). For the first of the objectives, KNN is one of the most widely used methods (e.g., Begueria et al. 2016). In this article KNN provided good results, but the algorithm depends on the density of the data. Further works should test the climate reconstruction based on the RF model presented in this work, and include data from the nearest stations (Vicente-Serrano et al. 2010), the average monthly values together with other proxies, for instance, low-frequency climate modes.

4.3 Distribution of the errors by CL and elevation range

Figure 5 shows the Tmin, Tmax and PP lapse rates (by 100 m) and the accuracy metrics grouped by elevation range. For PP, the difference by each 100 m is 41.59 mm/100 m. The average Tmin (Tmax) lapse rate is 0.31 ºC (0.40 ºC)/100 m. MLR and GAM overestimated the lapse rate, whereas the ML methods slightly underestimated it.

PP does not follow a Gaussian distribution and the interpolation accuracy should be evaluated by splitting the data using percentiles or elevation ranges (Serrano-Notivoli and Tejedor 2021). Our results show that the differences between the ML and linear approaches increased by elevation for all the analyzed variables (Fig. 5, 6 and Table S1). For Tmin and Tmax, the largest differences between the observed and MLR estimated values are found at > 2000 m. Whereas for PP, the largest bias is found at > 1500 m. The ML methods show small differences in the accuracy metrics. The most

Table 5 Accuracy and performance values, grouped by variable and method

Variable	METRIC	MLR	GAM	KNN	SVM	NN	GMB	RF
Tmin (ºC)	MAE	0.49	0.33	0.19	0.21	0.19	0.19	0.19
	RMSE	0.62	0.42	0.28	0.30	0.28	0.28	0.28
	D	0.98	0.99	1.00	1.00	1.00	1.00	1.00
	R^2	0.92	0.96	0.98	0.98	0.98	0.98	0.98
Tmax (ºC)	MAE	0.70	0.57	0.43	0.48	0.47	0.43	0.43
	RMSE	0.92	0.79	0.66	0.71	0.69	0.64	0.65
	D	0.98	0.98	0.99	0.99	0.99	0.99	0.99
	R^2	0.91	0.93	0.95	0.95	0.95	0.96	0.96
PP (mm)	MAE	240.18	178.61	80.26	108.55	120.59	82.98	70.44
	RMSE	315.21	239.04	136.77	173.31	178.48	128.83	118.89
	D	0.82	0.91	0.98	0.96	0.96	0.98	0.98
	R^2	0.52	0.72	0.91	0.86	0.85	0.92	0.93

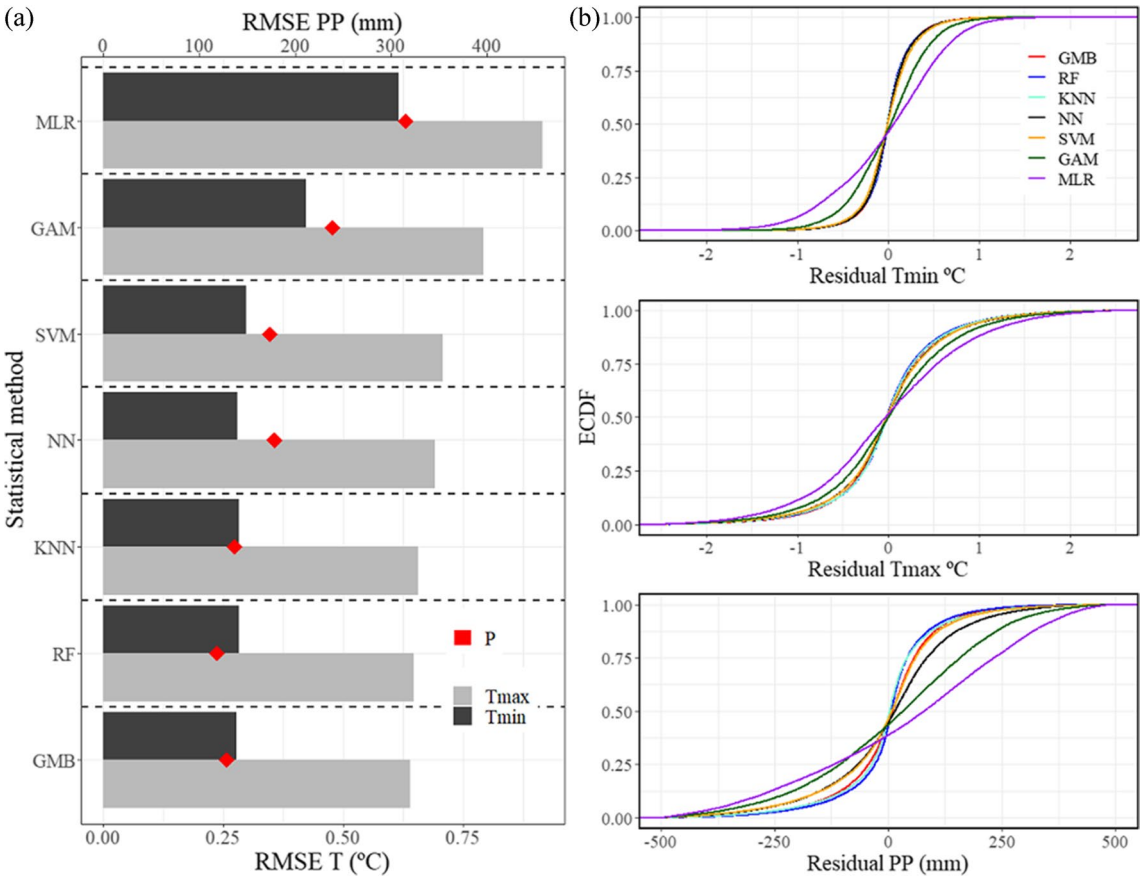

Fig. 2 **a** Average RMSE detected by each spatial interpolation method. **b** Empirical Cumulative Distribution Function (ECDF) of the residual values

accurate method for estimating Tmin is RF (MAE = 0.07 °C), followed by GMB (MAE = 0.10 °C). RF is the most accurate method for Tmax (MAE = 0.40 °C). GAM shows a good performance for PP (MAE = 145.42 mm), similar to ML methods, such as SVM (MAE = 151.84 mm). Nevertheless, the decision trees outperformed the spatial estimation of PP. The best results were obtained by RF (avg. MAE = 77.61 mm) followed by GMB (avg. MAE = 107.48 mm).

The decrease in performance in high-elevation areas could be because more than 80% of the data is found at < 1500 m (Table 2). The results agree with those of Herrera et al. (2019), who showed that the density of the meteorological stations explains a large variance (60%) of the climate interpolation accuracy. The linear approaches underestimate the climate values of the high lands, since they are based on extrapolations of climate data found at low elevations. Henn et al. (2018) compared the accuracy of six gridded PP datasets of the western United States, all of them based on weighted linear approaches. They found significant bias in high-elevation areas, which was mainly due to the lack of data in elevated sectors, inhomogeneities and missing data, together with the PP underestimation caused by undercatch. In the Pyrenees, the majority of the meteorological stations are placed in flat and valley zones (Batalla et al. 2016), and in high-elevation sectors of the eastern Pre-Pyrenees, meteorological instrumental records are only available since the earliest 2000s (Bonsoms et al. 2021b). Therefore, the low accuracy could be explained both by the interpolation methods and by the data uncertainty in elevated areas.

The spatial interpolation performance could be different depending on the spatial scale and the geographical sector of the range (Vicente-Serrano et al. 2003). Therefore, we have provided a spatial evaluation of the accuracy of Tmin, Tmax and PP for Pyrenean areas, based on the K-means unsupervised classification method (Fig. 7). For Tmin, GAM shows good results (RMSE = 0.40 °C), but all the ML methods are more accurate than GAM (RMSE < 0.30 °C) with negligible differences between then (Table 6). The same is observed with Tmax. GMB shows the lowest RMSE (0.64 °C) by CL, but only decimal differences are found in comparison with the other ML algorithms. For PP, the regression trees show the best results for all the CLs. The RF algorithm obtained the lowest error (average RMSE = 117.16 mm). The

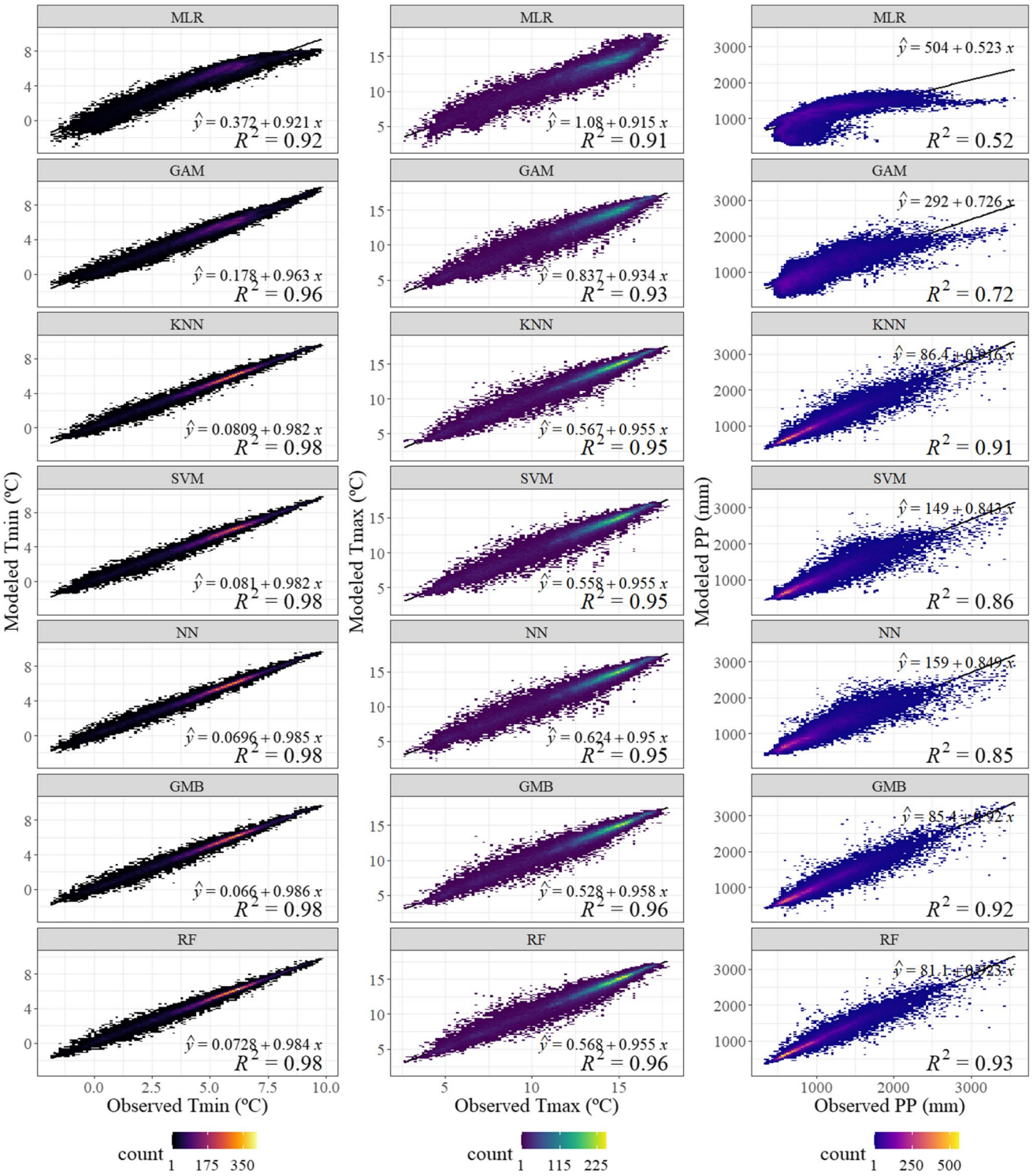

Fig. 3 Density scatterplot of the observed (horizontal axis) and predicted (vertical axis) values grouped by method

minimum error metrics are found in the driest zone (CL1, RMSE = 53.97 mm) and the maximum is found in the wet Atlantic area (CL5, RMSE = 208.07 mm). SVM and NN show a remarkable lack of accuracy in comparison with the regression trees at CL 5 (Fig. 8c). These methods are not able to estimate the extreme values of PP in the wet Atlantic area (RMSE = 304.58 mm and 308.57 mm, respectively). The results provide evidence that the RF algorithm

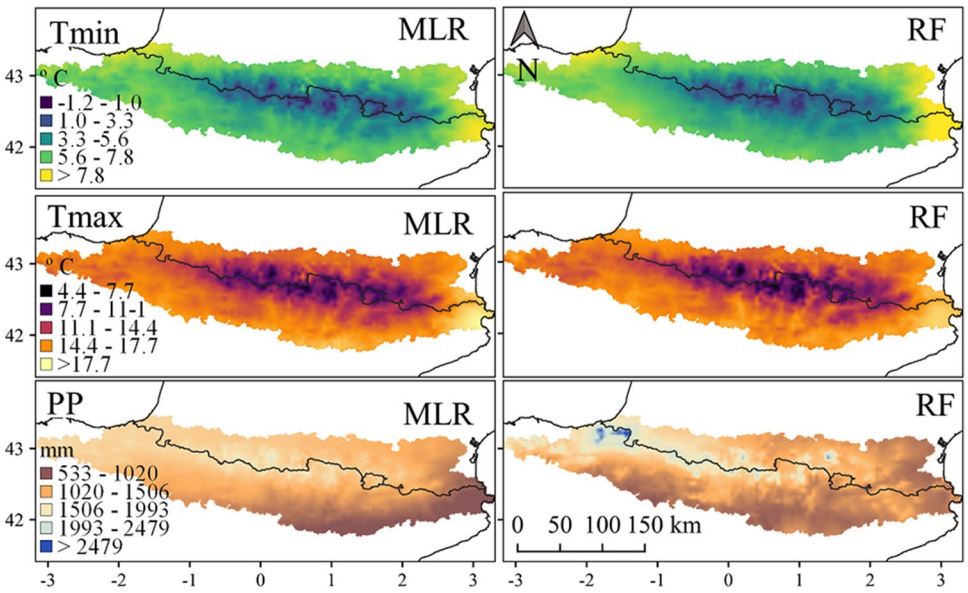

Fig. 4 Comparison of the average yearly values for Tmin, Tmax and PP with MLR (left column) and RF (right column). Mapping was performed with IDW (power = 2). The coordinate reference system is ETRS89 (UTM 30N)

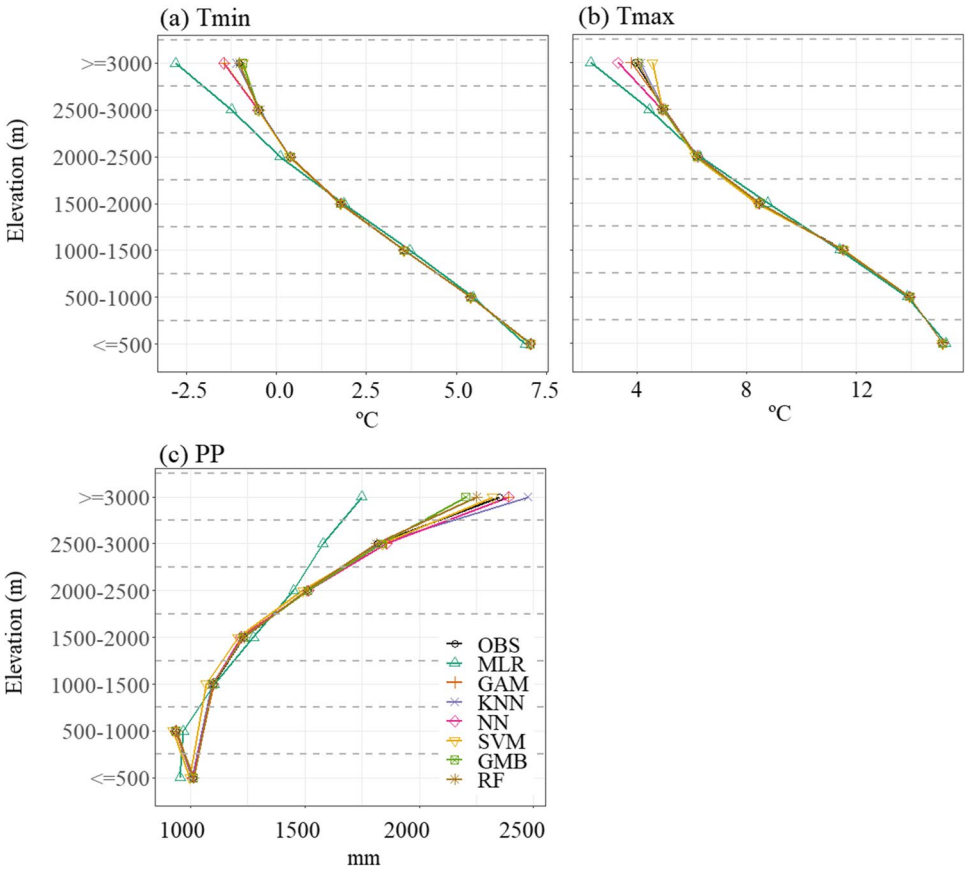

Fig. 5 Averaged values of each variable grouped by elevation and method

performs better than the other ML methods for the spatial prediction of the extreme low (CL 1) and high (Cl 5) values. The error observed at CL5 between the ML methods could be explained by the structure of each algorithm. RF is a non-parametric algorithm that builds a wide range of models, separates the nodes of the trees (mtry) and selects the best features for each point; nevertheless, this parameter decision is not equivalent in SVM (Pelletier et al. 2016). Moreover, that there are no differences between the training and testing datasets suggest that there is no overfitting of RF.

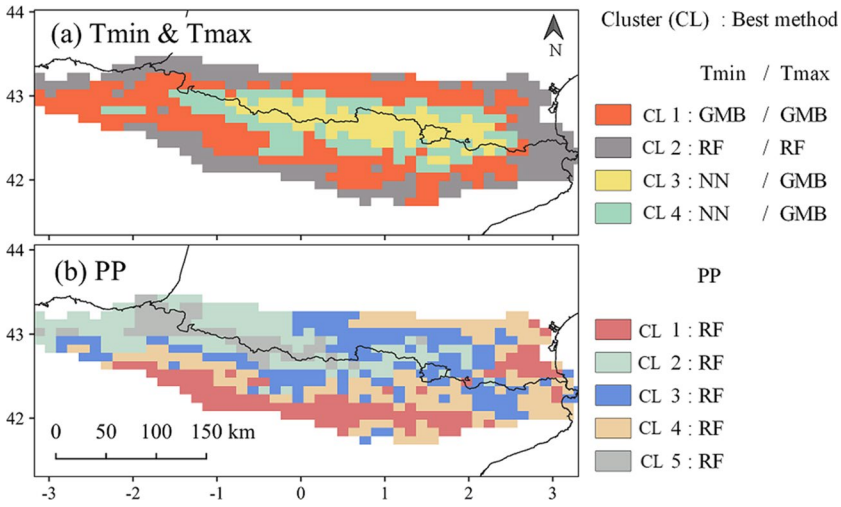

Fig. 6 Spatial representation of the best performance methods grouped by CL. The coordinate reference system is ETRS89 (UTM 30N)

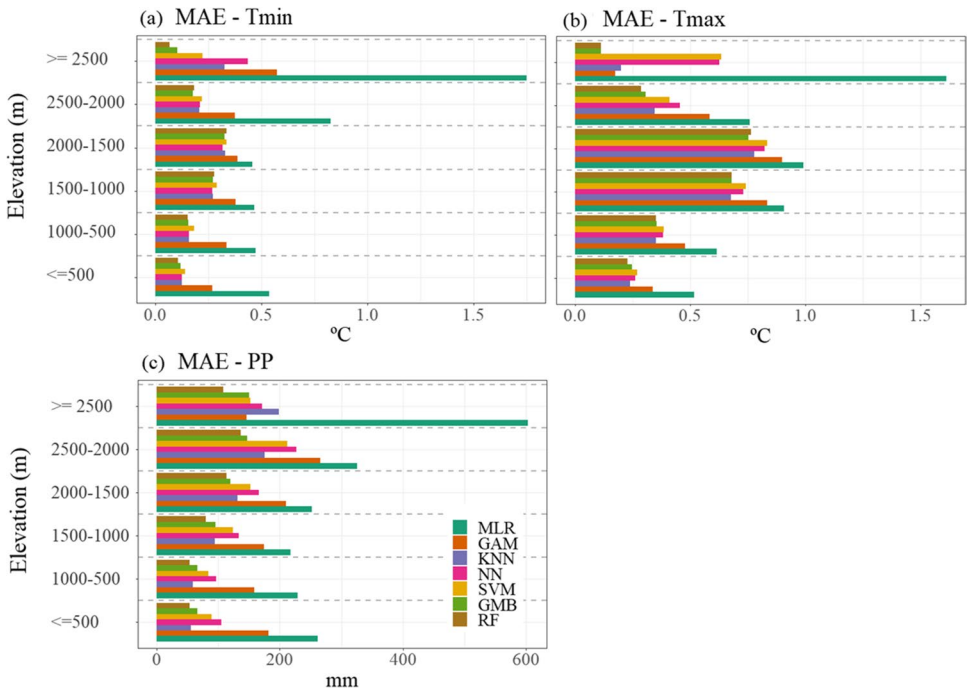

Fig. 7 MAE values of each variable grouped by elevation and method

The residuals are the difference between the observed and predicted values. The spatial representation of the residuals shows how different each CL is regarding the overall model (Fig. 9). The map of the residuals also shows the geographical settings where it is necessary for the meteorological stations to increase representativity (Herrera et al. 2019). For Tmin and Tmax, the K-means clustering followed an elevation pattern. The largest errors are found in the highest elevation area (CL 1). The climate maps based on MLR and GAM produce large errors in the estimation of PP (Fig. 9). A PP underestimation can be found on the leeward side. However, the results show an overestimation of PP values in the driest, rain-shadowed zone of the southern slopes of the Pyrenees (CL 1).

4.4 Variable of importance

The most important predictor variable for estimating PP, Tmin and Tmax is determined using the Percentage of Mean Decrease Accuracy (IncMSE) of RF (Fig. 10). The IncMSE measures the decrease in the model performance when one variable is excluded. The Tmax shows a higher dependence on elevation (IncMSE ≈ 76%) than Tmin (IncMSE ≈ 66%). The temperature could be estimated with the elevation; however, the incorporation of other geographical variables (e.g., latitude; Vicente-Serrano et al. 2003) improves the model's performance. For PP, the largest dependence is observed with

Table 6 RMSE of each method grouped by CL

Variable	Cluster	MLR	GAM	KNN	SVM	NN	GMB	RF
Tmin (ºC)	CL1	0.54	0.39	0.22	0.24	0.22	0.21	0.21
	CL2	0.73	0.39	0.18	0.19	0.17	0.16	0.16
	CL3	0.67	0.48	0.40	0.41	0.40	0.40	0.41
	CL4	0.58	0.46	0.36	0.37	0.35	0.35	0.36
Tmax (ºC)	CL1	0.78	0.66	0.52	0.56	0.55	0.50	0.50
	CL2	0.72	0.49	0.35	0.40	0.38	0.35	0.34
	CL3	1.12	1.00	0.89	0.91	0.91	0.87	0.87
	CL4	1.12	1.03	0.87	0.97	0.94	0.85	0.87
PP (mm)	CL1	239.37	166.89	61.08	80.54	84.42	69.55	53.97
	CL2	239.73	246.58	171.04	210.43	219.40	154.26	143.86
	CL3	275.79	218.58	122.58	152.21	160.21	121.09	106.38
	CL4	284.07	198.68	84.29	107.89	106.17	85.09	73.53
	CL5	576.47	390.10	232.43	304.58	308.57	211.23	208.07

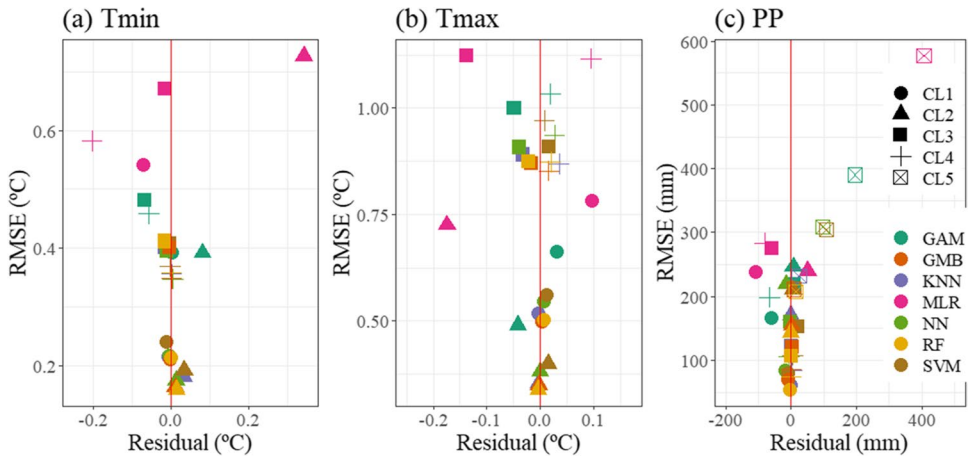

Fig. 8 Distribution of the average residuals (the difference between the observed and the estimated values; horizontal axis) and RMSE (vertical axis). The points are grouped by CL (shape) and method (colors)

latitude (IncMSE ≈ 42%), followed by elevation (IncMSE ≈ 32%) and longitude (IncMSE ≈ 25%).

There is a general decrease in PP towards the eastern strip (Fig. 4). This is because the western Pyrenees rain-shadows the eastern area (Pepin and Kidd 2006). Latitude, instead of elevation or longitude, is the geographical factor that governs PP. Latitude indirectly expresses the north or south face of the mountain range and the exposition of the CTs. On the southern slopes of the Pyrenees, PP is governed by west and south-west CTs during the cold half of the year; in addition, the influence of north and north-west CTs is reduced moving from north to south of the mountain range (Esteban et al. 2009; Lemus-Canovas et al. 2018). In summer the precipitation is convective (Xercavins, 1985), reducing the predictability of PP based on the independent variables included in this work. Similar patterns are observed with temperature. For the spatial interpolation of Tmin and Tmax, Ninyerola et al. (2000) found the lowest accuracy during winter months. Similar results are observed in Cristobal et al. (2008). In the Alps, the maximum error for the interpolation of temperature was also observed in winter (MAE = 1.5 ºC; Frei 2013). However, the highest accuracy is generally observed during the summer months (Ninyerola et al. 2000; 2007a). This is because the CTs that prevail in the study area during the coldest months (i.e., January) are anticyclonic (Bonsoms et al. 2021a). In these synoptic situations, basin scale-inversions usually occur (Pepin and Kidd 2006) and temperature does not follow a linear lapse rate with elevation. The climatological period analyzed in this article provides a robust mean of every grid site, averaging local effects. However, this supposes a loss of details at hourly scales when many weather situations between the valleys and the culmination zones are observed (e.g., nocturnal inversions in the boundary layer during stable weather conditions). In this sense, further works should address whether the non-linear structure of the ML algorithms can provide optimal results during temporal scales shorter than a day.

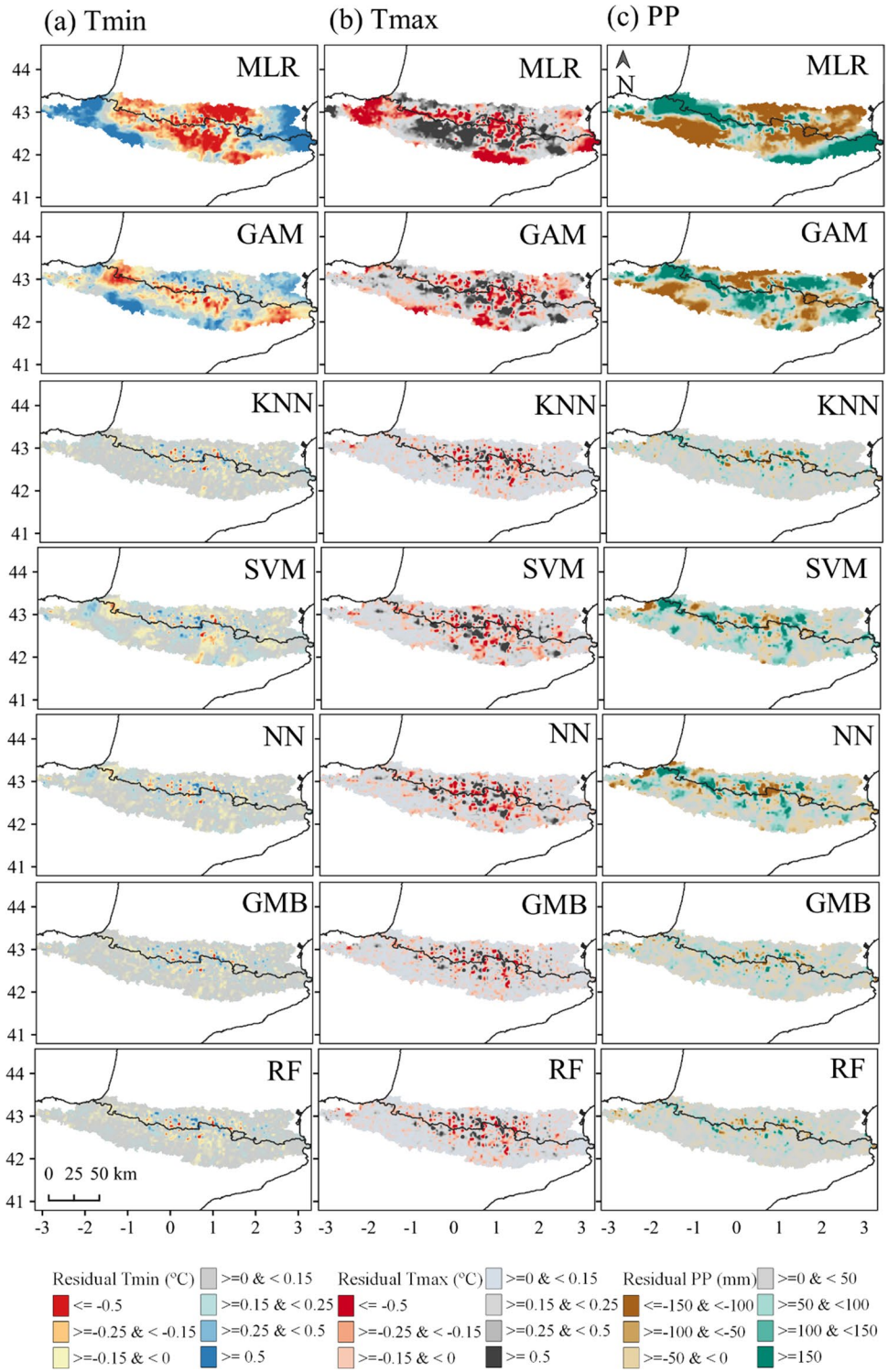

Fig. 9 **a** Tmin, **b** Tmax and **c** PP residual maps. The spatial interpolation of the bias was performed with IDW (power = 2). The coordinate reference system is ETRS89 (UTM 30N)

5 Conclusions

In a changing climate, evidence is accumulating that it is essential to assess the accuracy of the spatial reconstruction, interpolation and projection of the climate variables. In this article we provide insights about the spatial interpolation of climatological variables in a mountain range, comparing the traditional linear approaches (MLR and GAM) with ML methods. Our results show that the ML methods outperform the linear approaches. At the same time, no significant differences between the ML methods are observed for the interpolation of Tmin and Tmax. Our results highlight that ML regression trees better estimate the

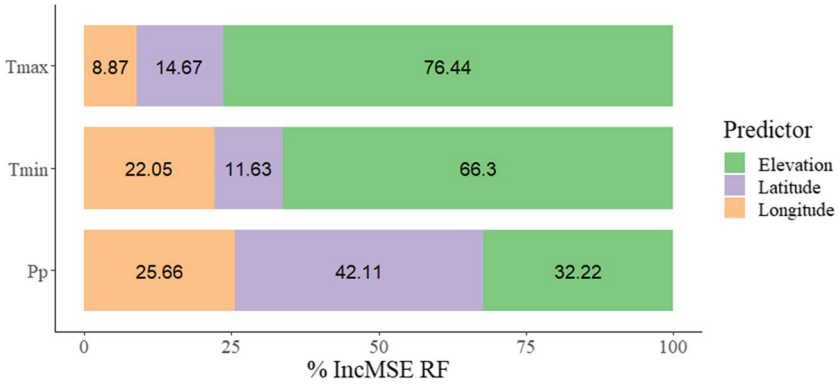

Fig. 10 Percentage of Mean Decrease Accuracy of each predictor variable

non-linearities between the geographical factors (latitude, longitude or elevation) and the climate variables. If the evaluation metrics are applied for the entire Pyrenees, four ML methods, called KNN, NN, GMB and RF have shown the same accuracy for the spatial interpolation of Tmin (RMSE = 0.28 ºC). For Tmax, the most accurate method is GMB (MAE = 0.43 ºC), and for PP the most accurate method is RF (MAE = 70.44 mm). If the analysis is carried out for each Pyrenees cluster and by elevation range, then RF is the most accurate method for the spatial interpolation of all the variables. The largest difference between the linear approaches and ML techniques is found with PP. The MLR approach underestimated all the variables and only explained 52% of the PP variance. ML methods based on regression trees were able to reproduce almost the totality of the PP variance (i.e., 93% for RF).

The evidence given in this article is not restricted to the spatial interpolation of climate data. The results suggest that ML methods would also be more accurate for the climatological spatial reconstruction, prediction, and for better understanding the variable of importance of each covariable included in the spatial models. The different regression tree ML methods show similar levels of accuracy. Therefore, further works should test the spatial climate data interpolation by including other independent variables and improving the time and spatial resolution.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00704-023-04725-5.

Acknowledgements This work is within the research topics examined by the research group "Antarctic, Artic, Alpine Environments-ANTALP" (2017-SGR-1102) founded by the Government of Catalonia, and "Grumets Research Group" through the NEWFORLAND research project (RTI2018-099397-B-C22MCIU / AEI / ERDF, EU) funded by the Spanish MCIU Ministry. JB is supported by a pre–doctoral University professor FPI grant (PRE2021-097046) funded by the Spanish Ministry of Science, Innovation and Universities. The authors acknowledge the CLIMPY project (Characterisation of evolution of climate and provision of information for adaptation in the Pyrenees) for the data provided.

Authors contribution's JB and MN designed the research. JB analysed the data and wrote the original draft with contributions of MN. JB and MN analysed the results and interpreted the data. All authors read and approved the final manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. JB is supported by a pre–doctoral University professor FPI grant (PRE2021-097046) funded by the Spanish Ministry of Science, Innovation and Universities and MN is funded by "Grumets Research Group" through the NEWFORLAND research project (RTI2018-099397-B-C22MCIU / AEI / ERDF, EU) funded by the Spanish MCIU Ministry.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare that they have not conflict of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Agnew MD, Palutikof JP (2000) GIS-based construction of baseline climatologies for the Mediterranean using terrain variables. Climate Res 14:115–127

Alonso-González E, López-Moreno JI, Navarro-Serrano FM, Revuelto J (2020a) Impact of North Atlantic Oscillation on the Snowpack in Iberian Peninsula Mountains. Water 12:105. https://doi.org/10.3390/w12010105

Alonso-González E, López-Moreno JI, Navarro-Serrano F, Sanmiguel-Vallelado A, Aznárez-Balta M, Revuelto J, Ceballos A (2020b)

Snowpack Sensitivity to Temperature Precipitation and Solar Radiation Variability over an Elevational Gradient in the Iberian Mountains. Atmos Res 243:104973. https://doi.org/10.1016/j.atmosres.2020.104973

Appelhans T, Mwangomo E, Hardy DR, Hemp A, Nauss T (2015) Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro Tanzania, Spatial Statistics 14:91–113. https://doi.org/10.1016/j.spasta.2015.05.008

Batalla M, Ninyerola M, Trapero L, Esteban P (2016) ACDA: Andorran Climate Digital Atlas (period 1981–2010) Map server. Institut d'Estudis Andorrans (IEA) Universitat Autonoma de Barcelona (UAB). https://www.iea.ad/90-sigma/cartografia/cartografia-climatica. Accessed 22 Jan 2022

Begueria S, Tomas-Burguera M, Serrano-Notivoli R, González-Hidalgo JC (2019) Gap filling of monthly temperature data and its effect on climatic variability and trends. J Clim 32(7797–782):1. https://doi.org/10.1175/JCLI-D-19-0244.s1

Beniston M, Uhlmann B, Goyette S, López-Moreno JI (2010) Will snow abundant winters still exist in the Swiss Alps in an enhanced greenhouse climate? Int J Climatol 31:1257–1263

Bonsoms J, Gonzalez S, Prohom M, Esteban P, Salvador-Franch F, López-Moreno JI, Oliva M (2021a) Spatio-temporal patterns of snow in the Catalan Pyrenees (SE Pyrenees NE Iberia). Int J Climatol 41(12):5676–5697. https://doi.org/10.1002/joc.7147

Bonsoms J, Salvador-Franch F, Oliva M (2021) Snowfall and snow cover evolution in the Eastern Pre-Pyrenees (NE Iberian Peninsula). Geographical Research Letters - Cuadernos de Investigación Geográfica 47(2):291–307. https://doi.org/10.18172/cig.4879

Breiman L (2001) Random forests. IEEE Machine Learning 45:5–32

Brunetti M, Maugeri M, Monti F, Nanni T (2006) Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. Int J Climatol 26:345–381

Buisán ST, López-Moreno JI, Sanz MA, Korchendorfer J (2016) Impact of weather type variability on winter precipitation temperature and annual snowpack in the Spanish Pyrenees. Climate Res 69:79–92. https://doi.org/10.3354/cr01391

Burrough PA, McDonnell RA (1998) Principles of Geographical Information Systems. Oxford University Press, New York, p 333

Burrough PA, McDonnell RA (1998b) Principles of Geographical Information Systems. Oxford University Press, Oxford

Carvalho MJ, Melo-Gonçalves P, Teixeira JC, Rocha A (2016) Regionalization of europe based on a k-means cluster analysis of the climate change of temperatures and precipitation. Phys Chem Earth 94:22–28. https://doi.org/10.1016/j.pce.2016.05.001

Cattell RB (1966) The Scree Test for the Number of Factors. Multivar Behav Res 1:245–276. https://doi.org/10.1207/s15327906mbr0102_10

Cristobal J, Ninyerola M, Pons X (2008) Modeling air temperature through a combination of remote sensing and GIS data. J Geophys Res Atmos 113(D13):1–13. https://doi.org/10.1029/2007JD009318

Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. Int J Climatol 28:2031–2064. https://doi.org/10.1002/joc.1688

Duhan D, Pandey A (2015) Statistical downscaling of temperature using three techniques in the Tons River basin in Central India. Theor Appl Climatol 121:605–622

Esteban P, Ninyerola M, Prohom M (2009) Spatial modelling of air temperature and precipitation for Andorra (Pyrenees) from daily circulation patterns. Theoret Appl Climatol 96:43–56. https://doi.org/10.1007/s00704-008-0035-3

Feng Y, Cui N, Hao W, Gao L, Gong D (2019) Estimation of soil temperature from meteorological data using different machine learning models. Geoderma 338:67–77

Forgy EW (1965) Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics 21:768–769

Frei C (2013) Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. Int J Climatol 34:1585–1605. https://doi.org/10.1002/joc.3786

Friedl MA, Brodley CE (1997) Decision tree classification of land cover from remotely sensed data. Remote Sens Environ 61:399–409

Friedman J (2001) Greedy boosting approximation: a gradient boosting machine. Annals of Statistics 29: 1189–1232. 101214/aos/1013203451

Hartigan JA, Wong MA (1979) Algorithm AS 136: a k-means clustering algorithm. J Royal Stat Soc Ser C (Appl Stat) 28(1):100–108. https://doi.org/10.2307/2346830

Hastie T, Tibshirani R (1987) Generalised additive model: some applications. J Am Statisticians Assoc 82:371–386

Haykin S (1998) Neural Networks: a comprehensive foundation, 2nd ed. Prentice Hall PTR, USA

Hengl T, Nussbaum M, Wright MN, Heuvelink GBM, Gräler B (2018) Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ 6:1–47. https://doi.org/10.7717/peerj.5518

Herrera S, Kotlarski S, Soares PMM, Cardoso RM, Jaczewski A, Gutierrez JM, Maraun D (2019) Uncertainty in gridded precipitation products: Influence of station density interpolation method and grid resolution. Int J Climatol 39:3717–3729. https://doi.org/10.1002/joc.5878

Huss M, Fischer M (2016) Sensitivity of very small glaciers in the Swiss Alps to future climate change. Front Earth Sci 4:34. https://doi.org/10.3389/feart.2016.00034

Kuhn M (2020) Caret: classification and regression training. R package version 6: 0–86. https://CRAN.R-project.org/package=caret

Lemus-Canovas M, Ninyerola M, Lopez-Bustins JA, Manguan S, García-Sellés C (2018) A mixed application of an objective synoptic classification and spatial regression models for deriving winter precipitation regimes in the Eastern Pyrenees. Int J Climatol. https://doi.org/10.1002/joc.5948

Lemus-Canovas M, Lopez-Bustins JA, Trapero L, Martin-Vide J (2019) Combining circulation weather types and daily precipitation modelling to derive climatic precipitation regions in the Pyrenees. Atmos Res 220:181–193. https://doi.org/10.1016/j.atmosres.2019.01.018

Li J, Heap A, Potter A, Daniell JJ (2011) Application of machine learning methods to spatial interpolation of environmental variables. Environ Modell Softw 26:1647–1659

López-Bustins JA, Lemus-Canovas M (2020) The influence of the Western Mediterranean Oscillation upon the spatio-temporal variability of precipitation over Catalonia (northeastern of the Iberian Peninsula). Atmos Res 236:104819. https://doi.org/10.1016/j.atmosres.2019.104819

López-Moreno JI, Nogués-Bravo D (2005) A generalized additive model for modelling the spatial distribution of snowpack in the Spanish Pyrenees. Hydrol Process 19:3167–3176

López-Moreno JI, Nogués-Bravo D (2006) Interpolating snow depth data: a comparison of methods. Hydrol Process 20:2217–2232

López-Moreno JI, Vicente-Serrano SM, Lanjeri S (2007) Mapping snowpack distribution over large areas using GIS and interpolation techniques. Climate Res 33:257–270

López-Moreno JI, Vicente-Serrano SM, Morán-Tejeda E, Lorenzo J, Kenawy A, Beniston M (2011) NAO effects on combined temperature and precipitation winter modes in the Mediterranean mountains: Observed relationships and projections for the 21st century. Global Planet Change 77:62–76

Martín-Vide J, López-Bustins JA (2006) The Western Mediterranean Oscillation and rainfall in the Iberian Peninsula. Int J Climatol 26:1455–1475. https://doi.org/10.1002/joc.1388

Meyer H, Kühnlein M, Appelhans T, Nauss T (2016a) Comparison of four machine learning algorithms for their applicability in satellite-based optical rainfall retrievals. Atmos Res 169:424–433

Meyer H, Katurji M, Appelhans T, Müller MU, Nauss T, Roudier P, Zawar-Reza P (2016b) Mapping Daily Air Temperature for Antarctica Based on MODIS LST. Remote Sensing 8:732. https://doi.org/10.3390/rs8090732

Mira M, Ninyerola M, Batalla M, Pesquer L, Pons X (2017) Improving Mean Minimum and Maximum Month-to-Month Air Temperature Surfaces Using Satellite-Derived Land Surface Temperature. Remote Sens 9(12):1313. https://doi.org/10.3390/rs9121313

Ninyerola M, Pons X, Roure JM (2000) A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. Int J Climatol 20:1823–1841. https://doi.org/10.1002/1097-0088(20001130)20:14%3c1823::AID-JOC566%3e3.0.CO;2-B

Ninyerola M, Pons X, Roure JM (2005) Atlas Climático Digital de la Península Ibérica. Metodología y aplicaciones en Bioclimatología y Geobotánica. Universitat Autònoma de Barcelona, Bellaterra, Spain

Ninyerola M, Pons X, Roure JM (2007a) Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. Int J Climatol 27:1231–1242. https://doi.org/10.1002/joc.1462

Ninyerola M, Pons X, Roure JM (2007b) Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. Theoret Appl Climatol 89:195–209. https://doi.org/10.1007/s00704-006-0264-2

Nogués-Bravo D, Bastos-Araujo M, Lasanta T, López-Moreno JI (2008) Climate Change in Mediterranean Mountains during the 21st Century. Ambio 37:280–285

Noi P, Degener J, Kappas M (2017) Comparison of multiple linear regression cubist regression and random forest algorithms to estimate daily air surface temperature from dynamic combinations of MODIS LST data. Remote Sens 9:398. https://doi.org/10.3390/rs9050398

Oliva M, Serrano E, Gómez-Ortiz A, González-Amuchastegui MJ, Nieuwendam A, Palacios D, Pellitero-Ondicol R, Pérez-Alberti A, Ruiz-Fernández J, Valcárcel M, Vieira G, Antoniades D (2016) Spatial and temporal variability of periglaciation of the Iberian Peninsula. Quatern Sci Rev 137:176–199. https://doi.org/10.1016/j.quascirev.2016.02.017

Oliva M, Ruiz-Fernández J, Barriendos M, Benito G, Cuadrat JM, García-Ruiz JM, Giralt S, Gómez-Ortiz A, Hernández A, López-Costas O, López-Moreno JI, López-Sáez JA, Martínez- Cortizas A, Moreno A, Prohom M, Saz MA, Serrano E, Tejedor E, Trigo R, Valero-Garcés BL, Vicente-Serrano S (2018) The Little Ice Age in Iberian mountains. Earth-Science Review 177:175–188. https://doi.org/10.1016/j.earscirev.2017.11.010

Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. Comput Geosci 30:683–691. https://doi.org/10.1016/JCAGEO200403012

Pelletier C, Valero S, Inglada J, Champion N, Dedieu G (2016) Assessing the Robustness of Random Forests to Map Land Cover with High Resolution Satellite Image Time Series over Large Areas. Remote Sens Environ 187:156–168

Pellicciotti F, Helbing J, Rivera A, Favier V, Corripio J, Araos J, Sicart J, Carenzo M (2008) A study of the energy balance and melt regime on Juncal Norte Glacier semi-arid Andes of central Chile using melt models of different complexity. Hydrol Process 22:3980–3997. https://doi.org/10.1002/hyp.7085

Peña-Angulo D, Brunetti M, Cortesi N, González-Hidalgo JC (2016) A new climatology of maximum and minimum temperature (1951–2010) in the Spanish mainland: a comparison between three different interpolation methods. Int J Geogr Inf Sci 30:2109–2132. https://doi.org/10.1080/13658816.2016.1155712

Pepin N, Kidd D (2006) Spatial temperature variation in the Eastern Pyrenees. Weather 61:300–310

Pozdnoukhov A, Foresti L, Kanevski M (2009) Data-driven topoclimatic mapping with machine learning methods. Nat Hazards 50:497–518. https://doi.org/10.1038/nclimate2563

R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.r-project.org/. Accessed 22 Jan 2022

Sekulić A, Kilibarda M, Heuvelink GBM, Nikolić M, Bajat B (2020) Random forest spatial interpolation. Remote Sens 12:1687. https://doi.org/10.3390/rs12101687

Serrano-Notivoli R, Beguería S, Saz MÁ, Longares LA, De Luis M (2017) SPREAD: A high-resolution daily gridded precipitation dataset for Spain-An extreme events frequency and intensity overview. Earth Syst Sci Data 9:721–738. https://doi.org/10.5194/essd-9-721-2017

Serrano-Notivoli R, Tejedor E (2021) From rain to data: A review of the creation of monthly and daily station-based gridded precipitation datasets. Wiley Interdiscipl Rev: Water 8(6). https://doi.org/10.1002/wat2.1555

Tripathi S, Srinivas VV, Nanjundiah RS (2006) Downscaling of precipitation for climate change scenarios: a support vector machine approach. Journal of Hydrology 330: climate change scenarios: a support vector machine approach. J Hydrol 330:621–640

Turco M, Rosa-Cánovas JJ, Bedia J, Jerez S, Montávez JP, Llasat MC, Provenzale A (2018) Exacerbated fires in Mediterranean Europe due to anthropogenic warming projected with non-stationary climate-fire models. Nat Commun 9:3821

Vicente-Serrano SM, Saz-Sánchez MA, Cuadrat JM (2003) Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): Application to annual precipitation and temperature. Climate Res 24:161–180. https://doi.org/10.3354/cr024161

Vicente-Serrano SM, Lanjeri S, Lopez-Moreno JI (2007) Comparison of different procedures to map reference evapotranspiration using geographical information systems and regression-based techniques. Int J Climatol 27:1103–1118

Vicente-Serrano S, Begueria S, Lopez-Moreno J, García-Vera M, Stepanek P (2010) A complete daily precipitation database for Northeast Spain: Reconstruction quality control and homogeneity. Int J Climatol 30:1146–1163. https://doi.org/10.1002/joc.1850

Vicente-Serrano S, López-Moreno MJI, Vega-Rodriguez MI, Begueria S, Cuadrat JM (2010) Comparison of regression techniques for mapping fog frequency: application to the Aragon region (northeast Spain). Int J Climatol 30:935–945

Viviroli D, Durr H, Messerli B, Meybeck M, Weingartner R (2007) Mountains of the world water towers for humanity: Typology mapping and global significance. Water Resour Res 43:W07447. https://doi.org/10.1029/2006WR005653

Willmott CT (1982) Some comments on the evaluation of model performance. Bull Am Meteor Soc 63:1309–1313

Xercavins-Comas A (1985) Els climes del Pirineu Oriental: des de les terres gironines fins a la Catalunya Nord i Andorra. Documents D'anàlisi Geogràfica 7:81–102

Zhou J, Li XB, Mitri HS (2016) Classification of rockburst in underground projects: Comparison of ten supervised learning methods. J Comput Civil Eng 30:04016003