RMetS
Royal Meteorological Society

1961–1990 high-resolution monthly precipitation climatologies for Italy

A. Crespi,[a] M. Brunetti,[b] G. Lentini[a,c] and M. Maugeri[a,b]*

[a] *Department of Environmental Science and Policy, Università degli Studi di Milano, Italy*
[b] *Istituto di Scienze dell'Atmosfera e del Clima, CNR, Bologna, Italy*
[c] *Poliedra – Politecnico di Milano, Italy*

ABSTRACT: High-resolution monthly precipitation climatologies for Italy are presented. They are based on 1961–1990 precipitation normals obtained from a quality-controlled dataset of 6134 stations covering the Italian territory and part of the Northern neighbouring regions. The climatologies are computed by means of two interpolation methods modelling the precipitation-elevation relationship at a local level, more precisely a local weighted linear regression (LWLR) and a local regression kriging (RK) are performed. For both methods, local optimisations are also applied in order to improve model performance. Model results are compared with those provided by two other widely used interpolation methods which do not consider elevation in modelling precipitation distribution: ordinary kriging and inverse distance weighting. Even though all the four models produce quite reasonable results, LWLR and RK show the best agreement with the observed station normals and leave-one-out-estimated mean absolute errors ranging from 5.1 mm (July) to 11 mm (November) for both models. Their better performances are even clearer when specific clusters of stations (e.g. high-elevation sites) are considered. Even though LWLR and RK provide very similar results both at station and at grid point level, they show some peculiar features. In particular, LWLR is found to have a better extrapolation ability at high-elevation sites when data density is high enough, while RK is more robust in performing extrapolation over areas with complex orography and scarce data coverage, where LWLR may provide unrealistic precipitation values. However, by means of prediction intervals, LWLR provides a more straightforward approach to quantify the model uncertainty at any point of the study domain, which helps to identify the areas mainly affected by model instability. LWLR and RK high-resolution climatologies exhibit a very heterogeneous and seasonal-dependent precipitation distribution throughout the domain and allow to identify the main climatic zones of Italy.

KEY WORDS   high-resolution climatology; precipitation; Italy; interpolation methods

Received 28 July 2016; Revised 9 June 2017; Accepted 27 June 2017

1. Introduction

High-resolution precipitation climatologies are becoming increasingly important as the small-scale spatial distribution of normal precipitation is often needed now for both models and other decision-support tools applied to a wide range of fields, such as agriculture, engineering, hydrology, energy management and natural resource conservation (Daly *et al.,* 2002; Daly, 2006). High-density observational datasets must be integrated by interpolation methods to consider all major factors affecting spatial precipitation patterns and to provide reasonable estimates even for areas with complex topography or for remote regions, such as mountain areas, with poor station coverage (Daly *et al.,* 2008).

While 30-arc-second-resolution temperature climatologies for Italy have been recently provided by Brunetti *et al.* (2014), at present monthly precipitation climatologies are not available with such resolution for Italy as a whole.

This is likely due to the difficulty in gathering and checking very large amounts of observational data and in coping with the great spatial variability of precipitation due to the complex orography of Italy. Very few climatological maps are available at a national scale and they are mostly in a non-digitized form, such as the hand-drawn precipitation maps for 1921–1950 produced by the Italian Hydrographical Service (Servizio Idrografico (SI), 1957). More recently, ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale) provided monthly 5 km × 5 km precipitation climatologies for 1951–1980 (ISPRA, 2014) and CREA-CMA (Research unit for Climatology and Meteorology applied to Agriculture) published national monthly climatologies for 1961–1990, 1971–2000 and 1981–2010 (Esposito *et al.,* 2015). However, in both cases, the rather low density of stations was not suitable to completely account for the complexity of the Italian physiography. On the contrary, remarkable results have been achieved for smaller Italian areas, and several works focusing on both spatial distribution and temporal behaviour of precipitation have been produced. In particular, because of its key role as a water reservoir for a wide trans-national area, the Alpine region has been

* Correspondence to: M. Maugeri, Department of Physics, Università degli Studi di Milano, via Celoria 16, 20133 Milan, Italy. E-mail: maurizio.maugeri@unimi.it

extensively investigated in numerous climatological studies often encompassing parts of Northern and Central Italy (see e.g. Frei and Schär, 1998; Schwarb, 2000; Brunetti *et al.*, 2012; Isotta *et al.*, 2014). Moreover, the cartographic maps produced by Cati (1981) are still a reference study for the Po basin climatology, while Brunetti *et al.* (2009) provided 1961–1990 high-resolution monthly precipitation climatologies for Northern and Central Italy. In addition, several works concern regional or sub-regional domains, such as Biancotti *et al.* (1998) for Piedmont, Antolini *et al.* (2016) for Emilia-Romagna, Drago (2005) and Di Piazza *et al.* (2011) for Sicily and Secci *et al.* (2010) for Sardinia.

A large variety of interpolation techniques have been developed to work with spatially distributed data. Among the most used techniques in climatological studies we can mention inverse distance weighting (IDW), splining, polynomial regression, kriging (and its variants) and the parameter elevation regression on independent slopes model (PRISM). The performances of these techniques in gridding precipitation data are compared and discussed in several literature works showing that the application of interpolation methods taking into account the relationship between precipitation and topography, especially elevation, improves significantly the spatial prediction of rainfall (see e.g. Martinez-Cob, 1996; Goovaerts, 2000; Vicente-Serrano *et al.*, 2003; Diodato and Ceccarelli, 2005; Masson and Frei, 2014). In particular, PRISM is found to be one of the most suitable approaches to produce high-resolution climatologies in areas with complex orography (Schwarb, 2000; Daly, 2006). More details on PRISM can be found in Daly *et al.* (1994, 2002, 2008), Daly (2006) and on the official website of the PRISM group (http://www.prism.oregonstate.edu).

Within this context, the present work presents 30-arc-second-resolution precipitation climatologies for Italy based on 1961–1990 monthly normals from a quality-controlled dataset of more than 6000 stations covering the whole of Italy and part of the neighbouring countries. The climatologies are computed on the grid points of the USGS GTOPO30 digital elevation model (DEM) by means of two approaches, both capturing the precipitation-elevation relationship at a local scale: a local weighted linear regression (LWLR) of precipitation *versus* elevation and a local regression kriging (RK). The main assumptions of these methods are that spatial pattern of precipitation is closely linked to physiographical features of the Earth's surface and that this link is best captured considering small areas (Basist *et al.*, 1994; Daly *et al.*, 2002, 2008; Daly, 2006). These approaches lead us to an evaluation of climatological normals for a number of points several orders of magnitude larger than presently available series.

In this work, the features of the techniques used are extensively investigated and, in order to assess the importance of elevation in gridding precipitation normals, their performance is compared with those provided by two interpolation approaches – ordinary kriging (OK) and IDW – which do not model the precipitation-elevation relationship.

2. Data

The dataset used to produce the 1961–1990 precipitation climatologies is the result of more than 10 years of activities carried out at the Institute of Atmospheric Sciences and Climate of the Italian National Research Council (ISAC-CNR) and at the Department of Physics at Milan University to obtain the largest possible amount of precipitation records and metadata for Italy and the surrounding areas. It represents an extended and enhanced version of datasets presented in previous works (Spinoni, 2010; Brunetti *et al.*, 2012) and it is in continuous progress. Information on data sources considered for the present work and the number of available series is listed in Table 1.

Most series derive from the rain gauge network of the former SI. After its closure at the end of the 20th century, its personnel and duties were transferred to the individual Italian Regions which continued to manage the station network directly or assigned its management to external agencies; in a few cases, the network was abandoned because a regional one had already been set up and no resources were available to maintain both. The attribution of the SI competences to the regions generally brought new resources for the station network and data rescue activities, though at the cost of a greater difficulty in collecting data covering the entire national territory as indicated by the high number of data providers listed in Table 1. Moreover, the decline of the SI together with the transition from mechanical to automatic station networks lead to an inhomogeneous data availability for Italy after the 1980s. Therefore, our choice of 1961–1990 period as the reference for the climatologies has been suggested by the peculiar situation of the Italian precipitation network.

Actually, some activities, such as the organization of a national data archive, continued to be performed also at national level, firstly by APAT (Agenzia per la Protezione dell'Ambiente e i Servizi Tecnici) and then by ISPRA. In particular, the data digitisation of the archive of the SI has been performed both at the national scale and by some of the regional services in charge of the station network management. Moreover, the availability of digital data from this archive has increased over the last few years, also thanks to contributions from many digitisation projects carried out by several research institutions; therefore, we had to deal with a wide number of sources based mainly on the same data networks. However, duplicate series can provide different and often complementary information because data and metadata availability can vary significantly depending on the considered sources. In addition, digitisation can derive from different non-digital supports, such as yearbooks published by the former SI (http://www.acq.isprambiente.it/annalipdf/), data forms filled out by station observers and strip charts of recording rain gauges.

Therefore, collected data were first checked to identify and merge series retrieved in more than one source. In the case of overlapping time intervals, we gave priority to the most reliable series based on number of gaps, temporal coverage of data source, availability of daily data and, in some cases, expert judgement of authors.

Table 1. Sources of precipitation series.

Data source	Number of series and other information
ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale)	4951 monthly series covering the entire Italian territory. A total of 1527 of them are available also at daily resolution. This dataset is composed both of series that have been digitized at ISPRA and by series that have been digitized by Italian local services and sent to ISPRA. We received the dataset directly from ISPRA, together with corresponding metadata. The ISPRA dataset is also available on-line (see http://www.scia.isprambiente.it/home_new_eng.asp).
CNR-ISAC (Italian National Research Council-Institute of Atmospheric Sciences and Climate) – Milan University, Department of Physics	187 monthly series covering the entire Italian territory. A total of 108 of them are available also at daily resolution. Most of these series have been used to study the long-term evolution of precipitation over Italy (Brunetti *et al.,* 2004, 2006a). This dataset is managed by the authors of this paper.
HISTALP – Historical Instrumental Climatological Surface Time Series of the Greater Alpine Region	137 monthly series covering the Greater Alpine Region. HISTALP (http://www.zamg.ac.at/histalp/) data have been used to study the long-term evolution of precipitation over this area (Brunetti *et al.,* 2006b).
RICLIC Project	84 daily series covering the Adda river catchment. For more information see the RICLIC Project web site (http://www.riclic.unimib.it/). We received the data from Milano-Bicocca University (2007).
Nuovo studio dell'idrologia superficiale della Sardegna	415 monthly Sardinian series set up in a monographic study performed by the Ente Autonomo del Flumendosa for the Sardinia regional administration (http://pcserver.unica.it/web/sechi/main/Corsi/Didattica/IDROLOGIA/DatiSISS/index.htm).
Italian Air Force	104 daily series covering the entire Italian territory. They concern synoptic stations (refer to http://clima.meteoam.it/istruzioni.php for data access). We received the data in the frame of an agreement between Italian Air Force and the Italian National Research Council.
CREA – CMA (The Agricultural Research Council – Research unit for Climatology and Meteorology applied to Agriculture)	249 daily series covering the entire Italian territory. The CREA-CMA dataset is composed both by the digitisation of the archive of the former Italian Central Office for Meteorology and by more recent data acquired by CREA-CMA in real-time. The data are partially available at http://cma.entecra.it/homePage.htm. The other data have to be requested at CREA-CMA. We received the data in the frame of the CLIMAGRI Project (2001).
Italian local meteorological/hydrological/environmental services	73 daily series from the Ufficio Idrografico of the Provincia Autonoma of Bolzano (http://www.provincia.bz.it/meteo/dati-storici.asp).
	146 daily series of the Provincia di Trento from Meteotrentino (http://www.meteotrentino.it/dati-meteo/info-dati.aspx?id=3).
	122 daily series covering the Reno river catchment from the 'Autorità di Bacino del Reno' (http://ambiente.regione.emilia-romagna.it/suolo-bacino/sezioni/strumenti-e-dati/pioggia/dati-pioggia).
	845 Tuscany daily series from the former 'Servizio Idrografico e Mareografico – Ufficio di Pisa' (refer to http://www.sir.toscana.it/annali-idrologici for data access).
	106 daily stations from ARPA Piemonte (https://www.arpa.piemonte.gov.it/rischinaturali/accesso-ai-dati/annali_meteoidrologici/annali-meteo-idro/banca-dati-meteorologica.html)
	109 daily series received from 'Regione Autonoma Valle d'Aosta Centro Funzionale Regionale' (2014) in the frame of the Project of National Interest NextData.
	146 daily series from ENEL, the former Italian National electricity board. They mainly concern dams in the Alps and in the Apennines. We received the data in the frame of the Project CARIPANDA (2008), funded by Fondazione CARIPLO.
	539 Sicily monthly stations. A total of 307 of them are available also at daily resolution (refer to http://pti.regione.sicilia.it/portal/page/portal/PIR_PORTALE/PIR_LaStrutturaRegionale/PIR_AssEnergia/PIR_Dipartimentodellacquaedeirifiuti/PIR_Organigramma/PIR_SERVIZIO2OSSERVATORIODELLEACQUE for data access).
	443 Calabria and Basilicata daily stations from the former 'Servizio Idrografico e Mareografico – Ufficio di Catanzaro' (refer to http://www.cfd.calabria.it/index.php/dati-stazioni/dati-storici for data access).
	171 daily stations from ARPA Veneto. We received the data (2012) in the frame of the EU FP7 ECLISE Project (refer to http://www.arpa.veneto.it/ for data access).
	126 monthly stations from Puglia regional administration (http://www.regione.puglia.it/index.php?page=documenti&opz=getdoc&id=165).
	63 Campania monthly stations from the former 'Servizio Idrografico e Mareografico – Ufficio di Napoli' (refer to http://www.sito.regione.campania.it/agricoltura/meteo/agrometeo.htm for data access).
Slovenian Environment Agency	40 monthly Slovenian series (http://www.arso.gov.si/en/).
MeteoSwiss	173 Swiss daily series (https://gate.meteoswiss.ch/idaweb/).
European Climate Assessment & Dataset project	167 daily series (eca.knmi.nl/). They mainly refer to Emilia-Romagna.

The identification of duplicates was not straightforward, as station location is generally reported in metadata with just a 1′ resolution (i.e. 1–2 km) and station name may vary from one source to another. Moreover, duplicate series could have temporal lags (measures concerning a 24-h period, for example 0800 GMT – 0800 GMT, may have been assigned to the first or to the second day), or data with different rounding and/or incoherent entries for some sub-intervals of the common period. It was therefore necessary to check a number of cases individually to decide whether two series could be merged or not. After this procedure, the complete dataset included about 6000 monthly series, all associated to their corresponding metadata. Series with less than 10 years of available data were then discarded and the remaining 5119 were subjected to further quality controls. To this purpose, we estimated the monthly precipitation series at each station site by means of the series of the neighbouring stations and compared the estimated values with the measured ones (Brunetti *et al.*, 2009, 2012). The first step of the estimation procedure consists in identifying, for each monthly datum of each station (test station), the ten closest stations (reference stations) with a non-missing value in correspondence with the entry under consideration and with a sufficient number of data for that month in common with the test series. The threshold for the number of common data is set to 15 if the test series has more than 15 years of records, whereas it is reduced (down to 9 years) in case of lower data availability. In order to consider the same subset for all the reconstructions, the reference stations are selected in any case among the series with more than 15 years of data. After identifying the reference series, the test series monthly datum under consideration (p_{test_m}) is estimated from the corresponding datum of each of the ten reference series ($p_{\text{rif}_{m,i}}$) with the following relation:

$$\tilde{p}_{\text{test}_m,i} = p_{\text{rif}_m,i} \cdot \frac{\overline{p}_{\text{test}_m}}{\overline{p}_{\text{rif}_m,i}} \qquad (i = 1, \ldots, 10) \qquad (1)$$

where $\overline{p}_{\text{test}_m}$ and $\overline{p}_{\text{rif}_m,i}$ are the test's and reference series' averages in the considered month over their period of common data availability. The best estimation of p_{test_m} is finally obtained considering the median of the ten estimated values $\tilde{p}_{\text{test}_m,i}$. This estimation procedure is based on the well-known anomaly method, and the only peculiar feature is that the period to calculate the normals is selected for each month and for each pair of test and reference series on the basis of their common data availability.

The comparisons between observed and estimated series highlighted the sites showing the lowest agreement with the neighbouring stations in terms of mean absolute error (MAE) and root mean square error (RMSE), both in absolute and relative terms. High errors could be due to spatial distance or strong elevation differences with respect to the surrounding stations, but also to actual erroneous data, such as outliers caused by digitisation oversights, inhomogeneities and unreliable sequences of null values. The stations exhibiting high errors also after correcting these

spurious values were ultimately excluded from the analysis unless the errors could be ascribed purely to the remote location of the stations. A further control concerned the detection of wrong coordinates. To this aim, the correlation coefficients for all station pairs were computed and compared to verify, for each site, if the highest correlation values were associated with the closest stations. Correcting location coordinates is indispensable, as location mistakes may induce significant errors in the evaluation of the orographic features of the site and, consequently, in the estimation of the precipitation-elevation dependence which is the primary assumption of our study.

After quality checks, the dataset was reduced to 4751 stations. For each site, we computed the 1961–1990 monthly precipitation normals and, whenever the 1961–1990 period was completely or partially unavailable, missing data were reconstructed by the same procedure described for the quality-check and based on the neighbouring stations, but considering only five reference stations and taking the weighted mean of $\tilde{p}_{\text{test}_m,i}$ after discarding the maximum and minimum ones. The goal here was to improve the accuracy of the reconstruction, without reducing too much robustness. Since a significant fraction of the series (49%) have more than 20% of missing years in the 1961–1990 period, we assessed the reliability of the gap-filling procedure and, in particular, its suitability to handle series with a relevant amount of missing values. To do this, we selected the series with at least 80% (24 years) of valid records in 1961–1990 for each month, and then we discarded, from one station at a time, a certain number of randomly selected entries to obtain a series of 10 years only of available records in 1961–1990. We next evaluated its monthly normal after filling in gaps. This procedure was repeated 100 times for each series and the mean and standard deviation of the reconstructed normals were computed to check the accuracy and the stability of the filling method. As expected, the mean values turned out to be in good agreement with the observed values (mean errors close to zero) and the standard deviations ranged from 2.8 mm in July to 3.9 mm in October. The monthly mean of the ratios between the standard deviations and the corresponding model RMSEs (see paragraph 4.1 for discussion of model errors) was under 30% for both LWLR and RK, indicating that the normals based on only 10 years of data in 1961–1990 have a rather low gap-filling error as compared to other errors that may affect our precipitation climatologies.

After the control procedures, the calculated normals were integrated with the 1961–1990 precipitation normals of several Austrian, Swiss and French sites available at ZAMG (Zentral Anstalt für Meteorologie und Geodynamik), MeteoSwiss and MétéoFrance, respectively, resulting in a total of 6134 available stations.

The domain chosen to compute the 1961–1990 precipitation climatologies corresponds to the area within Italy's administrative boundaries and includes the trans-national portions of the Po basin (mainly located in the Swiss territory and, to a lesser extent, in France), considering the remarkable role of the Po river (with 74 000 km² of
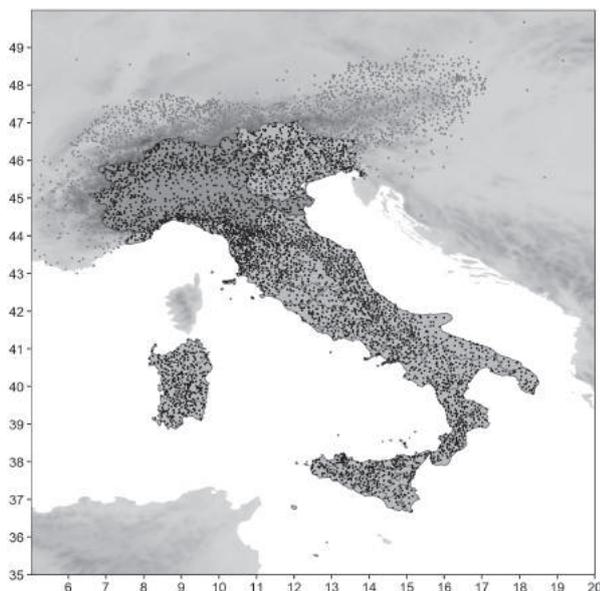
Figure 1. Study domain (light and dark grey bounded areas) and spatial distribution of the 6134 stations in the final version of the dataset. The dark grey portion corresponds to the Po basin and the black dots represent the 4525 stations included within the study domain. The sites outside this area (grey dots) were considered only in order to have a homogeneous station distribution also around the points at the boundaries of the domain.

drainage area) as the major water resource for agriculture and other economic activities in Northern Italy. The stations outside this area were considered to have a homogeneous station distribution also at the points close to the boundaries. The spatial distribution of the stations is shown in Figure 1 together with the study domain. A total of 4525 stations out of 6134 are located within the study area. The average spatial density is slightly less than one station per 70 km^2, with the greatest coverage between Liguria and Tuscany.

Figure 2 shows the vertical distribution of the stations compared to that of the DEM grid cells in the study domain. Except for a lower coverage concerning areas below 400 m, the station distribution is fairly homogeneous up to 2500 m. The mean station-to-grid-cell ratio is about 0.015 (see inset box in Figure 2), even though this ratio varies significantly over Italy, ranging from 0.003 to 0.026 for 1° sub-regions centred over the Italian grid points in Figure 3. At greater altitudes, data availability turns out to be scarce but, since the fraction of grid cells is also very low (<1% of the total), problems concerning rainfall estimation over these areas can be considered quite limited.

3. Methods

Precipitation climatologies presented in this paper are computed for grid cells of the 30-arc-second-resolution (∼800 m) GTOPO30 DEM (USGS, 1996). However, as direct effects of elevation on precipitation appear to be most important at larger scales (see e.g. Daly *et al.*, 2008), both for LWLR and RK a Gaussian filter was

applied to filter out terrain features, while retaining the 30-arc-second-resolution. The smoothing of the DEM was performed by assigning to each cell an elevation obtained as a weighted average of the elevations of the surrounding cells, with weights provided by a Gaussian function decreasing to 0.5 at a distance of *S* km from the cell itself, where *S* defines the degree of smoothing. Different values of *S* (*S* = 1, ..., 5) were tested assigning to each station all orographic parameters (including elevation) extracted from the smoothed DEMs and evaluating which degree of smoothing produces the lowest model error for LWLR and RK (see below for more details on model errors). Both for LWLR and RK, *S* = 3 turned out to give the best agreement between the model estimations and the station normals. Climatologies are therefore computed on this smoothed version of the GTOPO30 DEM (3smDEM, see Figure 3), which is also used to assign to the stations the orographic parameters required by LWLR and RK.

3.1. Local weighted linear regression

The LWLR 1961–1990 monthly precipitation climatologies are constructed by means of a PRISM-based procedure which estimates the local precipitation-elevation relationship at all grid cells of the domain, taking into account the topographic similarities between the stations and the grid cell itself. More precisely, precipitation normals at each cell (x, y) of the DEM are computed estimating a local weighted linear precipitation-elevation regression (LWLR) and assigning to the grid cell the value corresponding to its elevation by means of the following expression:

$$p(x, y) = a(x, y) + b(x, y) \cdot h(x, y) \qquad (2)$$

where $h(x, y)$ is the grid-cell elevation and $a(x, y)$ and $b(x, y)$ are the coefficients of the weighted linear regression of precipitation *versus* elevation. In this procedure, the definition of the regression weights is crucial. In our case, the weights of the stations selected for the evaluation of the regression coefficients at each grid cell are computed on the basis of the distances and the level of similarity (in terms of orographic features) between the stations and the grid cell itself. Thus, the weight of the *i*th station involved in the linear regression for the grid cell *(x,y)* is the product of several weighting factors:

$$w_i(x, y) = w_i^{\mathrm{rad}}(x, y) \cdot w_i^h(x, y) \cdot w_i^{\mathrm{st}}$$
$$(x, y) \cdot w_i^{\mathrm{facet}}(x, y) \cdot w_i^{d\mathrm{sea}}(x, y) \qquad (3)$$

All the weighting factors [radial distance (rad), vertical distance (*h*), slope steepness (st), slope orientation (facet) and distance from the sea (*d*sea)] range from 0 to 1 and are based on Gaussian functions of the form:

$$w_i^{\mathrm{par}}(x, y) = e^{-\left(\frac{\left(\Delta_i^{\mathrm{par}}(x,y)\right)^2}{c^{\mathrm{par}}}\right)} \qquad (4)$$

where par is the geographical parameter that is being considered, Δ_i^{par} is the difference between the values of the
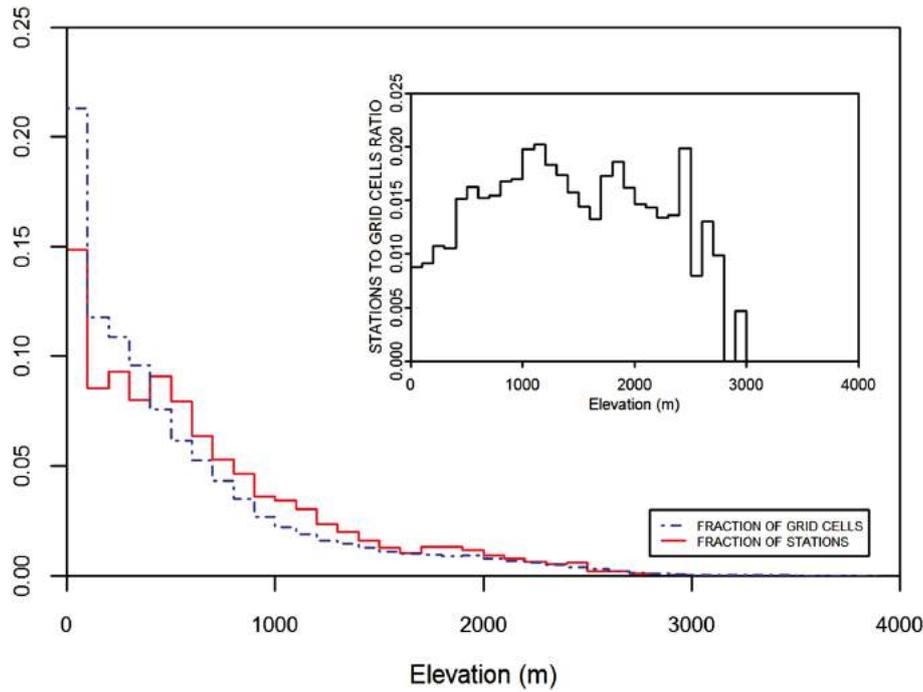
Figure 2. Vertical distribution of the 6134 stations (solid line) compared to the grid-cell elevation distribution (dashed line) in the domain for which we calculated the climatologies. The inset box shows the vertical distribution of the stations to grid-cells ratio. [Colour figure can be viewed at wileyonlinelibrary.com].
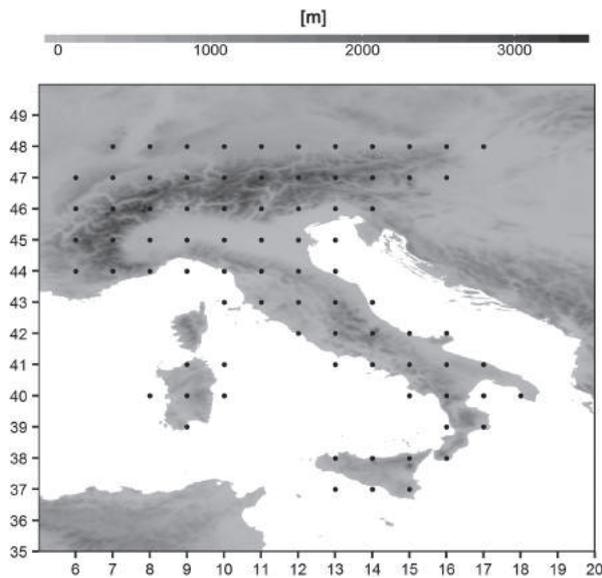

Figure 3. The 3smDEM and (superimposed) the 1° x 1° resolution grid used to locally optimize the decreasing coefficients of weighting factors.

considered parameter at the station i and at the grid cell (x, y), and c^{par} is the coefficient that regulates the weight decrease. For an easier interpretation of the weighting factors, the coefficient c^{par} can also be expressed in terms of the weighting halving distance value $\Delta^{par}_{\frac{1}{2}}$:

$$c^{par} = \frac{\left(\Delta^{par}_{\frac{1}{2}}\right)^2}{\ln 2} \qquad (5)$$

The LWLR method includes an algorithm which selects the 15 stations with the highest weights to be considered in the estimation of the regression coefficients. If fewer than five stations are found within a distance of 200 km from the grid point, the grid-cell precipitation is not evaluated.

Moreover, the coefficients regulating the decrease of the weighting factors in Equation (5) are locally optimized for each month by an iterative method. In fact, considering the orographic complexity of the domain, the influence of geographical features may not be the same for the whole area and may vary during the year. Therefore, at each point of a 1° × 1° resolution grid covering the whole study area (black dots in Figure 3), the normals of the stations in the range of 200 km were recursively computed to search for the optimal $\Delta^{par}_{\frac{1}{2}}$ values minimizing the error estimators. Due to computational time constraints, the optimized coefficients are first calculated on a 1° × 1° resolution grid. The optimized values are then interpolated on the high-resolution grid by IDW and used in the LWLR interpolation procedure described above to produce the climatologies.

Finally, with the LWLR scheme we can define a prediction interval for each grid-cell estimation. As explained in Daly *et al.* (2008), the procedure consists of estimating the variance of the precipitation values of a grid point at elevation h_{new} as:

$$s^2 \left\{ p_{h_{new}} \right\} = s^2 \left\{ \tilde{p}_{h_{new}} \right\} + \text{MSE} \qquad (6)$$

where $s^2 \left\{ \tilde{p}_{h_{new}} \right\}$ is the variance in the possible location of the expected precipitation for a given elevation and MSE

is the mean square error of the observed station precipitation values compared to those obtained by the regression model. $s^2 \left\{ \tilde{p}_{h_{\text{new}}} \right\}$ depends on the regression coefficients' errors, while MSE represents the fraction of the variability of the station precipitation normals which is not described by the precipitation *versus* elevation regression on which LWLR is based. Equation (6) can be written by expressing $s^2 \left\{ \tilde{p}_{h_{\text{new}}} \right\}$ in terms of MSE, station weights (w_i) and station elevations (h_i) as follows:

$$s^2 \left\{ p_{h_{\text{new}}} \right\} = \text{MSE} \cdot \left\{ 1 + \frac{1}{\sum w_i} + \frac{\left(h_{\text{new}} - \bar{h} \right)^2}{\sum \left(w_i h_i - \bar{h} \right)^2} \right\} \quad (7)$$

where i ranges over the stations involved in the regression and $\bar{h}$ is their weighted mean elevation.

The prediction interval (with confidence α) for the grid point with elevation h_{new} was then defined as:

$$p_{h_{\text{new}}} \pm t_{\frac{1-\alpha}{2}, df} \cdot s \left\{ p_{h_{\text{new}}} \right\} \quad (8)$$

where t is the value of a Student distribution with df degrees of freedom corresponding to a cumulative probability $(1 - \alpha)/2$. In this work, df was set to the number of stations considered in the linear regression, even though the choice of the appropriate value for df is not trivial (Daly *et al.*, 2008). In order to show one standard deviation around the model estimation, we set $(1 - \alpha) = 0.68$ and we called these prediction intervals PI68.

3.2. Regression kriging

In RK, a mixed approach is applied, combining a regression model of precipitation *versus* some chosen predictors (e.g. latitude, longitude and elevation) to a kriging-based geostatistical approach (Goovaerts, 2000). Since elevation is observed to be the most relevant predictor for precipitation data (see e.g. Secci *et al.*, 2010), we applied the following equation at each grid cell:

$$p(x, y) = a(x, y) + b(x, y) \cdot h(x, y) \quad (9)$$

where the coefficients $a(x, y)$ and $b(x, y)$ are estimated, for each grid cell, by the least-square method from the precipitation and elevation data at sample sites considering all the stations within a distance R from the considered grid point.

Then the station residuals from Equation (9) (ε) are interpolated on the grid by means of ordinary kriging (OK) and the precipitation value at each grid cell is estimated as:

$$p(x, y) = a(x, y) + b(x, y) \cdot h(x, y) + \mathbf{k}^{\mathbf{T}}(x, y) \cdot \varepsilon \quad (10)$$

where $\mathbf{k}(x, y)$ is the vector of kriging weights for the grid cell (x, y).

The exponential model was used to fit the semi-variogram of station residuals; the bin width was set to 10 km and all the station pairs within 300 km were considered. Moreover, an optimisation procedure was set up aiming at defining month-by-month the radius R of the area to consider for the precipitation-elevation regression and the best weights to be applied in the least-square estimation of the semi-variogram fit (Hengl, 2009). The R values providing the lowest station errors range from 125 to 200 km.

3.3. Other methods

In order to evaluate the benefit of considering the precipitation-elevation relationship at the local level, the LWLR and RK results were compared to those provided by two widely used interpolation techniques: IDW and OK. In the IDW approach, rainfall value at an unknown point is computed as a linear combination of a number of surrounding observations whose weights decrease with increasing distance from the point to estimate (Shepard, 1968). In our case, weights are defined by means of a Gaussian function whose decrease is locally optimized month-by-month following the same iterative procedure set up to estimate the best decreasing coefficients of LWLR weights.

4. Results

4.1. Performances of the interpolation models

LWLR, RK, OK and IDW were evaluated individually in terms of their ability to reconstruct the monthly 1961–1990 observed precipitation normals at station sites. More precisely, the monthly normals of the 4525 stations contained in the study domain were estimated by each model and then compared to the observed values. The reconstruction was performed in each case by means of the leave-one-out approach, i.e. by removing the station whose normals were being estimated, in order to avoid 'self-influence' of the station data to reconstruct. However, due to the requested computational time, the leave-one-out reconstruction in kriging-based models was performed by setting to 0 the kriging weight of the station to be estimated and by re-normalizing the remaining station weights, while the covariance matrix was obtained from the full dataset. The results of the comparison between estimated and observed values are listed in Table 2, where the accuracy of each method is expressed month-by-month in terms of mean error (BIAS, i.e. the mean difference between estimated and observed values), MAE and RMSE.

Bias values for LWLR, RK and OK are very low in all months, suggesting that these methods are not significantly affected by systematic errors when all the stations are considered, while IDW produces a small systematic positive bias indicating a global overestimation of station normals. MAE and RMSE averages over all months are smaller, and almost comparable, for LWLR (8.2 and 12.1 mm, respectively) and RK (8.1 and 11.9 mm) than for OK (8.8 and 12.9 mm) and IDW (9.1 and 13.3 mm). We also tried to subject the LWLR station residuals to OK in order to check whether errors could be further diminished. The semi-variogram (not shown) highlighted a very small

Table 2. Accuracy of the monthly climatologies obtained from the leave-one-out validation of the four methods for the 4525 stations in the study domain. All the values are expressed in mm.

	LWLR			RK			OK			IDW		
	BIAS	MAE	RMSE	BIAS	MAE	RMSE	BIAS	MAE	RMSE	BIAS	MAE	RMSE
1	0.0	9.6	14.2	−0.1	9.6	14.3	−0.1	10.7	16.2	0.4	10.9	16.3
2	0.0	8.8	12.9	−0.1	8.8	12.7	−0.1	9.9	14.5	0.4	10.0	14.5
3	−0.1	8.8	12.9	−0.1	8.8	12.8	−0.1	9.9	14.4	0.4	9.9	14.5
4	−0.1	8.5	12.9	0.0	8.6	12.6	−0.1	9.2	13.3	0.4	9.8	14.2
5	−0.2	7.4	11.5	0.0	7.5	11.5	0.0	8.1	12.3	0.3	8.5	12.9
6	−0.1	6.1	9.2	0.0	6.0	9.1	0.0	6.3	9.6	0.2	6.6	10.1
7	0.0	5.1	7.7	0.0	5.1	7.6	0.0	5.3	7.9	0.1	5.5	8.3
8	0.0	6.3	9.0	0.0	6.2	8.7	−0.1	6.4	9.0	0.2	6.6	9.4
9	0.0	7.1	10.2	0.0	7.1	10.0	−0.1	7.2	10.2	0.3	7.5	10.7
10	0.0	9.6	13.7	0.0	9.5	13.4	0.0	9.7	13.7	0.4	10.3	14.6
11	0.0	11.0	16.2	−0.1	10.9	15.7	−0.1	11.7	17.1	0.4	12.2	17.7
12	0.0	9.8	14.7	−0.1	9.8	14.6	−0.1	11.0	16.8	0.4	11.1	16.8

spatial variance suggesting that kriging interpolation is not useful to further reduce model errors.

The models were then investigated by evaluating the monthly bias of appropriate station clusters. More precisely, we first analysed both the station bias distribution for different 1° latitude bands covering the study domain and different intervals of other geographical parameters. The results (not shown) show that none of the methods has an evident bias for such station clusters. Then we selected the station subsets on the basis of both elevation and latitude, with thresholds chosen to consider the four combinations of high-level/low-level stations of Northern/Central-Southern Italy. We set 100 m a.s.l. as the low-level-station threshold, both for Northern and Central-Southern Italy, whereas for the high-level-station threshold we used 2000 m a.s.l. for Northern Italy and 1000 m a.s.l. for Central-Southern Italy. The box-plots of the monthly errors for these station subsets are reported in Figures 4(a), (b), 5(a) and (b). While LWLR and RK show quite similar results and their error medians are very close to zero in both cases, OK and IDW are affected by significant biases. In particular, high-level sites and low-level sites in Central-Southern Italy are under and overestimated, respectively, with the largest bias in winter (Figures 4(b) and 5(b)). A slight tendency to under/overestimate normals of high-level/low-level areas occurs also in Northern Italy, even though the biases are much less pronounced (Figure 4(a) and 5(a)).

These results indicate that when elevation is not considered, high-level/low-level station values, which are often computed by means of lower-level/higher-level neighbouring stations, are underestimated/overestimated. This problem is particularly relevant where the dependence of precipitation on elevation is more marked and confirms the importance of modelling the precipitation-elevation relationship, which has however a strong spatial variability over the Italian territory. In order to study it, we considered the distribution of the coefficients from the linear precipitation-elevation regression used in the RK procedure (see paragraph 3.2). The distribution of the grid point coefficients shows a very heterogeneous behaviour

in precipitation-elevation relationship in all months. The most marked variability occurs in winter when mean values range from 10 mm/km in the North (points above 43.7°N) to 70 mm/km in Southern Italy (points below 40.5°N). It is therefore quite evident that a global approach in modelling the precipitation-elevation relationship cannot be representative of the actual orographic influence on rainfall distribution over Italy.

We also compared the models by analysing the bias distribution for different ranges of station monthly normals in January and July (Figure 6). Even though all methods tend to underestimate the highest values, LWLR and RK are the least biased in January, while the performances of the different methods are more similar in July for any range of normals.

Therefore, the above observations suggest that considering the precipitation-elevation relationship and modelling it at a local scale, as performed in LWLR and RK, is the most suitable approach to produce Italian climatologies on a high-resolution grid. On the contrary, OK and IDW, by taking into account only station distance, turn out to be more suitable to interpolate data over regular domains where the local effects due to orographic features are negligible.

Even though LWLR and RK performances can be considered quite comparable, the two models show some peculiar features due to the different spatial scales at which the precipitation-elevation relationship is evaluated. Specifically, LWLR considers a very small scale, as the most relevant stations in the precipitation-elevation regression are generally located within 10 km from the considered grid point, whereas in RK the considered spatial scale is more than one order of magnitude larger (see paragraph 3.2). In this method, the small-scale effect is captured by the OK applied to the residuals from the precipitation-elevation relationship and if a too small scale were used to get this relationship, the kriging variogram would contain no significant signal.

Because of the small spatial scale adopted to perform the elevation-precipitation regression, LWLR shows the greatest ability in estimating the normal values for grid
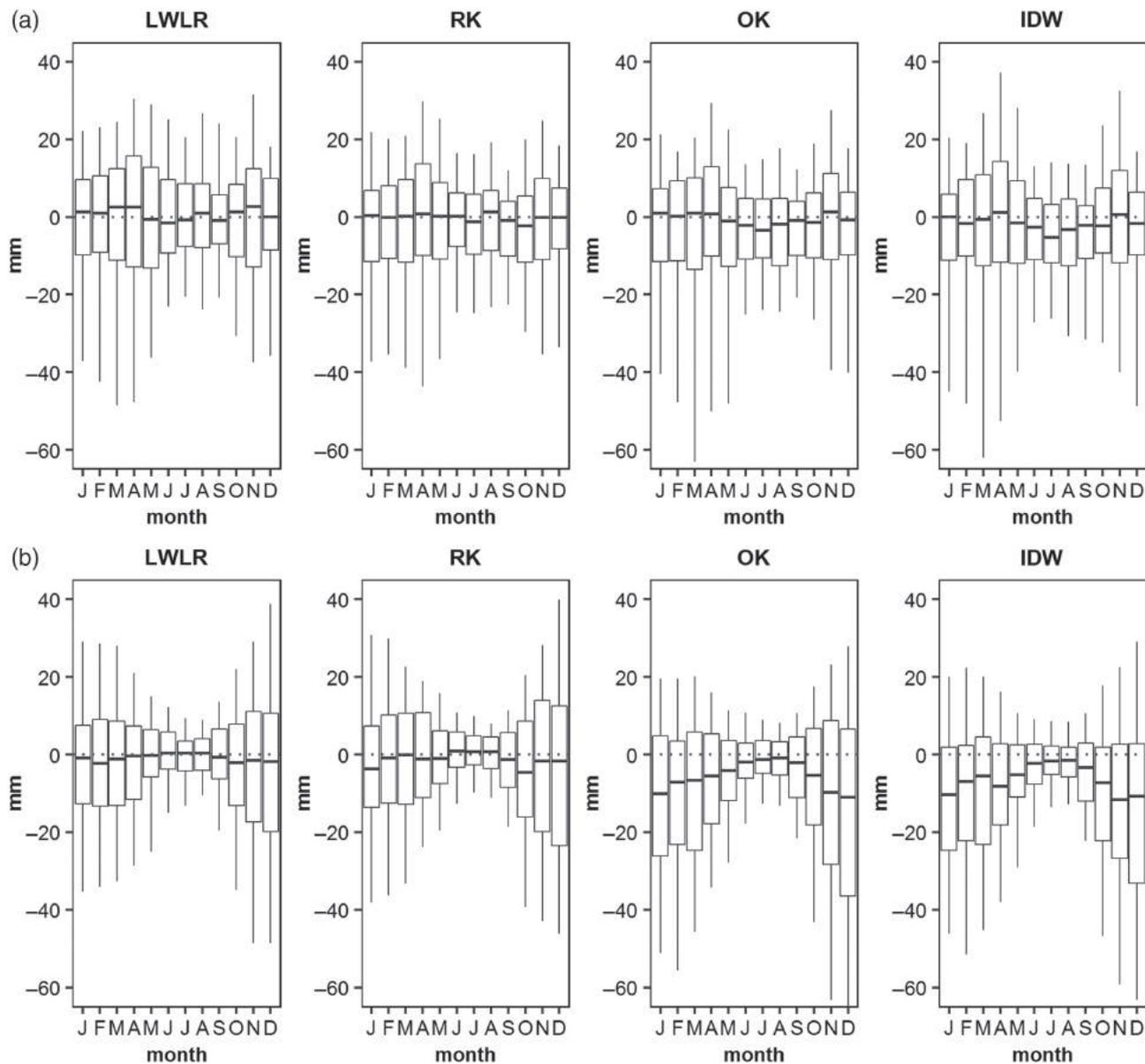
Figure 4. Monthly bias distribution of the reconstructed normals by the four methods for stations with (a) elevation >2000 m a.s.l. in Northern Italy and (b) elevation >1000 m a.s.l in Central-Southern Italy. The boxes range from the lower to the higher quartiles and are centred on the median; the whiskers represent the minimum and the maximum bias.

points at higher or lower elevation than the nearest stations. We checked this ability by computing the leave-one-out monthly errors of the stations that are located at higher or lower elevation than the ten neighbouring ones. Figure 7 shows the box-plots of the stations that are significantly higher (at least 50 m) than the neighbouring ones. For these stations LWLR turns out to be almost unbiased (the cumulated bias over all months is within −5 mm), whereas all the other models have greater bias with cumulated yearly values of about −50 mm for RK and about −150 mm for both IDW and OK. Lower mean values of MAE and RMSE also reflect the better performance of LWLR for these stations. Similar results are obtained for the stations that are significantly lower (at least 50 m) than the neighbouring ones, even though in this case the differences among the methods are less pronounced.

However, in spite of its better extrapolation ability, LWLR does not have lower errors than RK. This is because LWLR is found to be more affected by the low data availability in complex areas due to the small scale considered for the precipitation-elevation regression. When all station weights considered in the estimation of a certain grid point are negligible, the stations located far from it but showing a fortuitous similarity with the cell in few geographical features not dependent on the proximity of the station could prevail in the regression producing unrealistic results and marked discrepancy between precipitation values of adjacent cells. In these situations, even if the actual rainfall gradients of the areas cannot be precisely reconstructed due to the lack of stations, the RK interpolation generally provides smoother fields. An example of critical area for LWLR modelling is Mount Etna on the eastern coast of
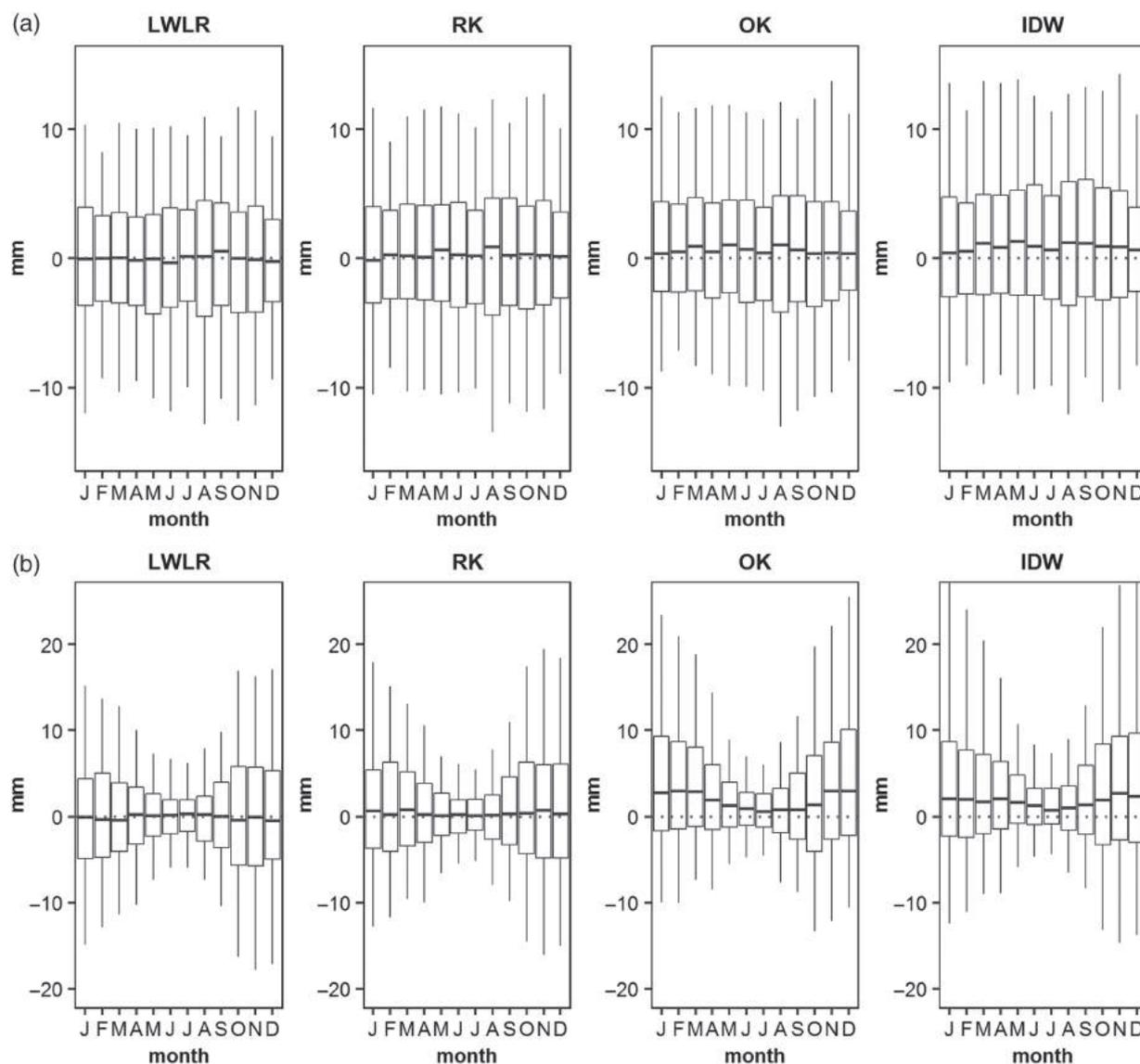
Figure 5. Monthly bias distribution of the reconstructed normals by the four methods for stations with elevation <100 m a.s.l. in (a) Northern Italy and (b) Central-Southern Italy. The boxes range from the lower to the higher quartiles and are centred on the median; the whiskers represent the minimum and the maximum bias.

Sicily. The region around Mount Etna (3329 m) is characterized by a remarkable pluviometric gradient with wetter conditions on the coastal area and drier conditions along the western inland side. As shown in Figure 8, where the map for January is shown as an example, the available stations are all located at the foot of the volcano and the highest station on the relief is installed at 1882 m on the southern side. Due to this poor station coverage, LWLR is not able to deal with the complexity of the precipitation variability of the area, and its output is a very heterogeneous rainfall distribution on the relief. A negative regression occurs in the grid cells close to the top, producing a band of low rainfall values. However, for the cells on the top, all station weights are negligible and the regression is only driven by those stations with similar slope orientation and, to a lesser extent, slope steepness, producing strong positive coefficients and, consequently, high precipitation values at the highest altitudes. On the contrary, RK

provides a more reasonable reconstruction and more reliable rainfall gradients.

A further difference between LWLR and RK is that the former method provides a straightforward approach to quantify model uncertainties for each grid point and thus to assess the reliability of the reconstructed precipitation fields. These uncertainties are expressed by means of the PI68 half-width, whose distribution over the domain for the central month of each season is reported in Figure 9. The PI68 half-widths calculated for the grid points closest to the stations turns out to be in good agreement with the leave-one-out station RMSE (Table 2) with a mean monthly difference of around 6%. However, as expected from the previous discussion, the greatest instability in the model results occurs where the interaction between circulations and very complex terrains leads to strong pluviometric gradients within rather limited areas, causing stations at the same elevation, but with different slope
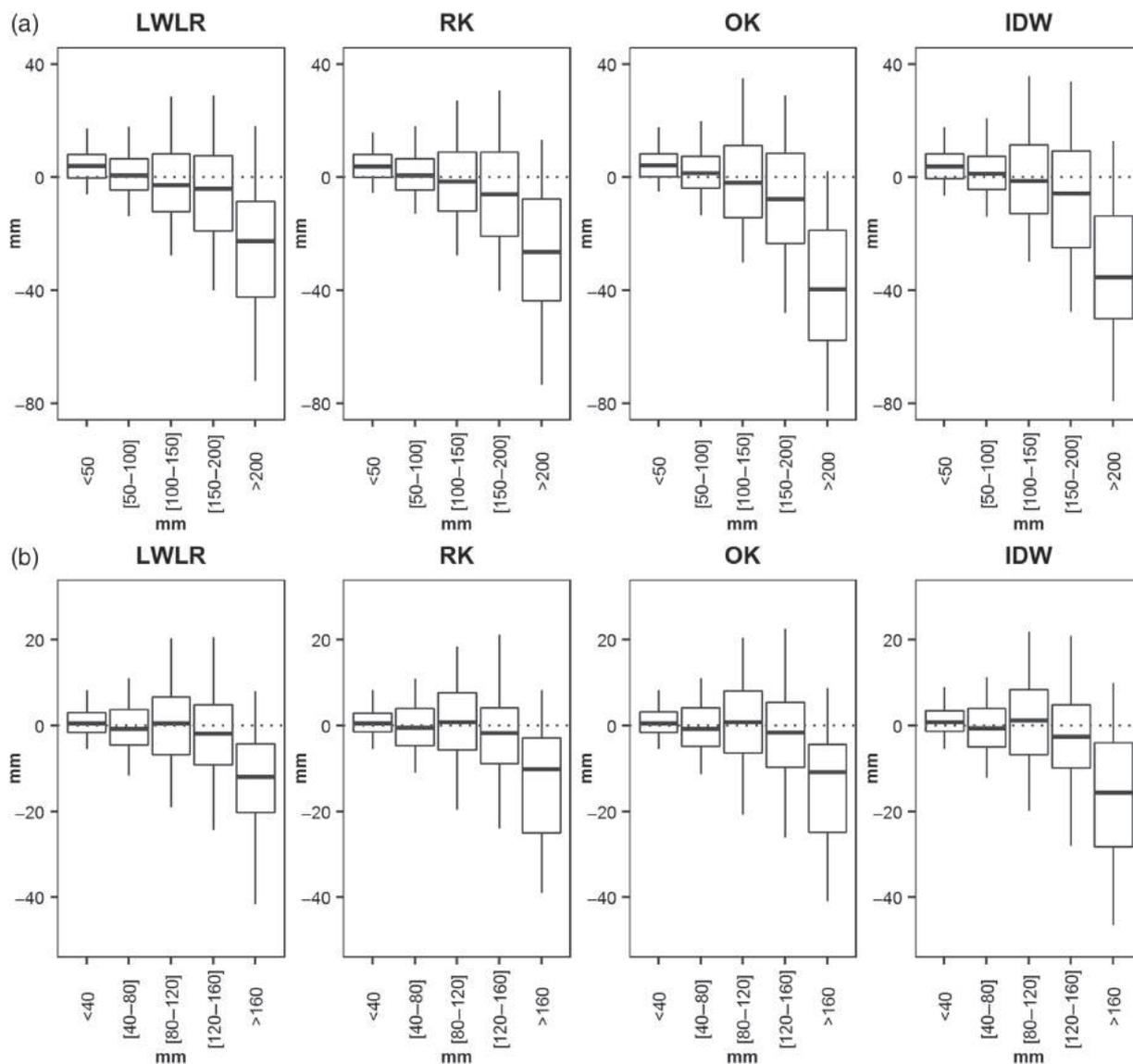
Figure 6. (a) January and (b) July bias distribution of the reconstructed normals by the four methods considering different ranges of station normals. The boxes range from the lower to the higher quartiles and are centred on the median; the whiskers represent the minimum and the maximum bias.

orientations, to have contrasting precipitation normals. This is mostly evident for Ligurian Apennines and for the southernmost part of the Apennine ridge, where slope orientation seems to have a much greater influence on precipitation than elevation, producing marked deviations from linear behaviour in the precipitation-elevation regression. Other areas exhibiting high PI68 values are Alpine and pre-Alpine regions, more evident during spring and autumn, and the Sicilian reliefs where the discussed issues in the precipitation reconstruction over Mount Etna are evident.

4.2. 1961–1990 high-resolution climatologies

LWLR and RK seasonal and annual precipitation climatologies are presented in Figures 10–13. The average seasonal LWLR and RK precipitation (winter–spring–summer-autumn) over the considered domain is very similar (253–243–179–289 mm and

252–242–178–288 mm, respectively). The agreement is generally good also for smaller areas and, splitting the domain into 1° sub-regions (centred on the nodes indicated in Figure 14), 47 out of the 50 sub-domains have LWLR–RK average differences in seasonal precipitation within 3% of LWLR–RK average seasonal area values. The most remarkable differences in seasonal amounts concern the region centred over grid point 7°E–46°N (northernmost part of Valle d'Aosta) which has 15–19% less precipitation for RK than for LWLR and those centred over grid points 11°E–47°N and 12°E–47°N which have 3–13% less precipitation for RK than for LWLR. For these sub-domains, however, more than half of the area falls outside Italy. Similar results are obtained even if we compare the two methods at a smaller scale of 0.5° areas. The agreement in terms of average seasonal precipitation is good also with OK and IDW (251–241–177–287 mm for both methods) whereas greater differences are present,
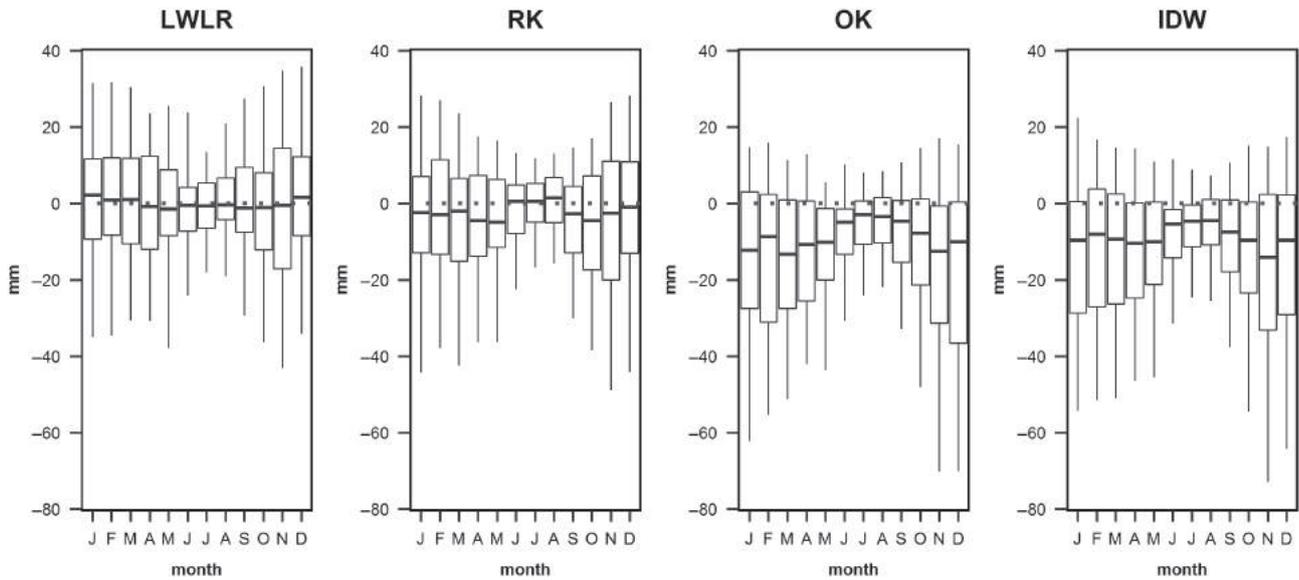
Figure 7. Monthly bias distribution of the reconstructed normals by the four methods for stations at higher elevation (at least 50 m) than the 10 nearest ones. The boxes range from the lower to the higher quartiles and are centred on the median; the whiskers represent the minimum and the maximum bias.
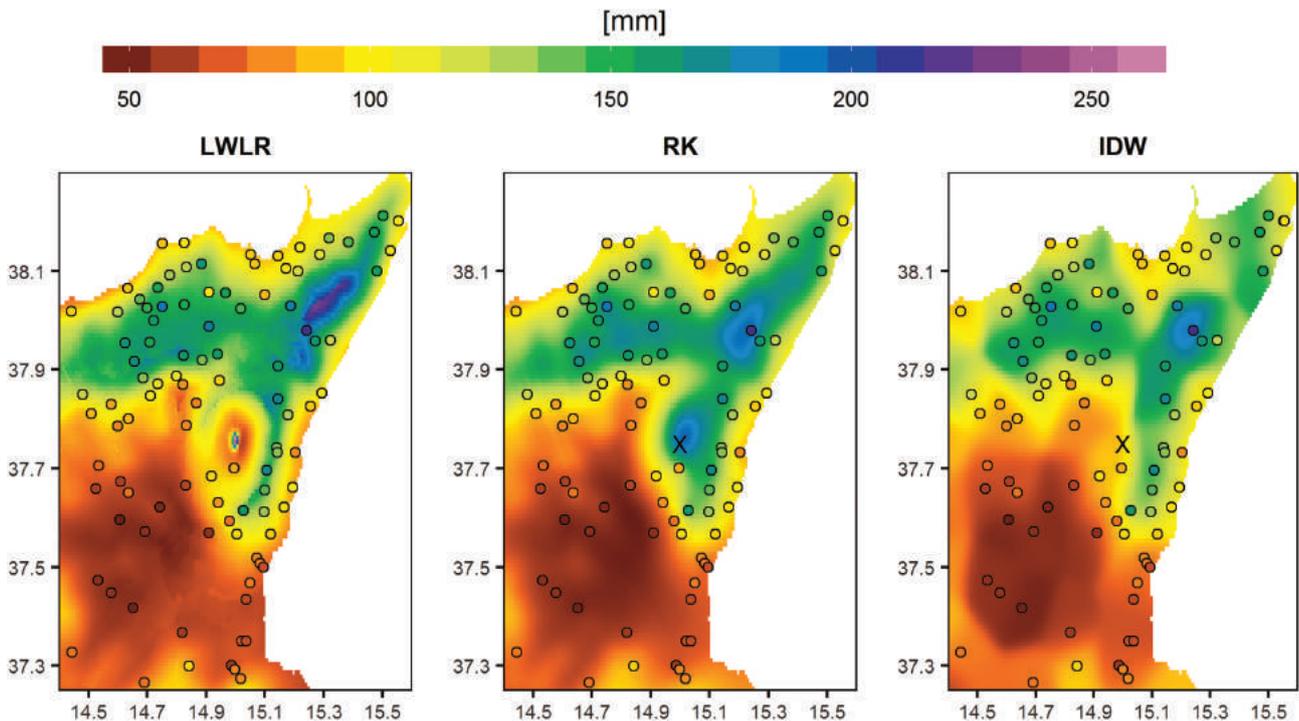

Figure 8. A detail of the January precipitation climatology over Mount Etna (Sicily) obtained by LWLR, RK and IDW. The points represent the available stations and the cross corresponds to the top of the volcano. OK climatology is not shown here as it provides a very similar precipitation distribution to that reported on IDW map.

especially in mountain areas, if we perform the comparison for 1° domains, bearing in mind the difficulties of these methods to deal with the precipitation-elevation relationship discussed above (see Figures 4 and 5). The agreement of LWLR and RK is good even when we consider the common variance of their precipitation fields: it is higher than 95% in all seasons, peaking up to 99% in summer. The agreement is slightly lower when LWLR is compared to IDW and OK with common variance ranging from 93% in winter to almost 99% in summer. On the contrary, RK shows a very good agreement with both OK and IDW in terms of common variance of precipitation fields, which is always above 98%. In addition to model agreement in terms of spatial precipitation fields, for all pairs of methods the correlation between yearly cycles reconstructed at each grid point was also considered. Yearly cycles provided by LWLR and RK result in very good agreement, with correlation coefficients very close
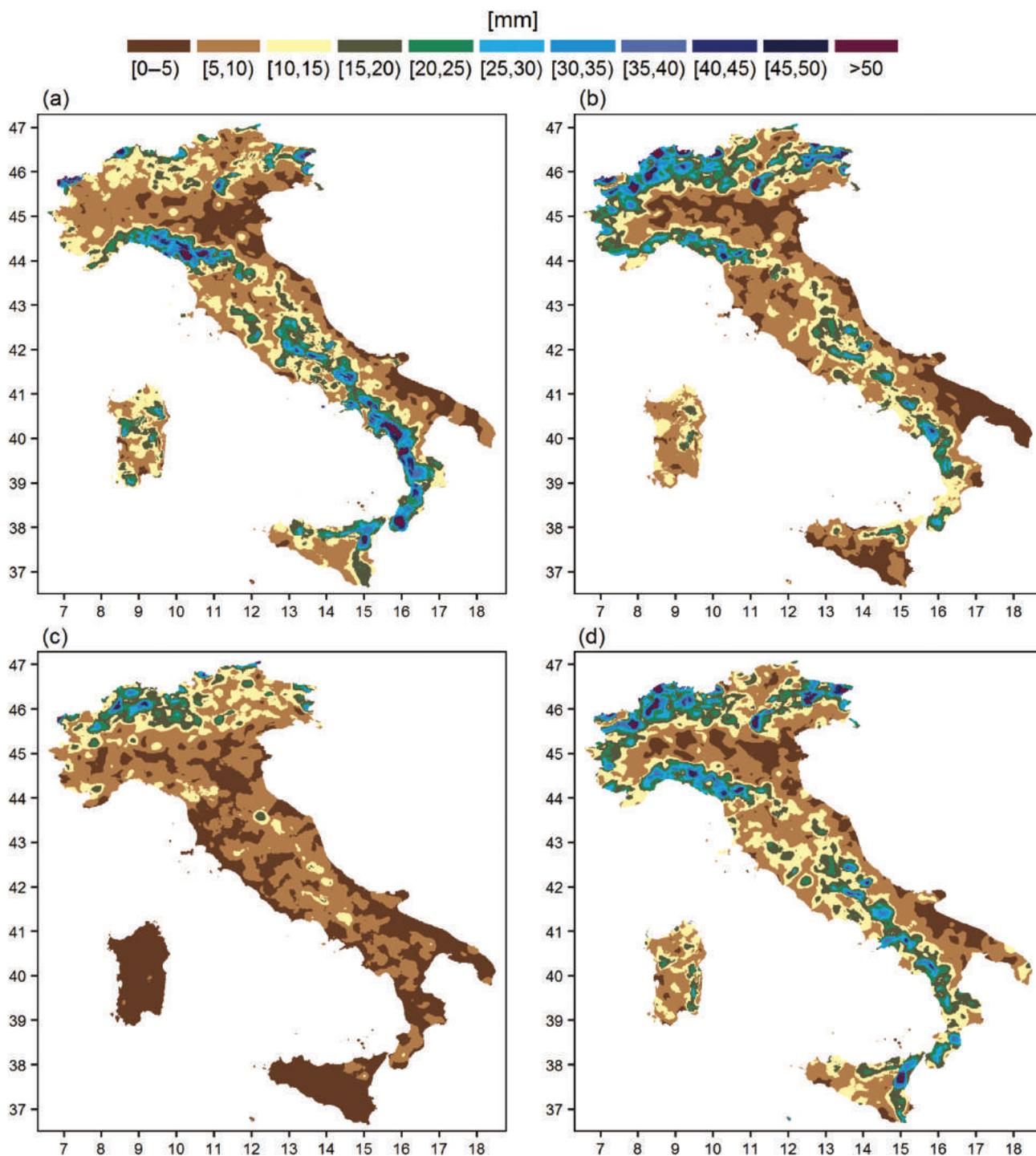
Figure 9. LWLR prediction interval half-width for (a) January, (b) April, (c) July and (d) October.

to 1 for most grid points and with 0.99 as average value over the domain. The agreement is very good with OK and IDW as well and the average of correlation coefficients over all grid points is 0.99 for any pair of methods.

Besides an overall good agreement, LWLR and RK climatologies also show interesting differences. As for the station normals (see Section 4.1), they are best emphasized when we focus on specific clusters of grid points. Limiting the comparison to the grid points that are at least 50 m higher than the ten closest stations, we get seasonal

LWLR average values of 185–292–315–285 mm, whereas the corresponding RK average values are 170–273–301–266 mm. These discrepancies are even greater if the same subsets of grid cells in OK and IDW climatologies are considered (their seasonal averages are 166–271–290–268 mm and 167–276–291–270 mm, respectively).

Some of the differences between LWLR and RK estimations seem therefore to depend on the better extrapolation ability of LWLR. In other cases, they could be
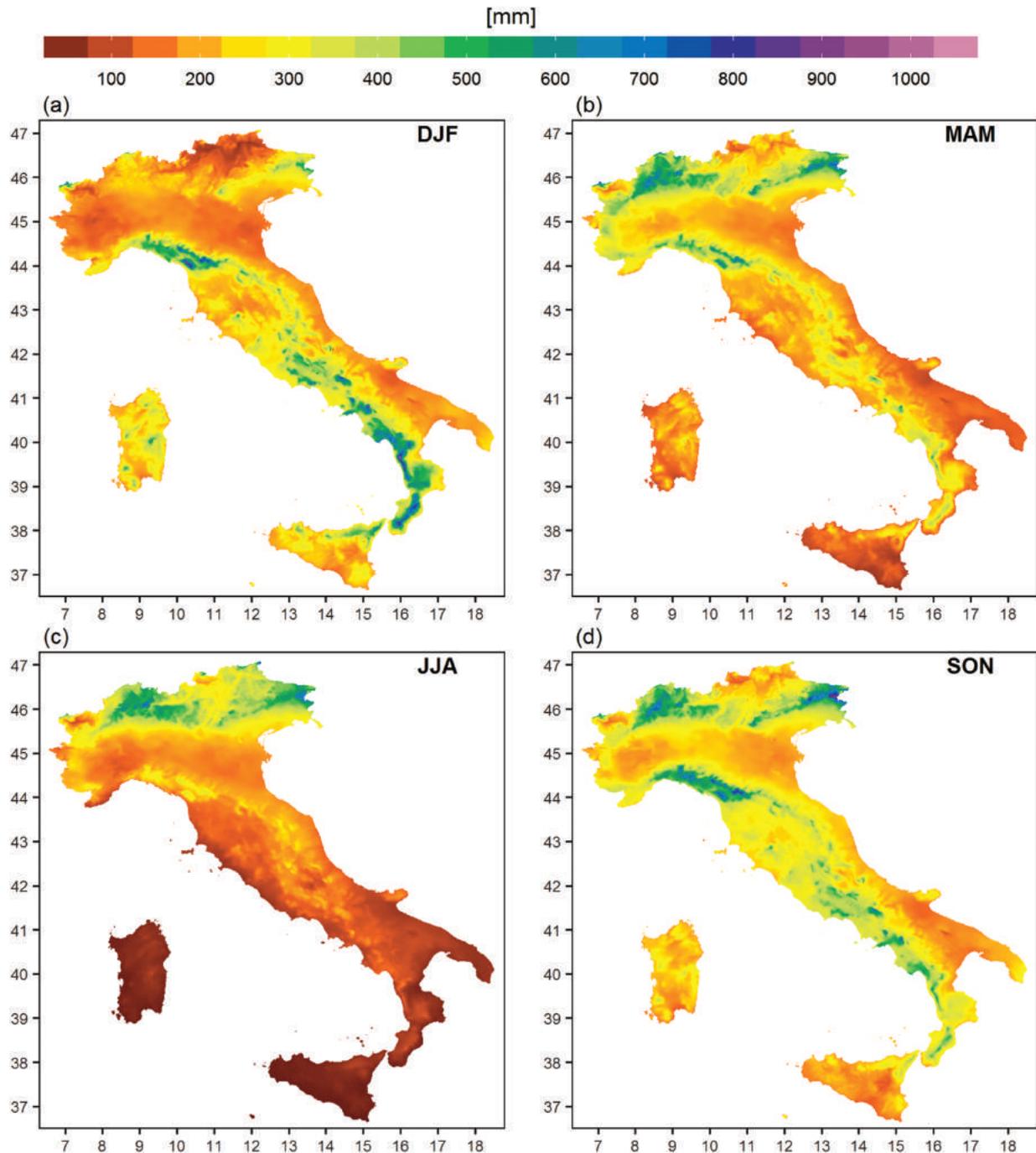
Figure 10. Seasonal LWLR precipitation climatologies.

due to the lower robustness of LWLR that may produce unrealistic results over areas with low data availability and with a complex precipitation-elevation relationship. An example concerning Mount Etna has been discussed in paragraph 4.1.

In spite of small differences, the LWLR and RK climatologies show the same main features, which are also present in OK and IDW climatologies. Both methods highlight a complex and strongly seasonal-dependent spatial distribution of precipitation. During the cold season, the Northern regions turn out to be generally drier than the Central-Southern areas, whereas the north-to-south gradient is reversed and more evident during summer, when very low values occur in Sicily, Sardinia and at the southernmost parts of the Italian peninsula, together with rather high values in the Alpine region.

Focusing on Northern regions, the precipitation-elevation gradient is very clear during spring and autumn, and even clearer in summer, especially from the Po Plain, with values lower than 250 mm, to the southern edge of Alps, where the summer averages range between 250 and 750 mm. During winter, these differences are much lower. In addition to elevation gradients, also the different exposure to moist-rich winds from south and
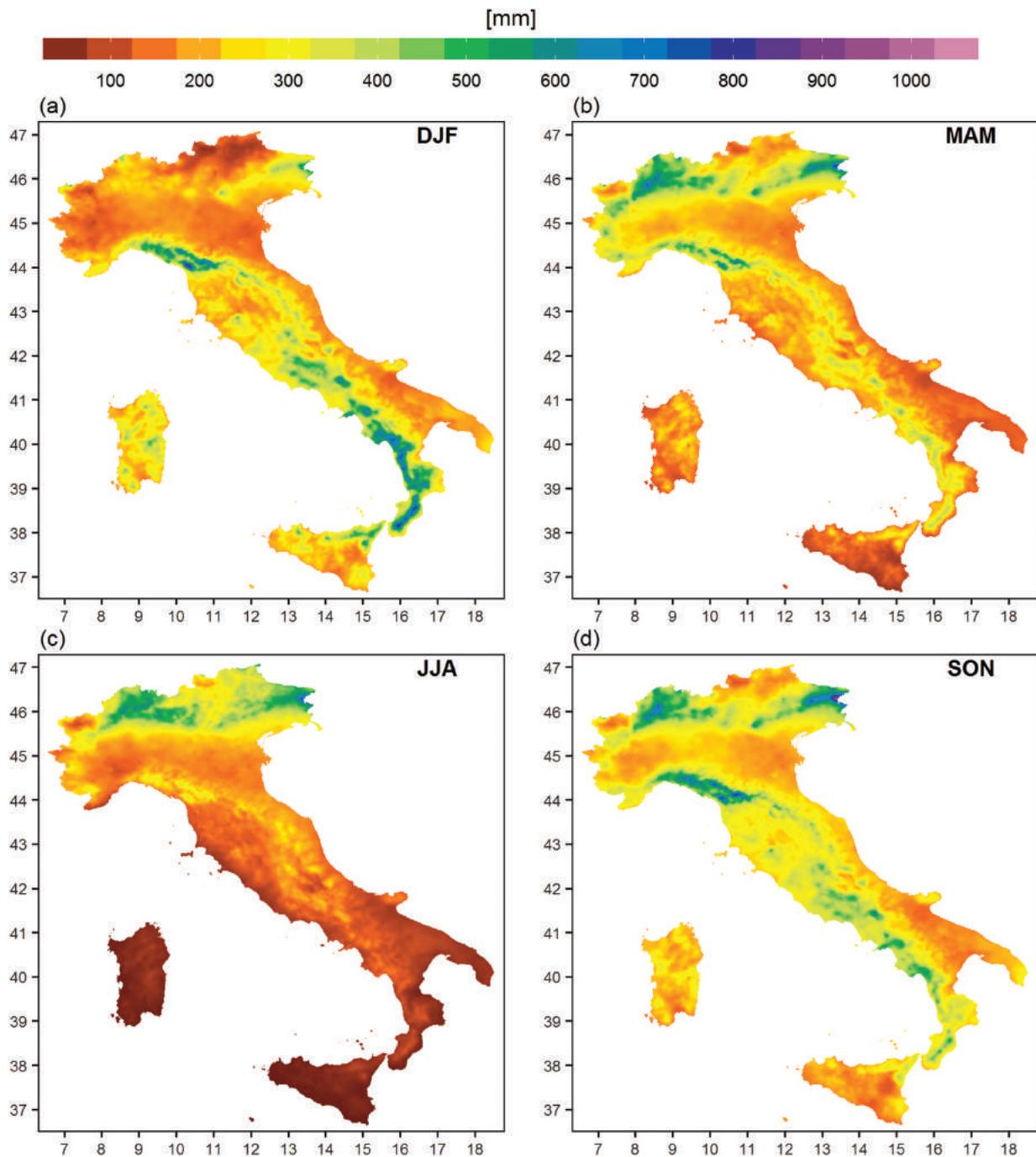
Figure 11. Seasonal RK precipitation climatologies.

south-west strongly influences precipitation distribution over Northern Italy. On the annual scale, the wettest sites in Northern Italy are located in Carnia (easternmost part of the domain), in Lepontine Alps and in the Apennines between Liguria and Tuscany with annual precipitation of slightly less than 3000 mm. Another very wet area in LWLR is the north-western part of Valle d'Aosta close to Grand Saint Bernard Pass. This area is however significantly less wet for RK, whose reconstruction has to be probably preferred here due to the very low station coverage of such a complex orography. Other very wet regions are Orobic Alps, central Ligurian Appenines

and Veneto Prealps where annual rainfall amounts reach 2000 mm. On the contrary, the Aosta plain and the inner Alpine valleys, particularly in the upper Adda and Adige river basins, feature drier conditions the whole year, with annual amounts around 500 mm. This feature is clearer in LWLR, whereas RK rather smooths it out.

As regards Central-Southern Italy, a west-to-east gradient is evident, especially during winter and autumn, with drier conditions along the Adriatic coast. Due to the short distance of the mountain chain from the sea, a noteworthy precipitation-elevation gradient occurs in all seasons along the southernmost part of Apennines
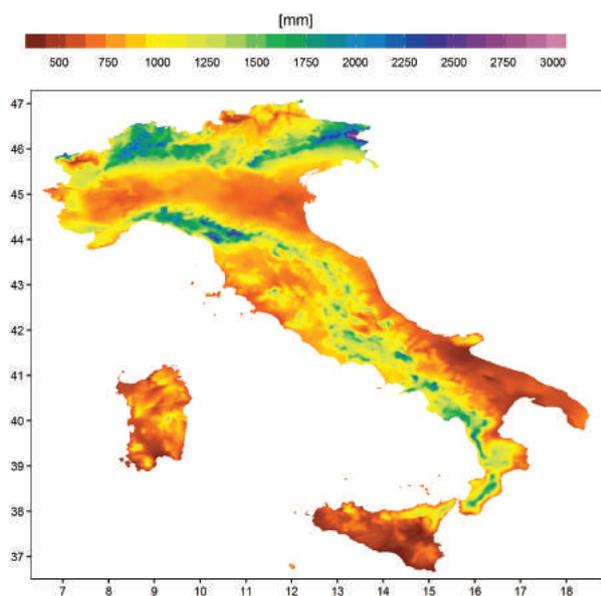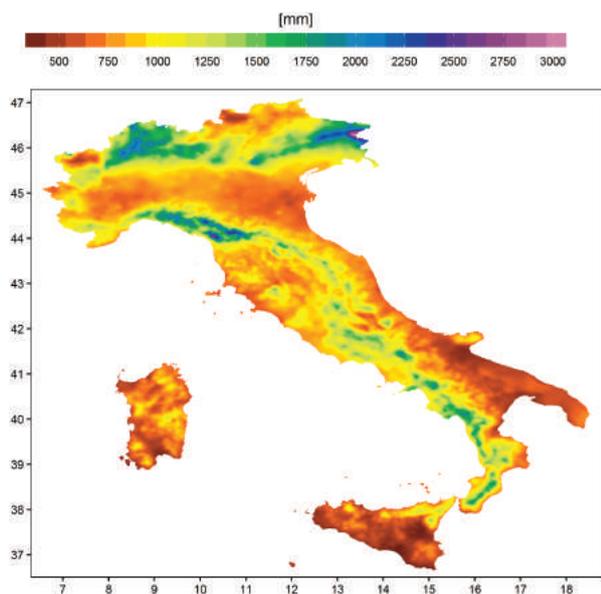
Figure 12. Annual LWLR precipitation climatology.

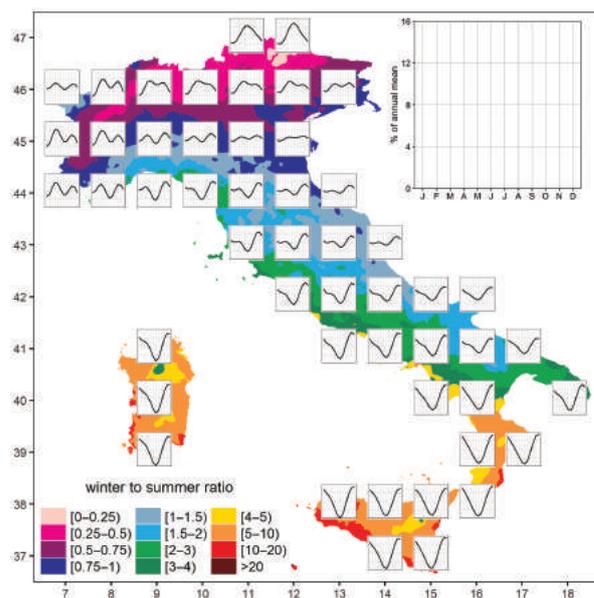

Figure 13. Annual RK precipitation climatology.

Figure 14. Distribution of winter (DJF) to summer (JJA) precipitation ratio over the domain and (superimposed) average yearly precipitation cycles over 1° sub-domains covering the whole study area. The inset box shows the range of the axes that is the same for all the plots.

(especially in Campania, Calabria and Basilicata) with weaker effects during summertime only. This feature is again more pronounced in LWLR, whereas RK smooths it a bit out. However, both climatologies describe the precipitation distribution over Italy with an improved spatial resolution in comparison with those available from ISPRA and CREA-CMA, which are the only other ones available at a national scale (ISPRA, 2014; Esposito *et al.*, 2015). If smaller regions are considered, the most important improvements with respect to existing regional climatologies concern Central-Southern Italy, where for many areas the available maps are based on poor data coverage and rather simple methods.

LWLR and RK climatologies also offer interesting information about Italian climate from the grid point precipitation annual cycles, whose patterns display a great heterogeneity over the domain. Figure 14 shows the average yearly precipitation cycles over 1° areas distributed throughout the Italian surface. They point out that the main Italian islands and the southern regions feature a very clear Mediterranean pattern with a precipitation maximum between late autumn and mid-winter and a very clear precipitation minimum in mid-summer, which becomes less marked moving from south to north and from the coast to the Apennines. On the eastern part of peninsular Italy, this pattern extends less to the north than on the western side, with the Adriatic coastal areas featuring a less pronounced precipitation decrease in summer and a shift of the highest contributions to autumn months. Central and Eastern Alps have a completely different pattern, with the highest precipitation in summer and the lowest one in winter. The remaining part of Northern Italy generally exhibits two maxima in spring and autumn and a decrease of rainfall in summer and winter. In this area, the prevailing of spring or autumn contribution defines the different climatic zones. A different behaviour is observed over the Po plain portion in North-Eastern Italy where annual cycles show a rather smooth pattern and quite constant rainfall contributions from all seasons. Figure 14 also shows the ratios between winter and summer precipitation normals at each grid point. They clearly show the Mediterranean climate of peninsular Italy, where winter to summer ratios are greater than 1 picking up to 20 in the southernmost Sicilian coasts, in contrast with the continental climate of Northern Italy, with values generally lower than 1 and largely below 0.5 over the North-Eastern Alps. Moreover, it is worth noting the distinction between the central Adriatic coastal areas, whose behaviour is more similar to that of northern Italian regions, from the

remaining part of peninsular Italy, where the decrease of summer precipitation is more pronounced.

5.   Conclusions

Monthly 30-arc-second-resolution precipitation climatologies for the 1961–1990 period were computed for Italy from a dense and quality-checked observational database by means of two interpolation approaches modelling the precipitation-elevation relationship at a local scale. More precisely, a LWLR and a local RK were performed. The ability of these methods to deal with the spatial distribution of precipitation normals over the Italian territory was evaluated by comparing their performances with those of two interpolating methods which do not apply any precipitation-elevation relationship: OK and IDW.

Even though all models reconstruct the station precipitation normals without significant global bias, LWLR and RK provide better results, with leave-one-out MAEs ranging from 5.1 mm (July) to 11 mm (November) for both models. Their better performance in reconstructing normals are even more evident when clustering stations by latitude and elevation, in particular considering high-level and low-level stations in Northern and Central-Southern Italy. While OK and IDW systematically show both overestimations and underestimations of the monthly precipitation, the median (and the mean) of LWLR and RK errors is almost null in every station subset, suggesting the greater suitability of the local approaches to describe the precipitation-elevation relationship within the orographically heterogeneous Italian domain. Although LWLR and RK have comparable performances, they show some specific features. In particular, LWLR is found to be strongly dependent on the domain features and on data availability. Especially over areas with complex orography and low station coverage, it may produce less reliable and discontinuous values, whereas RK provides more stable results. On the other hand, if the data density is high enough, LWLR shows a greater extrapolation ability, especially for the points located at higher elevation than the nearest stations. An additional advantage of LWLR is that it allows to directly define a prediction interval for each grid point of the study area which helps to identify the regions mostly affected by model uncertainty.

Despite small discrepancies, RK and LWLR high-resolution climatologies are in very good agreement and show very similar features. Seasonal and annual maps point out a heterogeneous and extremely seasonal dependent precipitation distribution, with remarkable spatial gradients between Northern and peninsular Italy, as well as between Tyrrhenian and Adriatic coasts. Moreover, the annual cycles reconstructed at the high-resolution grid and the distribution of the winter to summer precipitation ratios highlight a clear distinction among the different Italian climatic zones.

More data from new sites and updating the present database are needed to enhance the reconstruction of the most critical areas and to compute the climatological normals for more recent WMO reference periods (1971–2000 and 1981–2010). This requires the exploitation of the data collected in the last 20–25 years by the new automatic weather station networks set up by the Italian Regions that have to be merged with the data from the mechanical network of the former SI. One of the problems of this activity is that in some areas the quick transition to automatic stations and the change of most station sites make the merging of the two networks not easy, requiring the application of homogeneity tests to evaluate how the new records can properly be used to update the older ones. Thanks to the procedure we developed to estimate the precipitation normals of a period even when its data are partially or completely missing, the update of normals does however not require updating thousands of series, but only a subset of selected stations allowing to capture the spatial pattern of the temporal variability of precipitation over Italy.

References

Antolini G, Auteri L, Pavan V, Tomei F, Tomozeiu R, Marletto V. 2016. A daily high-resolution gridded climatic data set for Emilia-Romagna, Italy, during 1961-2010. *Int. J. Climatol.* **36**: 1970–1986. https://doi.org/10.1002/joc.4473.

Basist A, Bell GD, Meentemeyer V. 1994. Statistical relationships between topography and precipitation patterns. *J. Clim.* **7**: 1305–1315.

Biancotti A, Bellardone G, Bovo S, Cagnazzi B, Giacomelli L, Marchisio C, Motta L, Oliviero A, Rossino M, Turroni E, Vittorini S. 1998. Distribuzione regionale di piogge e temperature. Torino, Regione Piemonte, 80 pp., *Studi Climatologici in Piemonte* (in Italian).

Brunetti M, Buffoni L, Mangianti F, Maugeri M, Nanni T. 2004. Temperature, precipitation and extreme events during the last century in Italy. *Glob. Planet. Change* **40**(1–2): 141–149.

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

Brunetti M, Maugeri M, Nanni T, Auer I, Böhm R, Schöner W. 2006b. Precipitation variability and changes in the greater alpine region over the 1800-2003 period. *J. Geophys. Res.* **111**: D11107. https://doi.org/10.1029/2005JD006674.

Brunetti M, Lentini G, Maugeri M, Nanni T, Simolo C, Spinoni J. 2009. Estimating local records for northern and central Italy from a sparse secular temperature network and from 1961–1990 climatologies. *Adv. Sci. Res.* **3**: 73–80.

Brunetti M, Lentini G, Maugeri M, Nanni T, Simolo C, Spinoni J. 2012. Projecting north eastern Italy temperature and precipitation secular records onto a high resolution grid. *Phys. Chem. Earth* **40-41**: 9–22.

Brunetti M, Maugeri M, Nanni T, Simolo C, Spinoni J. 2014. High-resolution temperature climatology for Italy: interpolation method intercomparison. *Int. J. Climatol.* **34**: 1278–1296.

Cati L. 1981. Idrografia e idrologia del Po. Pubblicazione n. 19 dell'Ufficio Idrografico del Po. Roma, *Istituto Poligrafico e Zecca dello Stato* (in Italian).

Daly C. 2006. Guidelines for assessing the suitability of spatial climate data sets. *Int. J. Climatol.* **26**: 707–721.

Daly C, Neilson RP, Phillips DL. 1994. A statistical-topographic model for mapping climatological precipitation over Mountanious terrain. *J. Appl. Meteorol.* **33**: 140–158.

Daly C, Gibson WP, Taylor GH, Johnson GL, Pasteris P. 2002. A knowledge based approach to the statistical mapping of climate. *Clim. Res.* **22**: 99–113.

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

Di Piazza A, Lo Conti F, Noto LV, Viola F, La Loggia G. 2011. Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. *Int. J. Appl. Earth Obs. Geoinf.* **13**: 396–408.

Diodato N, Ceccarelli M. 2005. Interpolation processes using multivariate geostatistics for mapping of climatological precipitation mean in the sannio mountains (southern Italy). *Earth Surf. Process. Landf.* **30**: 259–268.

Drago A. 2005. Atlante climatologico della Sicilia – Seconda edizione. *Rivista Italiana di Agrometeorologia* **2**: 67–83 (in Italian).

Esposito S, Beltrano MC, De Natale F, Di Giuseppe E, Iafrate L, Libertà A, Parisse B, Scaglione M. 2015. Atlante italiano del clima e dei cambiamenti climatici. Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, Unità di ricerca per la climatologia e la meteorologia applicate all'agricoltura. Roma, pp. 264.

Frei C, Schär C. 1998. A precipitation climatology of the alps from high-resolution rain gauge observations. *Int. J. Climatol.* **18**: 873–900.

Goovaerts P. 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.* **228**: 113–129.

Hengl T. 2009. A Practical Guide to Geostatistical Mapping, ISBN 978–90–9024981-0. Licensed under a creative Commons Attribution-Noncommercial-No Derivative Works 3.0 license. Available at http://spatial-analyst.net/book/.

Isotta FA, Frei C, Weilguni V, Tadić MP, Lassègues P, Rudolf B, Pavan V, Cacciamani C, Antolini G, Ratto SM, Munari M, Micheletti S, Bonati V, Lussana C, Ronchi C, Panettieri E, Marigo G, Vertačnik G. 2014. The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-alpine rain-gauge data. *Int. J. Climatol.* **34**: 1657–1675.

ISPRA. 2014. Valori climatici normali di temperatura e precipitazione in Italia, Rome. ISBN: 978-88-448-0689-7 (in Italian).

Martinez-Cob A. 1996. Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. *J. Hydrol.* **174**: 19–35.

Masson D, Frei C. 2014. Spatial analysis of precipitation in a high-mountain region: exploring methods with multi-scale topographic predictors and circulation types. *Hydrol. Earth Syst. Sci.* **18**: 4543–4563.

Schwarb M. 2000. The Alpine precipitation climate: evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data. PhD thesis 13911, Swiss Federal Institute of technology (ETH), Zurich, Switzerland, 131 pp..

Secci D, Patriche CV, Ursu A, Sfîcă L. 2010. Spatial interpolation of mean annual precipitations in Sardinia. A comparative analysis of several methods. *Geographia Technica* **1**: 67–75.

Servizio Idrografico. 1957. Ministero dei Lavori Pubblici: Precipitazioni medie mensili ed annue e numero di giorni piovosi per il trentennio 1921–1950. Istituto Poligrafico dello Stato, 1-12c (in Italian).

Shepard D. 1968. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 ACM National Conference*, Blue RBS, Rosenberg AM (eds). ACM Press: New York, 517–524.

Spinoni J. 2010. 1961–90 high-resolution temperature, precipitation, and solar radiation climatologies for Italy, Ph.D. thesis , Milan University. http://air.unimi.it/bitstream/2434/155260/2/phd_unimi_R07883_1.pdf.

USGS (United States Geological Survey). 1996. GTOPO30 documentation. http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info.

Vicente-Serrano SM, Saz-Sanchez MA, Cuadrat JM. 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Clim. Res.* **24**: 161–180.