UNIVERSITÀ DEGLI STUDI DI MILANO

PhD Course in Environmental Sciences

XXXI Cycle

Precipitation climatologies and long-term record reconstruction: studies in a changing climate

Scientific tutor
Prof. Maurizio MAUGERI

Scientific co-tutor
Dr. Michele BRUNETTI PhD

PhD Thesis
Alice CRESPI
R11342-R24

Academic Year 2017-2018

This thesis was performed at the Department of Environmental Science and Policy, Università degli Studi di Milano (Milan, Italy), in collaboration with the Institute of Atmospheric Sciences and Climate - National Research Council (ISAC-CNR, Bologna, Italy).
SSD: FIS/06

Contents

Abstract

The availability of high-resolution datasets describing the spatio-temporal evolution of precipitation is becoming increasingly important in order to analyse the long-term variability and trends of the climatic signal and possible impacts over specific areas of interest. These datasets are therefore crucial not only for research purposes, but also for decision-makers in a wide range of fields, such as agriculture, energy production, hydrology, natural risk monitoring and resource management. The reconstruction of high-resolution climate descriptions requires both accurate in-situ observations and suitable interpolation schemes to project the station data onto regular grids.
The study focuses on the development and improvement of interpolation methods for monthly precipitation data and on the reconstruction and analysis of gridded datasets from dense and high-quality rain-gauge records covering specific study domains.
The first goal of the work was the reconstruction of the high-resolution monthly precipitation climatologies over Italy for the 1961-1990 period. The observation database was set up thanks to a relevant effort of data rescue and collection from a great number of national and international sources and led to more than 4500 quality-checked monthly records available for the climatological reconstruction. Considering the heterogeneous Italian orography and the influence of surface features on precipitation distribution, a local weighted linear regression (LWLR) of precipitation versus elevation was applied to interpolate the station monthly normals onto a 30-arc second resolution Digital Elevation Model (DEM). Leave-one-out (LOO) validation and inter-comparison proved that the approaches modelling the local precipitation-orography relationship provide more accurate results in respect with methods considering larger spatial scales or not including the topography at all. The computed 30-arc second resolution monthly precipitation climatologies for Italy were on-line released and they represent an updated and highly detailed description of precipitation normals over the whole national territory in digital form.
In order to evaluate the temporal evolution of precipitation over some of the most vulnerable Italian regions, gridded datasets of long-term precipitation series were produced by means of the anomaly-based approach. The secular precipitation series were reconstructed for the upper part of Adda river basin (Central Italian Alps), with an additional focus over the Forni glacier, and for Sardinia, as case-study for the Mediter-

ranean area. New records were collected, especially thanks to the integration of the recent automatic station records with those of the previous mechanical networks and to digitisation activities of the most ancient data from hardcopy archives. All the series underwent statistical procedures aiming at avoiding inhomogeneities due to non-climatic signals. The gridded dataset allowed to get the secular areal precipitation records for the study regions and to evaluate their trends. As regards the Adda basin, the reconstructed 1800-2016 series showed statistically significant negative trends for annual and autumn precipitation. The comparison with the 1845-2016 annual basin runoff record, which is one of the longest runoff series available in Italy, allowed both to depict a strong decrease in annual runoff driven by the increase of evapotranspiration and to evaluate the possible contribution of gauge undercatch, especially in mountainous sites, to the underestimation of basin precipitation.
The trend analysis over the 1922-2011 areal monthly precipitation record computed for Sardinia highlighted statistically significant decreases of -2.3% and -4.1% decade^{-1} in annual and winter values, respectively, and a positive but not statistically significant trend in summer precipitation. These outcomes agreed with other literature studies focusing on precipitation variability over Mediterranean area.
The final part of the work focused on the development of interpolation schemes to improve the accuracy of gridded precipitation fields for domains unevenly covered by station networks. Norway represents a very interesting case-study, where the severe climatic conditions and the complex orography limit the management of in-situ observations over the most remote regions leading to an unbalanced station distribution between North and South and between low and high elevation. At this aim, a new method to compute the Norwegian monthly precipitation climatologies (1981-2010) at 1 km resolution was implemented and the gridded dataset was on-line released for both research and operative purposes. In this scheme, named HCLIM+RK, the HCLIM-AROME climate numerical model fields, which are not based on rain-gauge data and describe the precipitation gradients also over unsampled areas, are combined with available in-situ observations by the kriging interpolation of station residuals. In HCLIM+RK the high-resolution reliable precipitation patterns provided by the numerical model are retained and the station data are used to correct the biases affecting the numerical fields. The LOO reconstruction errors of Norwegian station normals showed that the combined approach almost removes the biases affecting the original HCLIM-AROME dataset and it provides much lower errors than conventional interpolation procedures based on stations only.

Riassunto

La disponibilità di campi ad alta risoluzione che descrivono l'evoluzione spazio-temporale delle piogge sta assumendo una rilevanza crescente per l'analisi delle tendenze e della variabilità a lungo termine del segnale climatico e per la valutazione degli impatti su specifiche aree di interesse. Questi dataset rappresentano pertanto strumenti fondamentali non solo per scopi di ricerca, ma anche per scopi operativi e decisionali in diversi settori, quali l'agricoltura, la produzione energetica, l'idrologia e la gestione dei rischi e delle risorse naturali. La ricostruzione spazio-temporale del segnale climatico richiede accurati dati osservativi e schemi di interpolazione idonei a proiettare i valori puntuali su griglie regolari ad alta risoluzione.
Il presente studio si concentra sullo sviluppo di metodi interpolativi per valori di pioggia mensili e sulla ricostruzione ed analisi di dataset grigliati prodotti su specifiche aree di studio a partire da numerose e accurate serie osservative. Il primo obiettivo del lavoro ha riguardato la ricostruzione delle climatologie italiane ad alta risoluzione delle precipitazioni mensili per il periodo 1961-1990. Il database osservativo utilizzato è stato prodotto grazie ad un intenso lavoro di recupero e raccolta dei dati da numerose fonti nazionali ed extra-nazionali che ha portato a più di 4500 serie mensili disponibili per l'analisi climatologica, la cui qualità è stata valutata mediante specifici test statistici. Considerando l'eterogeneità dell'orografia italiana e l'influenza delle caratteristiche geografiche sulla distribuzione spaziale delle piogge, le normali mensili 1961-1990 delle stazioni sono state interpolate su un modello di elevazione digitale a risoluzione di 30 secondi d'arco mediante una regressione lineare pesata locale di precipitazione-quota. La validazione *leave-one-out* (LOO) ed il confronto hanno mostrato che i metodi che includono l'effetto dell'orografia a scala locale forniscono risultati più accurati rispetto a schemi interpolativi che considerano scale spaziali più ampie o che trascurano l'influenza orografica. Le climatologie 1961-1990 di precipitazione mensile per l'Italia ottenute e messe a disposizione on-line rappresentano una descrizione aggiornata e di elevato dettaglio spaziale delle normali pluviometriche sul territorio nazionale in formato digitale.
Al fine di valutare l'evoluzione temporale delle precipitazioni su alcune delle aree italiane più vulnerabili ai cambiamenti climatici, griglie regolari di lunghe serie pluviometriche sono state costruite mediante il metodo delle anomalie. In particolare, serie

secolari di piogge sono state ricostruite per la porzione superiore del bacino dell'Adda (Alpi centrali italiane), con un ulteriore focus sul ghiacciaio dei Forni, e per la Sardegna, scelta come caso di studio per l'area mediterranea. Nuovi dati sono stati raccolti grazie anche all'integrazione delle misure delle recenti stazioni automatiche con quelle delle precedenti reti meccaniche regionali e ad attività di digitalizzazione dei dati più antichi disponibili solo su supporti cartacei. Test statistici sono stati applicati alle serie al fine di eliminare eventuali disomogeneità dovute a fattori non climatici ed i valori interpolati hanno consentito di calcolare le serie areali di pioggia e di valutarne i trend. La serie 1800-2016 di piogge areali per il bacino dell'Adda ha mostrato decrescite statisticamente significative sia a scala annuale che per la stagione autunnale. Il confronto con la serie 1845-2016 di deflussi annuali, tra le più lunghe serie italiane di deflussi da bacino, ha permesso di evidenziare un significativo calo dei deflussi a lungo termine causato dall'aumento dell'evapotraspirazione e di valutare l'impatto della difficoltà di misurare correttamente le precipitazioni nevose da parte dei pluviometri, soprattutto nelle aree montane, sulla possibile sottostima degli afflussi al bacino.
La serie 1922-2011 delle piogge areali per la Sardegna ha mostrato decrescite statisticamente significative di -2.3% e -4.1% decade^{-1} rispettivamente per i valori annuali e invernali ed una tendenza positiva ma non significativa per le piogge estive. Questi risultati sono in accordo con altri studi di letteratura che analizzano la variabilità delle precipitazioni sulle regioni mediterranee.
L'ultima parte del lavoro ha riguardato lo sviluppo di metodi di interpolazione che consentano di migliorare l'accuratezza dei campi di precipitazione per regioni in cui la copertura della rete meteorologica è disomogenea sul territorio. La Norvegia rappresenta un caso di studio molto interessante dove il clima severo e la complessa orografia ostacolano la gestione di stazioni di misura in aree remote provocando un forte squilibrio nella disponibilità dati tra Nord e Sud e tra le aree di bassa e alta quota. Per questo motivo, le climatologie norvegesi di piogge mensili (1981-2010) a 1 km di risoluzione sono state calcolate mediante un nuovo metodo interpolativo e rese disponibili online sia per scopi di ricerca che operativi. In questo schema, denominato HCLIM+RK, i campi di pioggia forniti dal modello numerico HCLIM-AROME, che non utilizza i dati delle stazioni e descrive i gradienti pluviometrici anche per le aree più remote, vengono combinati con i dati osservativi disponibili mediante una procedura interpolativa basata sul kriging dei residui. In questo modo, si conservano i gradienti ad elevata risoluzione forniti dal modello numerico HCLIM-AROME ma correggendone il bias grazie alle osservazioni. Gli errori di ricostruzione delle normali mensili delle stazioni norvegesi hanno dimostrato che il metodo HCLIM+RK consente di rimuovere il bias del dataset originale del modello numerico e migliora l'accuratezza dei risultati rispetto a schemi interpolativi tradizionali basati esclusivamente sui dati osservativi.

Chapter 1

Introduction

1.1 The climate system and the study of climate change

According to the definition reported by the Working Group I in the Fifth Assessment IPCC (Intergovernmental Panel on Climate Change) report (IPCC, 2013), climate is usually defined as the average weather or, more rigorously, as the statistical description in terms of the mean and variability of relevant quantities over a period of time ranging from months to thousands or millions of years. The considered quantities are most often surface variables such as temperature or precipitation, while 30 years is the usual averaging period, as defined by the World Meteorological Organization (WMO). The climate should be distiguished from the weather, which is the description of the atmosphere conditions at a certain time and place with reference to the meteorological elements, e.g. temperature, pressure, humidity and wind. The term *climate change* in IPCC usage refers to a change in the state of the climate that can be identified (e.g. using statistical tests) by changes in the mean and/or by the variability of its properties, and that persists for an extended period, typically decades or longer. It refers to any change in climate over time, whether due to natural variability or as a consequence of human activity. The *climate system* is in fact the result of the complex interaction among atmosphere, hydrosphere, cryosphere, land surface and biosphere (Figure 1.1).
The climate processes on Earth are mainly driven by the solar radiation. Almost half of the incoming solar radiation, i.e. the shortwave radiation, is absorbed by the Earth's surface, while the remaining fraction is reflected back by clouds, by the atmosphere and by the Earth's surface itself or directly absorbed by the atmosphere. Earth, atmosphere and clouds emit themselves the absorbed radiation in form of infrared radiation (longwave) which is absorbed in large part by the so-called greenhouse gases (GHG_s), such as water vapour, carbon dioxide (CO_2), nitrous oxide (N_2O), and methane (NH_4), which are natural parts of the atmospheric composition. Atmosphere emits in turn infrared radiation in all directions including a downward component to the Earth's

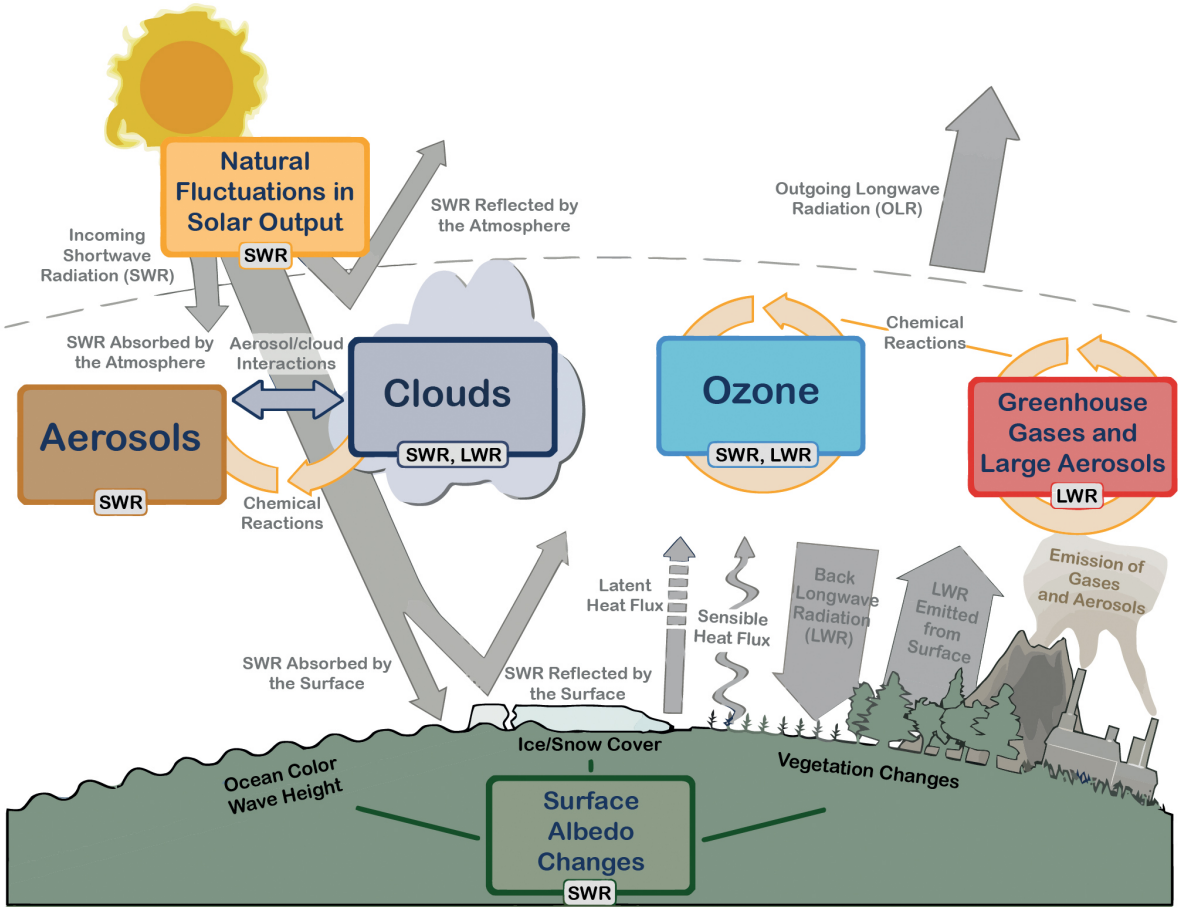

Figure 1.1: The main components of climate system.

surface which produces a heating effect of the surface and of the low atmospheric layers. This mechanism is called the *natural greenhouse effect* and it allows the temperature of the Earth's surface to be kept at a temperature of on average 15° C. Due to the Earth curvature, tropical and subtropical regions receive the greatest fraction of incoming solar energy. The global circulation is therefore determined by the atmospheric and oceanic transport processes which partly redistribute the energy surplus of tropical regions towards the Poles.

All natural and anthropogenic changes that affect the net shortwave and/or longwave radiation, i.e. the climate forcings, could alter the Earth's radiative equilibrium leading to temperature increments or decreases. The natural climate forcings include changes in the solar radiation due to Solar activity (Crowley, 2000), Milankovitch cycles, i.e. small variations in Earth's orbit and its axis of rotation occurring on millenial scale (Berger, 1988), natural oscillations, such as the ENSO (El-Niño Southern Oscillation) and the NAO (North Athlantic Oscillation) (Hurrel and van Loon, 1997; Beniston and Jungo, 2002) and large volcanic eruptions that inject light-reflecting particles into the stratosphere (Stenchikov et al., 1998). Anthropogenic forcings concern the increment

in aerosol concentration, which alters incoming and outgoing radiation (Twomey et al., 1984), the land-use changes, e.g. deforestation, which modify the surface albedo, i.e. its reflectance, and the increment in the emissions of greenhouse gases in the atmosphere, which enhances its ability in absorbing and emitting backwards the longwave radiation (Kvalevåg and Myhre, 2007). Each forcing can in turn trigger feedbacks that intensify or weaken the original effect on climate. Since the second half of the 20^{th} century, a significant increase in the mean global temperature of 0,12° C per decade (1951-2012) has been registered (Figure 1.2) (IPCC, 2013). The remarkable impact of

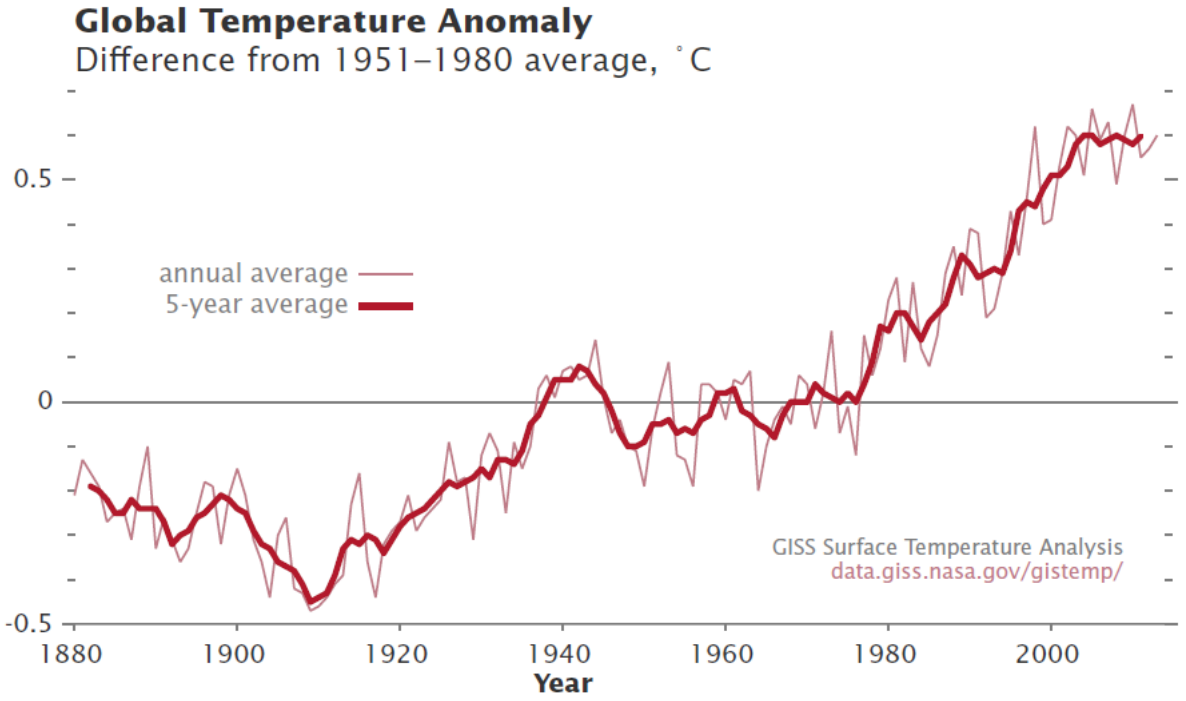

Figure 1.2: 1880-2013 series of global surface temperature anomalies in respect with 1951-1980 average.

human activities on this recent and fast temperature increase has been proved by simulations, since the inclusion of natural forcings alone in the modelling is not enough to explain the observed global temperature trend (IPCC, 2013). The increment in the concentration of GHG_s in the atmosphere mainly driven by human activities such as fossil fuel exploitation and industrial development, is considered the primary responsible of the warming effect. In particular, the global concentrations of CO_2 increased significantly since the beginning of the industrial era passing from about 270 ppm before 1750 to the actual 400 ppm (Figure 1.3). Evidences of the fast changing climatic process can be easily found all around the world for example in the variations in ocean circulation and temperature, as well as in the fast reduction of glaciers and ice-sheets and in the rise of sea level.

Due to the role of climate in determining the environment, the ecosystems, the econom-

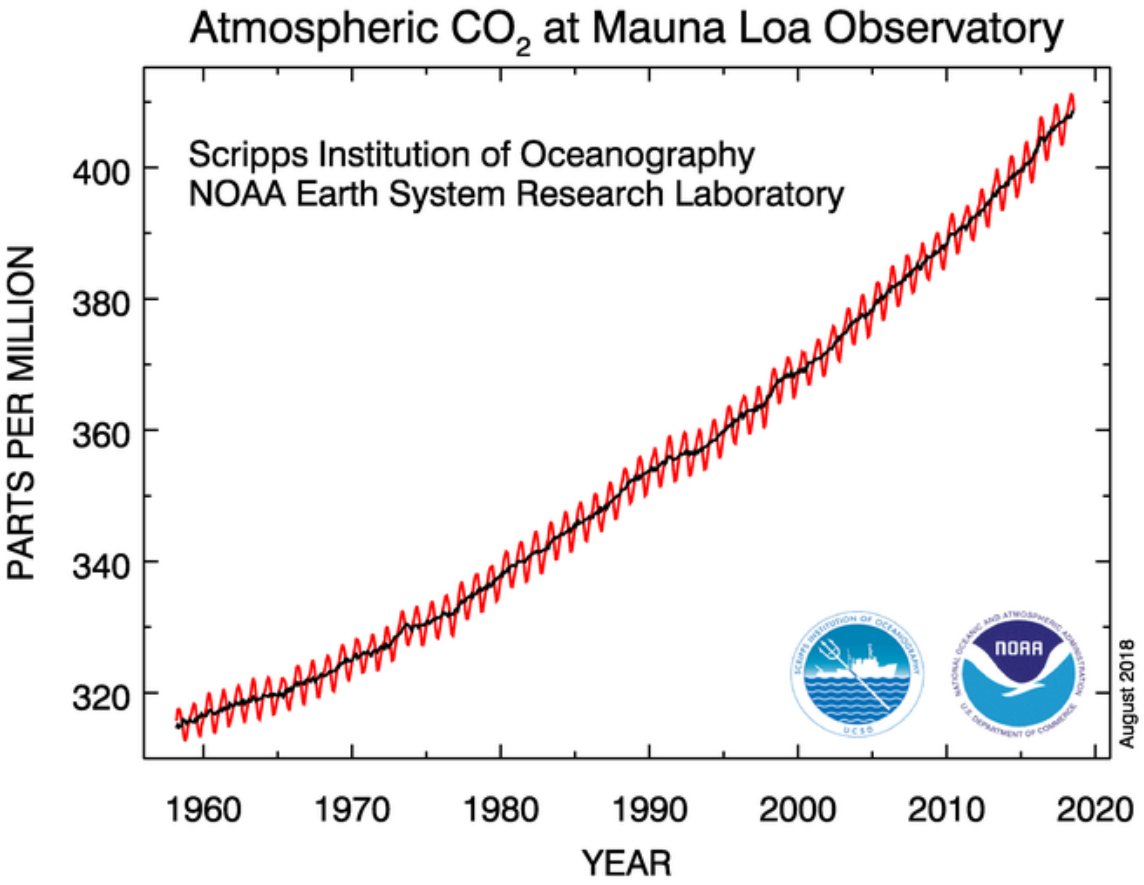

Figure 1.3: The carbon dioxide data (red curve), measured as the mole fraction in dry air, on Mauna Loa constitute the longest record of direct measurements of CO_2 in the atmosphere. The black curve represents the seasonally corrected data. Data are reported as a dry mole fraction defined as the number of molecules of carbon dioxide divided by the number of molecules of dry air multiplied by one million. Source: https://www.esrl.noaa.gov/gmd/ccgg/trends/.

ical activities as well as the human health and development, the evaluation of both the spatial distribution and time evolution of climate variables and their influence on the natural systems is therefore a crucial, and currently relevant, research topic at both global and regional scales. As regards the temporal evolution of climate, the reconstructions are performed for both the remote past and the recent past (from the last two-three centuries to present). Due to the lack of direct meteorological observations, the climate over the past millennia has to be evaluated by proxy data such as pollens (Davis et al., 2003), tree-rings (Frank and Esper, 2005) and the isotopic composition of snow or ice cores (Jouzel et al., 1987) and of speleothemes (Constantin et al., 2007). The climate of the recent past is reconstructed by means of observational databases of instrumental records of the main climatological variables such as temperature, precipitation and solar radiation. For Italy the evolution of the climate over the last centuries

has been recently analysed by Brunetti et al. (2004; 2006) and Manara et al. (2015). The future climate is computed by numerical models based on a complex system of physical equations taking into account all the acting processes, the future role of both natural and human forcings and their interactions. The climate scenarios for the next decades highlight a further strengthen of the global warming, with temperature trends strongly dependent on the adopted Representative Concentration Pathway (RCP), i.e. the expected curve of GHG emissions for the next years (van Vuuren et al., 2011). Beside to the temperature increase, the future climate scenarios give also evidence of a reduction of the water resource availability, as well as of an increase in precipitation extremes both in terms of intensity and duration, so that long dry spells and short periods of intense precipitation are expected to be more likely in the next decades. In addition, the warming will enhance the hydrological cycle which implies higher rates of evaporation and a greater proportion of liquid to solid precipitation; these physical mechanisms, associated with potential changes in precipitation amount and seasonality, could affect soil moisture, groundwater reserves, and the frequency of flood or drought episodes. However the effects of the global climate system variations and their magnitude could be extremely different on regional and local scales in relation with the specificities and vulnerability of the areas, so that local trends could turn out to be double or opposite in comparison with the global climate changes. For this reason, the studies on climate changes need to be managed on different spatial resolution, ranging from global to regional scales, by developing methods able to describe the climate at local level and to capture the heterogeneity of the climate signal. This could be particularly crucial for precipitation, which are highly influenced by the topography, and over the mountainous regions. The IPCC Fifth Assessment Report indicates in fact the mountain environments as highly fragile ecosystems where the rich biodiversity, hydrolgical complexity and human communities are more exposed to the climate changes. The Alps, in particular, have experienced in the last 150 years a double rate in temperature increase in comparison with the global average, which has been recorded at both the low and high-elevation areas of this region (Böhm et al., 2001; Brunetti et al., 2009a). The analysis of proxy records, mainly from tree-rings, showed that this warming has led in the last decades to decadal average summer temperatures that have never been observed in the last millennium (Masson-Delmotte et al., 2013 and references therein). The rising temperature is inducing a fast glacier reduction, the melt of permafrost inside the soil which enhances slope instability and higher evapotranspiration rates. In such a changing environment, the precipitation regime could significantly affect the safety of life and civil infrastructures of mountainous communities. The knowledge and reconstruction of the spatio-temporal variability and distribution of precipitation at fine scale represent therefore fundamental tasks to manage future resources, to preserve and promote the mountain environment and to prevent

water-related hazards, such as floods, landslides and avalanches.

1.2 The reconstruction of the recent climate

High-resolution information on the recent climate evolution (from the last two centuries to present) over areas of interest is a crucial requirement to analyse with spatial continuity the temporal variability of specific meteorological variables and the occurrence of long-term trends at local scale. Gridded datasets of secular series of relevant quantities, e.g. temperature and precipitation, based on station observations are relevant not only for climate-related research but also as decision-support tools in a wide range of fields, such as energy production, agriculture, engineering, civil protection and natural resource preservation. In addition they are useful to retrieve areal information about the target variable or specific related quantities, which would not be directly computed by station records only. Station-based averages could be in fact highly vulnerable to gaps in the observation series and unevenless in spatial data coverage, for example between mountains and valley bottoms, which could influence the actual pattern of the climatic signal. High-resolution grids of precipitation records allow for example to provide total areal precipitation series over a catchement or a specific region and to identify changes and tendencies in water input and disposal.
At the state of the art, one of the most applied schemes to reconstruct the temporal signal of a meteorological variable is the so-called *anomaly method* (e.g. New et al., 2001; Mitchell and Jones, 2005), which is based on the assumption that the temporal variation of any meteorological variable can be computed by superimposing two separate fields: a spatial field of constant values of reference, i.e. the climatologies, and the departures from them, i.e. the anomalies. The two fields can be computed independently, even starting from different databases. In fact climatologies are strongly influenced by the orographic features of the surface and they require a dense station distribution in order to properly capture and reconstruct the spatial meteorological gradients; on the contrary, anomalies are characterised by a larger spatial coherence so that their field can be obtained using a fewer number of observations. In anomaly interpolation, continuous and high-quality records are required in order to extend the reconstruction over long periods and avoid the influence of non-climatic signals which could mask the actual climate evolution. The anomaly-based approach aims at limiting the biases in interpolated fields due to the uneven distribution of stations over both time and space as the climate signal reconstruction over low-sampled areas takes advantage from the large spatial coherence characterising the anomalies and from the high-resolution information supplied by the climatological background.

1.3 High-resolution climatologies

Climatologies report the mean values of a specific meteorological variable over a period of reference at a certain temporal and spatial resolution. Several spatial interpolation techniques have been proposed so far in order to project the climatological information from station sites onto high-resolution grids. The interpolation allows to obtain the climatological data for a number of points of several orders of magnitude higher than the number of available observations. The interpolation methods can be classified into i) deterministic models such as splines, inverse distance weighting and linear regression ii) statistical probability models, including kriging, its variants and Baesyan-based approaches, and iii) expert-based models, such as data-driven techniques and neural network. The choice of the model largely depends on the study variable, the availability of observation data and the target spatial resolution. In particular, precipitation are strongly related to the geographical features of the Earth's surface, and this relationship is known to vary at a very local level (Daly et al., 1994). For this reason, the construction of high-resolution datasets of precipitation climatologies requires the application of interpolation approaches modelling the links between precipitation and orography as well as of rain-gauge networks properly covering the study area in order to capture the fine-scale precipitation gradients.

Interpolation grids are generally based on Digital Elevation Models (DEMs) at specific spatial resolutions. The most recent DEM versions released by U.S. Geological Survey and used in the presented works are available at 30, 15 and 7.5 arc second resolutions (USGS GMTED2010, GTOPO30). However, the surface description provided by the DEMs is generally finer than the actual spatial scales at which the atmospheric circulation is influenced by the topography (Henn et al., 2018), so that a smoothing of elevation details by means of proper filters is generally performed for most spatialisation procedures considering the topography as precipitation predictor.

Over the last decades several precipitation climatologies based on rain-gauge observations have been produced in Europe with different spatial resolutions, on both national and regional scales. 1 km resolution climatologies have been recently released for Finland (Aalto et al., 2017) and Croatia (Perčec Tadić, 2010), while Frei and Schär (1998) provided a 1971-1990 monthly climatological dataset at about 25 km resolution covering the European Alps and, more recently, the daily precipitation climatologies for the same Alpine region have been described at 1 km resolution by Isotta et al. (2014). In the following some of the most widely used approaches for precipitation interpolation are described.

1.3.1 Regression-based models

Regression-based models assume the occurrence of a linear relationship between precipitation, which is considered the dependent variable y, and one or more specific predictors, or independent variables, x. The least squared method is applied to define the straight line which allows to minimise the squared distance from each point. The number of independent variables defines the type of regression, which could be simple if only one variable is used, or multiple, if more predictors are included:

$$y = a + bx + e \tag{1.1}$$

$$y = a + b_1 x_1 + ... + b_n x_n + e \tag{1.2}$$

where a is the intercept, i.e. the value of y when the independent variables are zero, b describes the rate of change (increase or decrease) of y for the unitary increment of each independent variable and e are the residuals. Since precipitation spatial gradients are mostly driven by geographical conditions, some of the main predictors which are generally taken into account in regression-based interpolations are coordinates, elevation, sea distance, slope orientation and slope steepness.
The linear regression could be global or local in relation to the area over which the fit is evaluated. A global regression performs a single linear fit over the whole study area, i.e. the same relationship between the dependent variable and predictors is assumed to exist at all the points of the grid. This approach is preferable for domains featuring homogeneous physiographic conditions and shallow elevation gradients. On the contrary, the projection of precipitation normals onto orographically complex surfaces could require the application of a local regression approach in which different linear relationships are applied over the different portions of the domain by means of moving windows spanning the whole study area. The extent of the windows depends on the grid resolution, geographical conditions and, most of all, station coverage which has to be dense enough to properly represent the local climatic regimes.
One of the most used methods for spatial interpolation of precipitation and based on regression is the PRISM approach (Daly et al., 2002 and references therein) which was recently implemented to study the monthly precipitation distribution over Northern and Central Italy by Brunetti et al. (2009b). In this method, a local weighted linear regression of precipitation versus elevation (LWLR) is applied at each point of the grid. The precipitation-elevation relationship at each cell is defined by means of a weighted linear fit performed on neighbouring stations. In fact, in order to prioritise stations which are mostly representative of the orographic conditions at target location, they enter in the regression with weights depending on their nearness and orographic similarity (e.g. elevation, slope steepness, slope orientation and sea distance) to the grid cell under consideration. The range of the searching radius for stations to be included

in the fit as well as the weighting functions have to be evaluated in relation to the features of the study domain.

1.3.2 Kriging-based models

Kriging is a geostatistical interpolation scheme which was firstly proposed by Daniel Krige in 1951 (Krige, 1951) and developed further by George Matheron (Matheron, 1969). In kriging model, prediction is defined as the combination of a constant stationary function, i.e. a global mean, and a spatially correlated stochastic part of variation. Following the description reported in Hengl (2009), if P is the target variable, $p(\underline{s}_1), p(\underline{s}_2), ..., p(\underline{s}_N)$ is a set of observations and $\underline{s}_i = (x_i, y_i)$ are the geographical locations, the predictions are based on the model:

$$P(\underline{s}) = \mu + \epsilon(\underline{s}) \tag{1.3}$$

where μ is the constant stationary function (global mean) and $\epsilon(\underline{s})$ is the spatially correlated stochastic part of variation.
Therefore, the predictions at an unknown location $\underline{s}_0$ are computed as:

$$\tilde{p}(\underline{s}_0) = \sum_{i=1}^{N} w_i(\underline{s}_0) \cdot p(\underline{s}_i) = \lambda_0^T \cdot \underline{p} \tag{1.4}$$

where λ_0 is the vector of kriging weights and $\underline{p}$ is the vector of N observations at selected locations. The weights are chosen in order to minimise the prediction error variance and they depend on the spatial autocorrelation structure of the variable. The first step consists in fact in the definition of an experimental semivariogram which is a function describing the degree of spatial dependence of the spatial random field and it is defined as the half of the average squared difference between all the data pairs with increasing distance:

$$\gamma(\underline{h}) = \frac{1}{2m(\underline{h})} \cdot \sum_{i=1}^{m(\underline{h})} [p(\underline{s}_i) - p(\underline{s}_i + \underline{h})]^2 \tag{1.5}$$

where $\gamma(\underline{h})$ is the estimated semivariance for the distance $\underline{h}$ and $m(\underline{h})$ is the number of measured point pairs at distance $\underline{h}$. The experimental semivariogram is fitted by means of a theoretical model, e.g. Gaussian, spherical or exponential (Goovaerts, 2000), and the theoretical semivariogram is then used to extract the kriging weights. The weights are extracted by minimising the error variance and by assuring the unbiasedness of the estimator, i.e. the sum of the weights must be equal to unity. The basic assumptions of kriging are the stationarity, unbiasedness and normal distribution of data and the absence of global trends.

Besides the basic kriging scheme above described, also called Ordinary Kriging (OK), several variants are used. Regression (or Residual) kriging (RK) is a kriging variant used to interpolate non stationary data. In this case the value of the study variable at a certain location is represented as the sum of deterministic and stochastic components that can be modelled separately:

$$P(\underline{s}) = m(\underline{s}) + \epsilon^{'}(\underline{s}) + \epsilon^{''} \tag{1.6}$$

and the predictions at an unknown location are computed:

$$\tilde{p}(\underline{s}_0) = \sum_{k=1}^{n} \tilde{\beta}_k \cdot q_k(\underline{s}_0) + \sum_{i=1}^{N} \lambda_i \cdot \epsilon(\underline{s}_i) \tag{1.7}$$

where $\tilde{\beta}_k$ are the regression coefficients, $\underline{q}_k$ are the regression predictors, n is the number of predictors, $\underline{\epsilon}$ are the residuals of regression at sample locations and $\underline{\lambda}$ are the kriging weights obtained from the spatial dependence structure of the residuals. The regression coefficients $\tilde{\beta}_k$ can be evaluated by ordinary least squared methods and, once the deterministic part is computed, the residuals are interpolated by OK and added to the estimated trend at target location. RK is useful whenever the target variable is expected to feature a specific relationship with some selected predictors, as it is the case for precipitation with the geographical conditions of the domain surface. As for the regression-based models described in the previous section, the regression coefficients of RK can be evaluated on both global and local scales.

1.4 Anomaly interpolation and long-term series reconstruction

Anomalies are defined as the deviations of a meteorological series from the corresponding climatological values of reference. While temperature anomalies are expressed as the difference from the normals, for precipitation relative anomalies are considered, i.e. the ratios to the corresponding climatologies. The spatial structure of anomalies shows a larger spatial coherence than the absolute values so that reliable interpolated estimates can be retrieved even from a lower number of stations. However continuous and high-quality records are required in order to extend the reconstruction over long periods and to avoid the influence of non-climatic signals which could mask the actual climate evolution. In particular, the homogeneity of the observed records is a crucial requirement in order to assure the absence of *breaks* in the series due to external factors, e.g. station relocations, malfunctions or replacement of sensors (see section 1.6).

Most of the approaches proposed in scientific literature to interpolate station anomalies onto a high-resolution grid are based on the weighting average of surrounding observations (e.g. Efthymiadis et al., 2006; Isotta et al., 2014). In particular the weight of the $i-$th station at each target cell (x, y) is computed as a product of Gaussian functions, one for each factor which is expected to mostly influence the precipitation distribution over the study domain, such as the distance or the elevation difference from the point:

$$w_i^{par}(x,y) = e^{-\frac{\Delta_i^{par}(x,y)}{c^{par}}} \tag{1.8}$$

where *par* represents the considered factor, $\Delta_i^{par}(x,y)$ is the difference in the factor between the considered point and the station i and c^{par} is the rate of the weight decay. The decreasing rate of the weighting functions could be dependent on both location and time step.
The high-resolution field of anomalies can be finally converted into absolute values by the superimposition of climatologies (sum or product in relation to the variable under consideration) in order to provide a long-term series for each point of the regular grid covering the study area.

1.5 Validation of interpolation methods

The validation of the interpolation methods applied to project station information onto high-resolution grids allows to quantify the accuracy and reliability of the computed fields. The validation is performed by the cross-validation of station data and by the evaluation of reconstruction errors. One of the most applied cross-validation methods is based on the Leave-One-Out procedure (LOO) in which each station in turn is discarded from the database, its value is computed by applying the interpolation method to the remaining stations and then compared to the measured one. The LOO validation allows to avoid the self-influence of the station data under reconstruction and to increase the robustness of the validation estimates. In particular, the agreement between observed values (O) and the model predictions (P) can be quantified by means of several error estimators. The most common error estimators are:

- BIAS, it is the mean difference between modelled and measured values and allows to detect systematic errors in model estimates.

$$BIAS = \frac{1}{N}\sum_{i=1}^{N}(P_i - O_i) \tag{1.9}$$

- Mean Absolute Error (MAE), it provides the mean value of the absolute differ-

ence between modelled and observed data.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|P_i - O_i| \tag{1.10}$$

- Mean Absolute Percentage Error (MAPE), it provides the percentage of the mean absolute differences between modelled and measured data normalised to the observations.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|P_i - O_i|}{O_i}\cdot 100\% \tag{1.11}$$

- Root Mean Squared Error (RMSE), it emphasises the influence of high differences between predictions and observations, since the errors are squared before they are averaged. For this reason RMSE could be useful to highlight the cases showing the most problematic reconstructions.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2} \tag{1.12}$$

- Determination coefficient or R-squared (R^2), it explains the proportion of the variance in the observations that is predictable by the model:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}(O_i - \overline{O})^2} \tag{1.13}$$

1.6 Data preparation: quality control and homogenisation

In order to assure the quality of observation data from meteorological stations entering into climatic studies and temporal reconstruction, preliminary processes on the database are required. The quality-check procedures applied to station records allow in fact to detect and remove the most spurious entries, such as outliers or, in case of precipitation, sequences of null values, due to sensor malfunction or digitisation oversights, which could lead to relevant artefacts in the interpolated fields. Erroneous station metadata (i.e. coordinates or elevation) have to be identified too and, whenever it is possible, corrected. One of the most used procedures for quality-check is based on station anomalies and consists in comparing each measured series with a simulated one by means of the neighbouring records. More specifically, in case of monthly precipitation series, for each datum of each measured series (p_{test}) the n 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 are selected. The p_{test} for the month m is then reconstructed from each reference value p_{ref} as follows:

$$\tilde{p}_{test_m,i} = p_{ref_m,i} \cdot \frac{\bar{p}_{test_m,i}}{\bar{p}_{ref_m,i}} \qquad i = (1, \ldots, n) \tag{1.14}$$

where $\bar{p}_{test_m,i}$ and $\bar{p}_{ref_m,i}$ are the test 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 by considering the median of the n estimated values. The comparisons between observed and estimated series allows to detect the records showing a low agreement with the neighbouring stations and to point out the most relevant discrepancies (Figure 1.4).

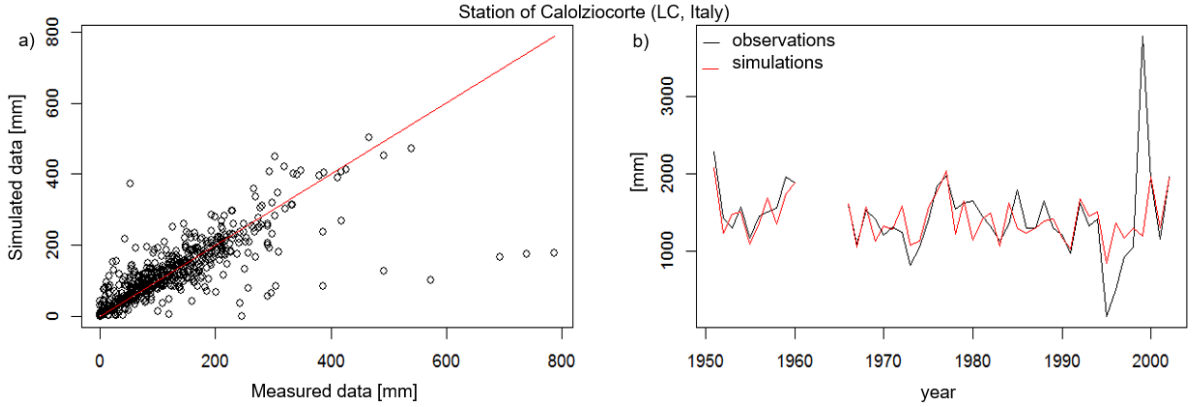

Figure 1.4: Comparison between the observed precipitation record and the simulated values by means of equation 1.14 at the Calolziocorte station (LC, Italy). Panel a) reports the scatter plot between simulated and observed monthly values, in panel b) the modelled (red line) and measured (black line) annual precipitation series are compared.

The quality-check procedure allows also to detect the most evident inhomogeneities which could be contained in a meteorological series, i.e. changes in the climatic regime caused by external factors, such as changes in station locations or in the surrounding environment. The identification and correction of breaks is particularly crucial when the records are used for temporal recostruction and trend analyses since they could introduce non-climatic signals altering the actual evolution of the target variable. Station metadata reporting details about the site history help to identify the periods of possible *breaks* in the station record, even though in most cases these information are not easily available. One of the best performing methods to control series homogeneity (Venema et al., 2012) and to detect *breaks* is the *Craddock test* (Craddock, 1979), which is based on the assumption that the difference/ratio between two homogeneous records should be constant along the time. If monthly precipitation data are considered, for each datum

of the test series p_{test} the ratio with the corresponding entry of each one of n selected reference stations is computed as follows:

$$R_m = \frac{\sum_j \frac{p_{test,m,j}}{N}}{\sum_j \frac{p_{ref,m,j}}{N}} \tag{1.15}$$

where m is the month under consideration, j represents the common years of valid data and N is the total number of common years. 12 values for R are thus obtained. The monthly series of cumulated differences C (*Craddock series*) is computed between the test series and the normalised values of each reference one by means of the corresponding R_m estimates. The j-th element of the Craddock series is:

$$C_j = \sum_{m=1}^{j} (p_{test,m} - R_m \cdot p_{ref,m}) \tag{1.16}$$

It is possible to verify that the last element of the Craddock series is null. The breakpoints in the test series could be identified by simultaneous changes (first-derivative discontinuities) in the slope of all the Craddock series obtained from the comparison with the n reference stations (Figure 1.5).

The time series containing breaks can be subjected to a procedure of data homogenisation by following the same assumption of constant ratio among homogeneous records.

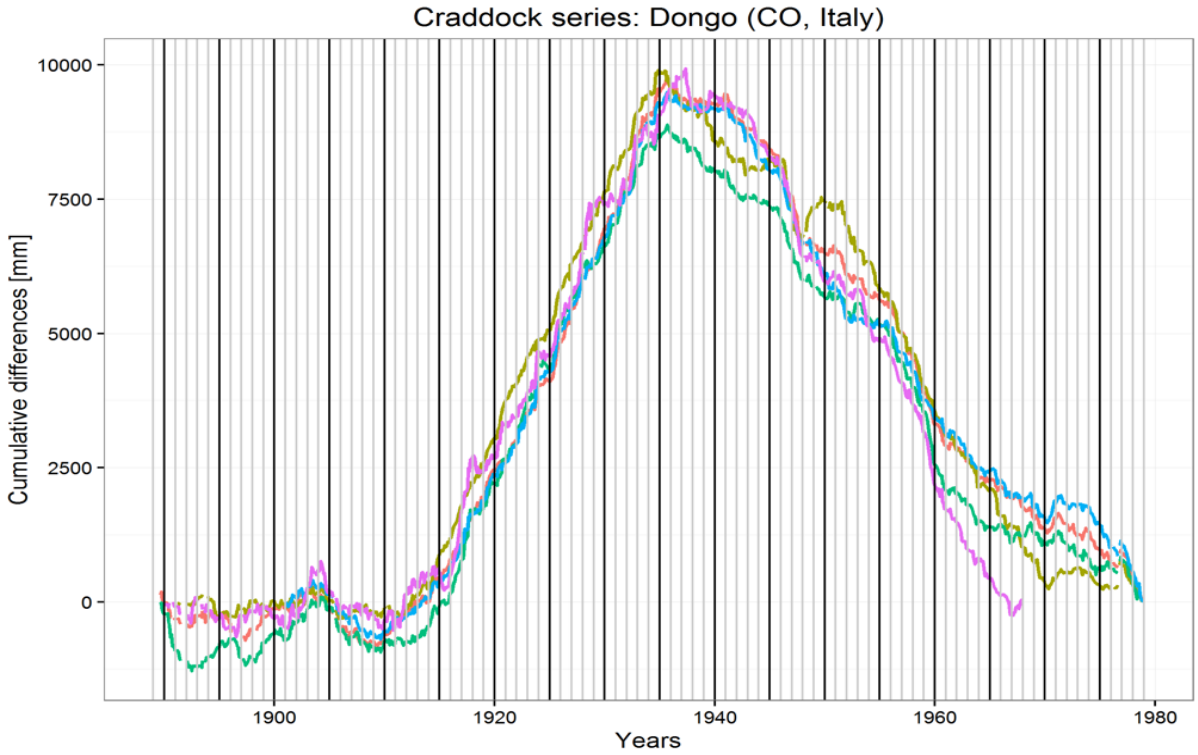

Figure 1.5: Craddock series for the monthly precipitation record of Dongo station (CO, Italy) obtained from five different reference stations.

By considering a homogeneous reference series on the whole period and the inhomogeneous test series on the subperiod A, the following relation can be considered:

$$\frac{\overline{ref}_{A,m}}{\overline{ref}_{B,m}} = \frac{\overline{test}_{A,hom,m}}{\overline{test}_{B,m}} = \frac{C_m \cdot \overline{test}_{A,inhom,m}}{\overline{test}_{B,m}} \tag{1.17}$$

where

- B refers to the subperiod in which both series are homogeneous;
- m is the month under consideration;
- $\overline{test}_{A,inhom,m}$ is the mean precipitation for the month m of test series over the inhomogenous subperiod$A$;
- $\overline{test}_{A,hom,m}$ is the expected mean of test series for the month m in order to obtain a constant ratio over the whole period;
- C_m is the correcting factor for the month m.

In order to increase the robustness of the correction, C_m is generally computed from several reference stations and the mean value is finally applied. In some cases a different correcting factor for each month should be applied, in others a single value for all the months is preferable by considering for example the average of the monthly corrections normalised for the corresponding monthly climatologies. However, at the state of the art, a standard approach to identify and correct breaks in meteorological data has not been yet provided in scientific literature and the homogenisation procedure is largely based on expert-based choices.

1.6.1 Trend analysis of meteorological series

The availability of long and continuous series of meteorological variables for a specific region allows to investigate the occurrence of long-term variability and trends. Trends in a time series, i.e. increasing or decreasing tendency in the time evolution of the values assumed by the variable, are generally assessed by linear regression where time is the independent variable and the studied meteorological quantity is the dependent variable. However, the outcomes of the ordinary least squared method could be strongly affected by the presence of outliers masking the actual correlation between the variables while the efficiency of the test could be strongly limited by the heteroskedasticity of data, i.e. different variances for different subpopulations (Wilcox, 2010). In order to overcome these limitations, the Theil-Sen non-parametric estimator was proposed (Theil, 1950; Sen, 1968), which computes the median of the slopes obtained from

all the pairs of sample points without requiring any assumption about data distribution. This approach allows to deal better with outliers and asimmetry of data in comparison with the simple linear regression. A confidence interval for the slope of Theil-Sen regression can be computed as the interval containing a certain fraction of the slopes obtained from all the data pairs.
In order to evaluate the statistical significance of the trends obtained for a given time series several methods have been proposed. One of the most applied methods is the Mann-Kendall test, which is a non-parametric test requiring no assumption about the probability distribution of the variable of interest (Kendall, 1970). The null condition (H_0) assumes that no trend exists and data are independent and randomly distributed. Each element of the ordered time series T is compared with all the following ones to compute the Mann-Kendall statistics S:

$$S = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} sgn(T_j - T_i) \tag{1.18}$$

where $sgn()$ is the sign function and N is the number of entries of the series. By the null condition H_0 and for $N > 10$ it is possible to verify that S assumes a Gaussian distribution with null average and variance:

$$\sigma^2 = \frac{N(N-1)(2N+5)}{18} \tag{1.19}$$

The normalised statistics in Mann-Kendall test is then computed as follows:

$$Z = \begin{cases} \frac{S-1}{\sigma} & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sigma} & S < 0 \end{cases} \tag{1.20}$$

which is asimptotically distributed as a normal standard distribution. The relevance of trend can be assessed by setting a certain significance level. If H_0 is considered valid, the probability to obtain a value of S different from zero at least as much as the computed one is:

$$\alpha = 1 - \frac{1}{\sqrt{2\pi}} \cdot \int_{-Z}^{Z} e^{-\frac{t^2}{2}} dt \tag{1.21}$$

The significance level is the value of α below which H_0 is rejected. A common choice for α is 0.05 corresponding to about 2σ.

1.7 State of the art and aims of the PhD project

My PhD activities were undertaken within the research group composed by the Department of Environmental Science and Policy of Università degli Studi di Milano and the Institute of Atmospheric Sciences and Climate of the National Research Council (ISAC-CNR). The group focuses on the reconstruction of the Italian climate and on the study of its variability by recovering the available long series of Italian meteorological data, in particular temperature, precipitation and solar radiation, and by implementing new interpolation schemes and analysis procedures.
Italy is characterised by an exceptional availability of long meteorological series. Regular meteorological observations started during the 18^{th} century in Bologna, Padua, Milan, Turin and Palermo. A relevant contribution to the systematic collection of meteorological data was provided since 1913 by the National Hydrographic Service, which managed more than 4000 stations after the I World War. This relevant heritage of observations is today available but it has been only partially transferred to a digitised form which could be used as input in computer analyses. Moreover, after 1980 the Hydrographic Service was closed so that its competences were transferred to the regional services together with the historical records in hardcopy archives. The regionalisation of the national network together with the concurrent transition from mechanical to automatic stations caused the fragmentation of data sources and discontinuities in observation series due to dismission or relocation of a relevant number of monitoring sites. The recovering activities of the research group led to gather most of the secular Italian meteorological series with the digitisation of data from hardcopy archives for the past together with station metadata and to merge the new automatic records for the most recent decades in order to overcome the discontinuities in data coverage characterising the last decades of the 20^{th} century. One of the results of these activities was the set up of a dense database of temperature series covering Italy and neighbouring countries which was used to reconstruct the national monthly temperature climatologies for the period 1961-1990, which corresponds to the 30-year interval of best data availability for Italy (Brunetti et al., 2014).
In this framework, my PhD project aims at improving the study of precipitation in Italy with the set up of a dense database of precipitation series for the country and the implementation of new methodologies to analyse both the climatological distribution and the long-term variability of this meteorological variable also in relationship with other hydrological quantities, especially catchement runoff. In particular, the work was focused on the construction of a national precipitation climatology of reference for Italy and on the evaluation of the precipitation behaviour over mountainous regions where the analyses are generally affected by a relevant reduction of available in-situ observations and by the possible underestimations of solid precipitation collected dur-

ing the winter season by rain-gauges at the highest elevations (Sevruk et al., 2009).
At these aims, a national database of monthly precipitation records was constructed to be used for the climatology reconstruction by exploiting a relevant number of Italian data sources, while a database covering the Northern Italy and including the longest available records was set up in order to investigate the long-term evolution of the climate signal over specific Alpine areas.
The largest part of the precipitation records used to set up the two databases was retrieved from the former Hydrographic Service and from ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale). In addition, some of the most ancient data were digitised from the available hardcopy yearbooks of Hydrographic Service and partly retrieved from previous digitisation projects. The Italian regional and subregional services which are currently in charge of the management of the regional meteorological networks are the data providers of the most recent records, especially of those collected by the new automatic stations. In some cases, the merging between the automatic and mechanical records was evaluated by an accurate analysis of metadata and potential breaks in the series in order to extend their temporal coverage to present.
All the series were subjected to quality-check procedures and the longest series entering in trend analyses were controlled for homogeneity.
From the quality-checked national database the 1961-1990 monthly precipitation normals for all the available station sites were computed and entered into the interpolation procedures to provide the high-resolution climatologies for Italy. The obtained gridded dataset is useful to investigate the spatial distribution of precipitation over Italy, to highlight the main Italian climatic zones and to validate numerical model output, radar measurements or the results of hydrological models.
The Northern Italian database was used to investigate the long-term trend of precipitation over specific mountainous areas after projecting the station series onto a high-resolution grid by means of the anomaly-based approach and by computing the corresponding regional averaged series. This approach was found to be particularly useful to reconstruct the spatio-temporal precipitation fields of mountainous areas and to deal with the low data coverage affecting the remote regions.
In order to evaluate the most suitable approaches to provide reliable climatological fields for regions affected by sparse station networks, a six-month period of my PhD course was spent by the Norwegian Meteorological Institute of Oslo (Norway). The Norwegian meteorological network is in fact characterised by a relevant reduction of in-situ observations over high-elevated areas and in the northernmost part of the country, preventing from getting accurate climatological estimates over the whole territory by means of conventional interpolation schemes. The combination of numerical model outputs with observations allowed to provide a reference high-resolution dataset of Norwegian monthly precipitation climatologies which is available for hy-

drological comparisons as well as to provide information about the distribution of water disposals, which is particularly crucial for the hydropower production sector of the country.

1.8 PhD thesis organisation

The thesis is organised as a collection of published or submitted papers describing in detail the scientific contributions of my PhD research activities. Each chapter focuses on a specific issue and it is composed by one or more published or submitted articles presenting the problem and the obtained results. More specifically, the thesis is structured as follows:

- Chapter 2: The 1961-1990 high-resolution monthly precipitation climatologies over Italy are presented. The observation database and the interpolation methods used to project the station normals onto a 30-arc second resolution grid are fully described. In particular two reference versions of the Italian climatologies are obtained by means of two different interpolation schemes: a Local Weighted Linear Regression of precipitation *versus* elevation and a Local Regression Kriging. The accuracy of the methods is discussed and an inter-comparison with other widely used approaches is performed. The article is published in open-access version on the *International Journal of Climatology* (Crespi et al., 2018) and the dataset of gridded monthly climatologies is freely available at http://www.isac.cnr.it/climstor/CLIMATE_DATA/.

- Chapter 3: The reconstruction of a gridded dataset of monthly precipitation normals and of 1913-2015 monthly precipitation records over a high-mountain area centred on the Forni Valley (Central Italian Alps) is described. The ability to provide high-resolution information over an orographically complex region starting from a sparse rain-gauge network is discussed. In particular, the model estimates over the Forni Glacier are compared with the in-situ observations available from the Automatic Weather Station (AWS) operating on the glacier tongue and not used as input in the interpolation procedure. The gridded dataset is also used to provide the 1913-2015 areal precipitation series of the study region which is investigated for variability and trends. This application aims at providing a suitable methodology to study the climate evolution over vulnerable high-mountain regions where direct in-situ observations are often sparse and mainly located at the lowest elevations. The presented work is published in open-access version on *Advances in Meteorology* (Golzio et al., 2018).

- Chapter 4: The high-resolution dataset of secular monthly precipitation records for the upper part of Adda river basin (Central Italian Alps) is presented. The

dataset is obtained by applying an anomaly-based approach to a dense and homogenised archive of monthly precipitation series covering Northern Italy and spanning more than two centuries. The dataset is used to retrieve the 1800-2016 monthly areal precipitation series over the basin whose temporal evolution is investigated by trend analysis. The influence on the reconstruction accuracy of the variation in data coverage occurred over the study period is estensively discussed and the evolution of the basin precipitation series is then compared with that of the corresponding 1845-2016 series of monthly runoff in order to detect possible changes in the hydrological cycle. The dataset is also useful to reconstruct and analyse the spatial structure of past intense precipitation events occurred over the region and some case-studies are reported. The results of this activity are published in open-access version on *Advances in Science and Research* (Crespi et al., 2018) and submitted to *Water Resources Research* (Crespi et al., 2019).

- Chapter 5: The spatio-temporal reconstruction of monthly precipitation over the period 1922-2011 for Sardinia (Italy) is presented. The work aims at evaluating the long-term evolution and possible trends of average precipitation regime over one of the main Mediterranean islands. A database including more than 350 quality-checked and homogenised monthly precipitation records spanning more than one century was set up by recovering the most ancient data from non-digitised yearbooks and the recent observations from the automatic meteorological network. The 1961-1990 monthly precipitation climatologies are computed and the anomaly method is applied to project the 1922-2011 station records onto the 30-arc second resolution grid. The average annual and seasonal total precipitation series for Sardinia are presented and long-term trends are investigated. Further analyses are planned in the next future and the work will be object of a forthcoming paper.

- Chapter 6: The 1981-2010 monthly precipitation climatologies over Norway at 1 km resolution are described. The climatologies are obtained by a combined interpolation approach joining in-situ observations with the precipitation fields provided by the climate model version of HARMONIE NWP model, which was run over the country at 2.5 km resolution. The improvements in reconstruction accuracy in respect with conventional interpolation schemes are discussed. The application aims at providing a methodology to improve the climatological studies over countries covered by uneven meteorological station networks. The presented work has been submitted to *International Journal of Climatology* (Crespi et al., 2019).

All the presented works are the results of my original research and I contributed as first author in all the published and submitted articles, except for the study discussed

in chapter 3 which was managed in collaboration with another PhD student and in which I contributed with the data interpolation and trend analysis.

1.9 References

- Aalto, J., Pirinen, P., Heikkinen, J., and Venäläinen, A. (2013). Spatial interpolation of monthly climate data for Finland: comparing the performance of kriging and generalized additive models. Theoretical and Applied Climatology, 112, 99-111. doi:10.1007/s00704-012-0716-9

- Beniston, M., and Jungo, P. (2002). Shifts in the distributions of pressure, temperature and moisture and changes in the typical weather patterns in the alpine region in response to the behaviour of the North Atlantic Oscillation. Theoretical and Applied Climatology, 71, 29-42. doi:10.1007/s704-002-8206-7

- Berger, A. (1988). Milankovitch Theory and climate. Reviewes of Geophysics, 26, 624–657. doi:10.1029/RG026i004p00624

- Böhm, R, Auer, I., Brunetti, M., Maugeri, M., Nanni, T., and Schöner, W. (2001). Regional temperature variability in the European Alps 1760–1998 from homogenized instrumental time series. International Journal of Climatology, 21, 1779–1801. doi:10.1002/joc.689

- Brunetti, M., Buffoni, L., Mangianti, F., Maugeri, M., and Nanni, T. (2004). Temperature, precipitation and extreme events during the last century in Italy. Global and Planetary Change, 40, 141-149. doi:10.1016/S0921-8181(03)00104-8

- Brunetti, M., Maugeri, M., Monti, F., and Nanni, T. (2006). Temperature and precipitation variability in Italy in the last two centuries from homogenized instrumental time series. International Journal of Climatology, 26, 345-381. doi:10.1002/joc.1251

- Brunetti, M., Lentini, G., Maugeri, M., Nanni, T., Auer, I., Böhm, R., and Schöner, W. (2009a). Climate variability and change in the Greater Alpine Region over the last two centuries based on multi-variable analysis. International Journal of Climatology, 29, 2197-2225. doi:10.1002/joc.1857

- Brunetti, M., Lentini, G., Maugeri, M., Nanni, T., Simolo, C., and Spinoni, J. (2009b). 1961–1990 high-resolution Northern and Central Italy monthly precipitation climatologies, Advances in Science and Research, 3, 73-78. doi:10.5194/asr-3-73-2009

- Brunetti, M., Maugeri, M., Nanni, T., Simolo, C., and Spinoni, J. (2014). High-resolution temperature climatology for Italy: interpolation method intercomparison. International Journal of Climatology, 34, 1278-1296. doi:10.1002/joc.3764

- Constantin, S., Bojar, A., -V., Lauritzen, S., -E., Lundberg, J. (2007). Holocene and Late Pleistocene climate in the sub-Mediterranean continental environment: A speleothem record from Poleva Cave (Southern Carpathians, Romania). Palaeogeography, Palaeoclimatology, Palaeoecology, 243, 322-338. doi:10.1016/j.palaeo.2006.08.001

- Craddock, J. (1979). Methods of comparing annual rainfall records for climatic purposes, Weather, 34, 332-346. doi:10.1002/j.1477-8696.1979.tb03465.x

- Crespi, A., Brunetti, M., Lentini, G., and Maugeri, M. (2018). 1961-1990 high-resolution monthly precipitation climatologies for Italy. International Journal of Climatology, 38, 878–895. doi:10.1002/joc.5217

- Crespi, A., Brunetti, M., Maugeri, M., Ranzi, R., and Tomirotti, M. (2018). 1845–2016 gridded dataset of monthly precipitation over the upper Adda river basin: a comparison with runoff series, Advances in Science and Research, 15, 173-181. doi:10.5194/asr-15-173-2018

- Crespi, A., Lussana, C., Brunetti, M., Dobler, A., Maugeri, M., Tveito, O.E. (2019). High-resolution monthly precipitation climatologies over Norway (1981–2010): joining numerical model datasets and in-situ observations. International Journal of Climatology (submitted).

- Crespi, A., Brunetti, M., Michailidi, E. M., Ranzi, R., Tomirotti, M., and Maugeri, M. (2019). A multi-century gridded dataset of monthly precipitation records for the Adda river basin (Central Alps) based on in-situ observations. Water Resources Research (submitted).

- Crowley, T.J. (2000). Causes of the climate change over the past 1000 years. Science, 289, 270-277. doi:10.1126/science.289.5477.270

- Davis, B. A. S., Brewer, S., Stevenson, A. C., and Guiot, J. (2003). The temperature of Europe during the Holocene reconstructed from pollen data. Quaternary Science Reviews, 22, 1701-1716. doi:10.1016/S0277-3791(03)00173-2

- Daly, C., Neilson, R. P., and Phillips, D. L. (1994). A statistical-topographic model for mapping climatological precipitation over Mountainous Terrain. Journal of Applied Meteorology, 33, 140-158. doi:10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2

- Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P. (2002). A knowledge based approach to the statistical mapping of climate. Climate Research, 22, 99–113. doi:10.3354/cr022099

- Efthymiadis, D., Jones, P. D., Briffa, K.R ., Auer, I., Böhm, R., Schöner, W., Frei, C., and Schmidli, J. (2006). Construction of a 10-min-gridded precipitation data set for the Greater Alpine Region for 1800–2003, Journal of Geophysical Research, 111, D01105. doi:10.1029/2005JD006120

- Frank, D., and Esper, J. (2005). Temperature reconstructions and comparisons with instrumental data from a tree-ring network for the European Alps. International Journal of Climatology, 25, 1437-1454. doi:10.1002/joc.1210

- Frei, C., and Schär, C. (1998). A precipitation climatology of the Alps from high-resolution rain-gauge observations. International Journal of Climatology, 18, 873-900. doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9

- Golzio, A., Crespi, A., Bollati, I. M., Senese, A., Diolaiuti, G. A., Pelfini, M., and Maugeri, M. (2018). High-Resolution Monthly Precipitation Fields (1913–2015) over a Complex Mountain Area Centred on the Forni Valley (Central Italian Alps). Advances in Meteorology, 2018, ID 9123814, pp. 17. doi:10.1155/2018/9123814

- Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology, 228, 113–129.doi:10.1016/S0022-1694(00)00144-X

- 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.

- Henn, B., Newman A. J., Livneh, B., Daly, C., and Lundquist, J. D. (2018). An assessment of differences in gridded precipitation datasets in complex terrain. Journal of Hydrology, 556, 1205–1219. doi:10.1016/j.jhydrol.2017.03.008

- Hurrell, J. W., and van Loon, H. (1997). Decadal variations in climate associated with the North Atlantic Oscillation. Climatic Change, 36, 301-326. doi:10.1023/A:1005314315270

- IPCC (2013): Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press.

- Isotta, F.A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and 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, International Journal of Climatology, 34, 1657-1675. doi:10.1002/joc.3794

- Jouzel, J., Lorius, C., Petit, J. R., Genthon, C., Barkov, N. I., Kotlyakov, V. M., and Petrov, V. M. (1987). Vostok ice core: a continuous isotope temperature record over the last climatic cycle (160,000 years). Nature, 329, 403–408. doi:10.1038/329403a0

- Kendall, M. G. (1975). Rank Correlation Methods. Griffin, London.

- Krige, D. G. (1951). A Statistical Approaches to Some Basic Mine Valuation Problems on the Witwatersrand. Journal of the Chemical, Metallurgical and Mining Society of South Africa, 52, 119-139.

- Kvalevåg M. M., and Myhre, G. (2007). Human impact on direct and diffuse solar radiation during the industrial era, Journal of Climate, 20, 4874-4883. doi:10.1175/JCLI4277.1

- Manara, V., Beltrano, M. C., Brunetti, M., Maugeri, M., Sanchez-Lorenzo, A., Simolo, C., and Sorrenti, S. (2015). Sunshine duration variability and trends in Italy from homogenized instrumental time series (1936-2013). Journal of Geophysical Research, 1, 3622-3641. doi:10.1002/2014JD022560.

- Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., González Rouco, J. F., ... and Osborn, T. Information from Paleoclimate Archives. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, 2013.

- Matheron, G. (1969). Le krigeage universel (Universal kriging) Vol. 1. Cahiers du Centre de Morphologie Mathematique, Ecole des Mines de Paris, Fontainebleau, 83. doi:10.1016/j.geoderma.2003.08.018

- Mitchell, T. D., and Jones, P. D. (2005). An improved method of constructing a database of monthly climate observations and associated highresolution grids. International Journal of Climatology, 25, 693–712. doi:10.1002/joc.1181

- New, M., Todd, M., Hulme, M., and Jones, P. (2001). Precipitation measurements and trends in the twentieth century. International Journal of Climatology, 21, 1899–1922. doi:10.1002/joc.680

- Perčec Tadić, M. (2010). Gridded Croatian climatology for 1961–1990. Theoretical and Applied Climatology, 102, 87-103. doi:10.1007/s00704-009-0237-3

- Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall's tau. Journal of the American Statistical Association, 63, 1379-1389.

- Sevruk, B., Ondrás, M., and Chvíla, B. (2009). The WMO precipitation measurement intercomparisons. Atmospheric Research, 92. 376-380. doi:10.1016/j.atmosres.2009.01.016

- Stenchikov, G. L., Georgiy, L., Kirchner, I., Robock, A., Graf, H. F., Antuna, J. C., Grainger, R. G., Lambert, A., and Thomason, L. (1998). Radiative forcing from the 1991 Mount Pinatubo volcanic eruption. Journal of Geophysical Research. 103(D12), 13837-13857. doi:10.1029/98JD00693

- Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. I. Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Series A, 53, 386–392.

- Twomey, S. A., Piepgrass, M., and Wolfe, T. L. (1984). An assessment of the impact of pollution on global cloud albedo, Tellus B, 36, 356-366. doi:10.1111/j.1600-0889.1984.tb00254.x

- USGS (United States Geological Survey). website: http://eros.usgs.gov/

- van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J., -F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K. (2011). The representative concentration pathways: an overview. Climatic Change, 109: 5. doi:10.1007/s10584-011-0148-z

- Venema, V. K. C., Mestre, O., Aguilar, E., Auer, I., Guijarro, J. A., Domonkos, P., Vertacnik, G., ... and Brandsma, T. (2012). Benchmarking homogenization algorithms for monthly data, Climate of the Past, 8, 89-115. doi:10.5194/cp-8-89-2012

- Wilcox, R. R. (2010). Fundamentals of modern statistical methods. 2nd Edition. Springer, New York.

Chapter 2

1961-1990 Italian monthly precipitation climatologies

In this chapter the reconstruction of the Italian monthly precipitation climatologies for the reference period 1961-1990 is presented. More than 5000 series of monthly precipitation covering the whole national territory and surrounding countries were retrieved, together with their metadata, and all the series were checked for quality. The 1961-1990 monthly station normals were interpolated onto a smoothed version of a 30-arc second resolution Digital Elevation Model by means of a Local Weighted Linear Regression and a Local Regression Kriging modelling the link between precipitation and orography. The local approaches considering elevation as the main precipitation predictor were proved to deal better than other tested methods with the complex spatial precipitation gradients occurring on the rough Italian surface.
The obtained climatologies allowed to identify the locations of Italian precipitation extremes and to characterise the main climatic zones.
The resulting dataset represents an improved version in both resolution and accuracy of previous national precipitation climatologies available in digital form and it takes advantage of a very dense observation database evenly covering the whole country. The gridded monthly dataset is now available at the web-site of ISAC-CNR (http://www.isac.cnr.it/climstor/CLIMATE_DATA/). The chapter is based on the peer-reviewed article published on *International Journal of Climatology*:

Crespi, A., Brunetti, M., Lentini, G., and Maugeri, M. (2018). 1961-1990 high-resolution monthly precipitation climatologies for Italy. *International Journal of Climatology*, 38, 878–895. doi:10.1002/joc.5217

The study was also presented during two international scientific conferences and it was published on a national review in Italian language in order to disseminate the contribution of the research on Italian climate to a larger public:

- Crespi, A., Brunetti, M., and Maugeri, M. (2106). 1961-1990 high-resolution monthly precipitation climatologies for Italy. 16^{th} European Meteorological Society (EMS) Annual Meeting and 11^{th} European Conference on Applied Climatology (ECAC), Trieste (Italy), Vol. 13, EMS2016-88.

- Crespi, A., Brunetti, M., and Maugeri, M. (2106). 1961-1990 high-resolution monthly precipitation climatologies for Italy. Mediterranean Climate Variability and Predictability (Medclivar) conference 2016, Athens (Greece), ID: 2016/2-027.

- Brunetti, M., Crespi, A., Lentini, G., and Maugeri, M.(2017). Nuove carte termometriche e pluviometriche per l'Italia. Nimbus, vol. 78, pp.5.

INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2017)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.5217

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 km^2 of

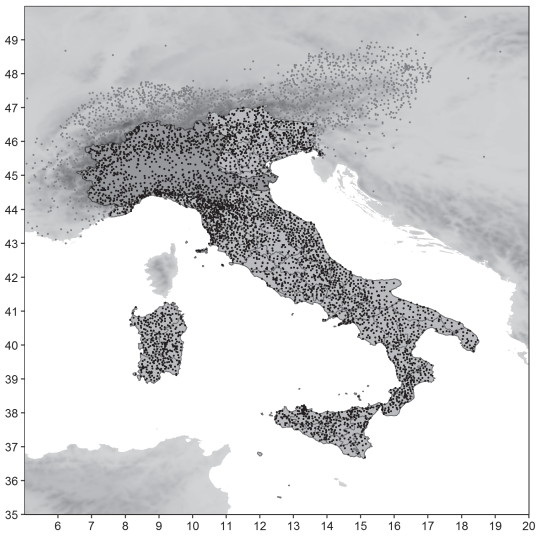

Figure 1. Study domain (light and dark grey bounded areas) and spatial distribution of the 6134 stations in the final version of the dataset. The dark grey portion corresponds to the Po basin and the black dots represent the 4525 stations included within the study domain. The sites outside this area (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:

$$p(x,y) = a(x,y) + b(x,y) \cdot h(x,y) \tag{2}$$

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 ith station involved in the linear regression for the grid cell (x,y) is the product of several weighting factors:

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

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:

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

where par is the geographical parameter that is being considered, Δ_i^{par} is the difference between the values of the

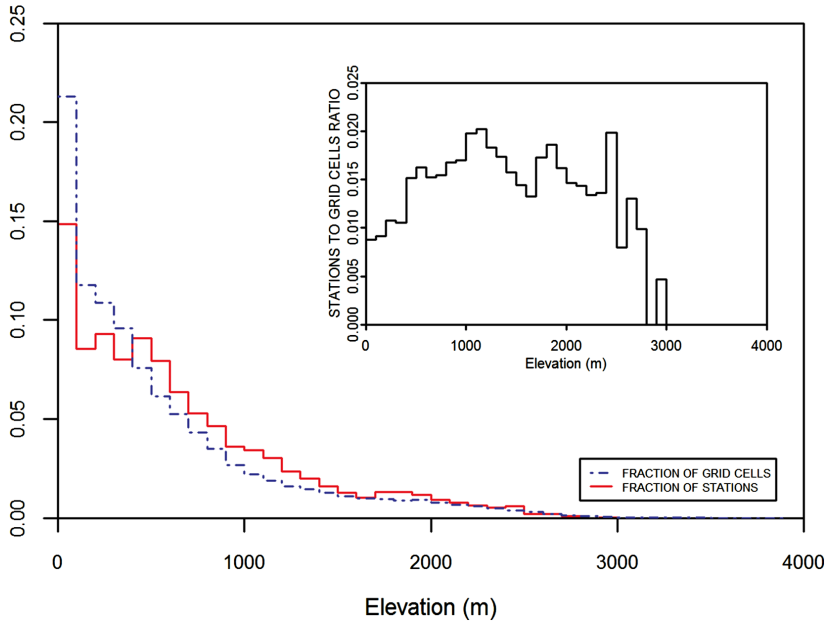

Figure 2. Vertical distribution of the 6134 stations (solid line) compared to the grid-cell elevation distribution (dashed line) in the domain for which we calculated the climatologies. The inset box shows the vertical distribution of the stations to grid-cells ratio. [Colour figure can be viewed at wileyonlinelibrary.com].

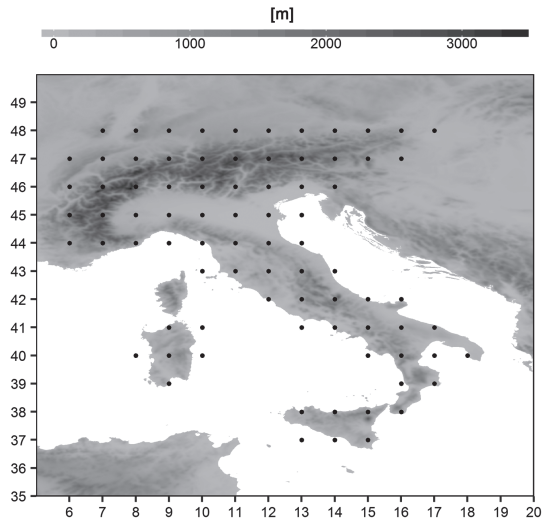

Figure 3. The 3smDEM and (superimposed) the 1° x 1° resolution grid used to locally optimize the decreasing coefficients of weighting factors.

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 $\Delta^{\text{par}}_{\frac{1}{2}}$:

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

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 $\Delta^{\text{par}}_{\frac{1}{2}}$ 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:

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

where $s^2\left\{\tilde{p}_{h_{\text{new}}}\right\}$ 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:

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

where i ranges over the stations involved in the regression and $\overline{h}$ 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) ($\boldsymbol{\varepsilon}$) 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

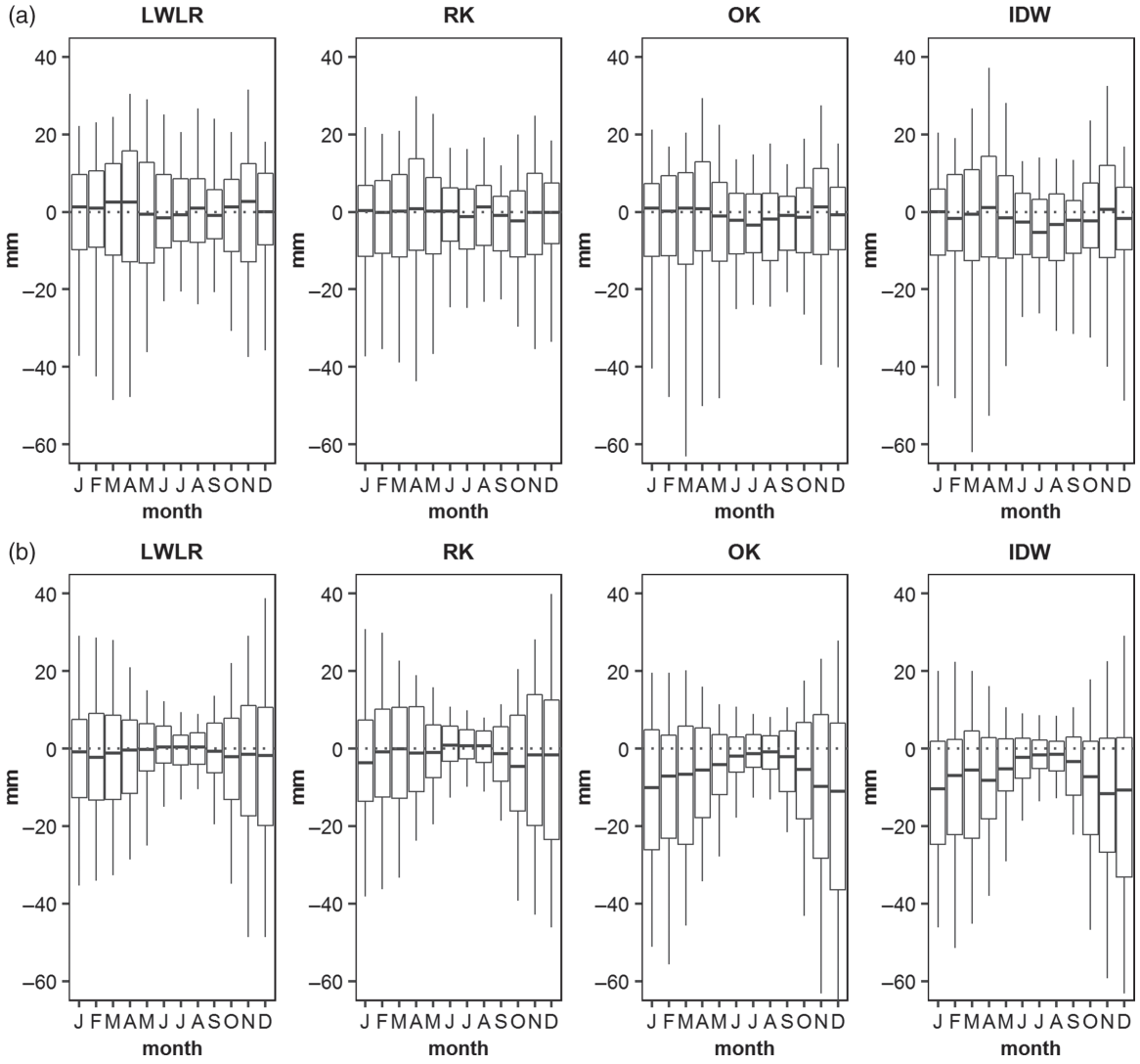

Figure 4. Monthly bias distribution of the reconstructed normals by the four methods for stations with (a) elevation >2000 m a.s.l. in Northern Italy and (b) elevation >1000 m a.s.l in Central-Southern Italy. The boxes range from the lower to the higher quartiles and are centred on the median; the whiskers 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

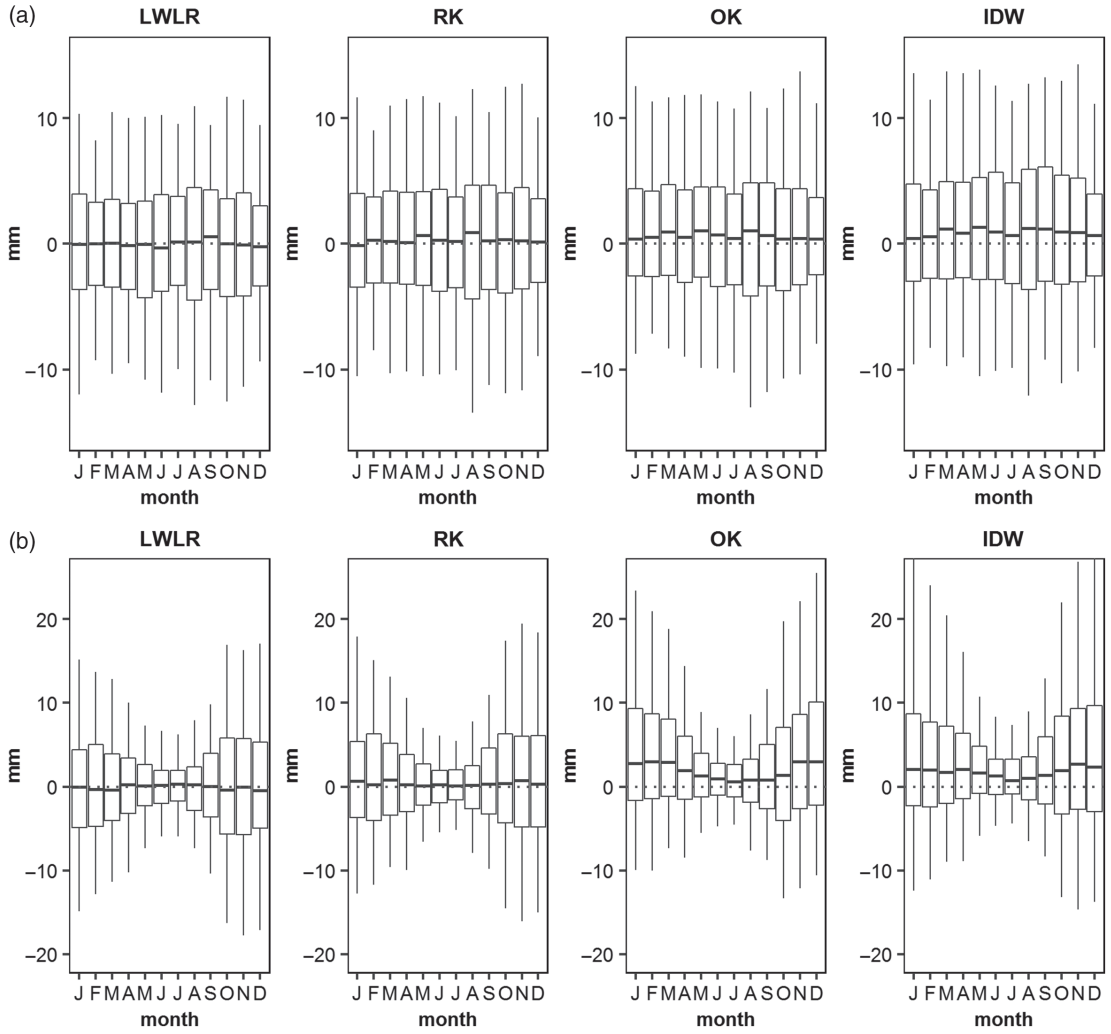

Figure 5. Monthly bias distribution of the reconstructed normals by the four methods for stations with elevation <100 m a.s.l. in (a) Northern Italy and (b) Central-Southern Italy. The boxes range from the lower to the higher quartiles and are centred on the median; the whiskers represent the minimum 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

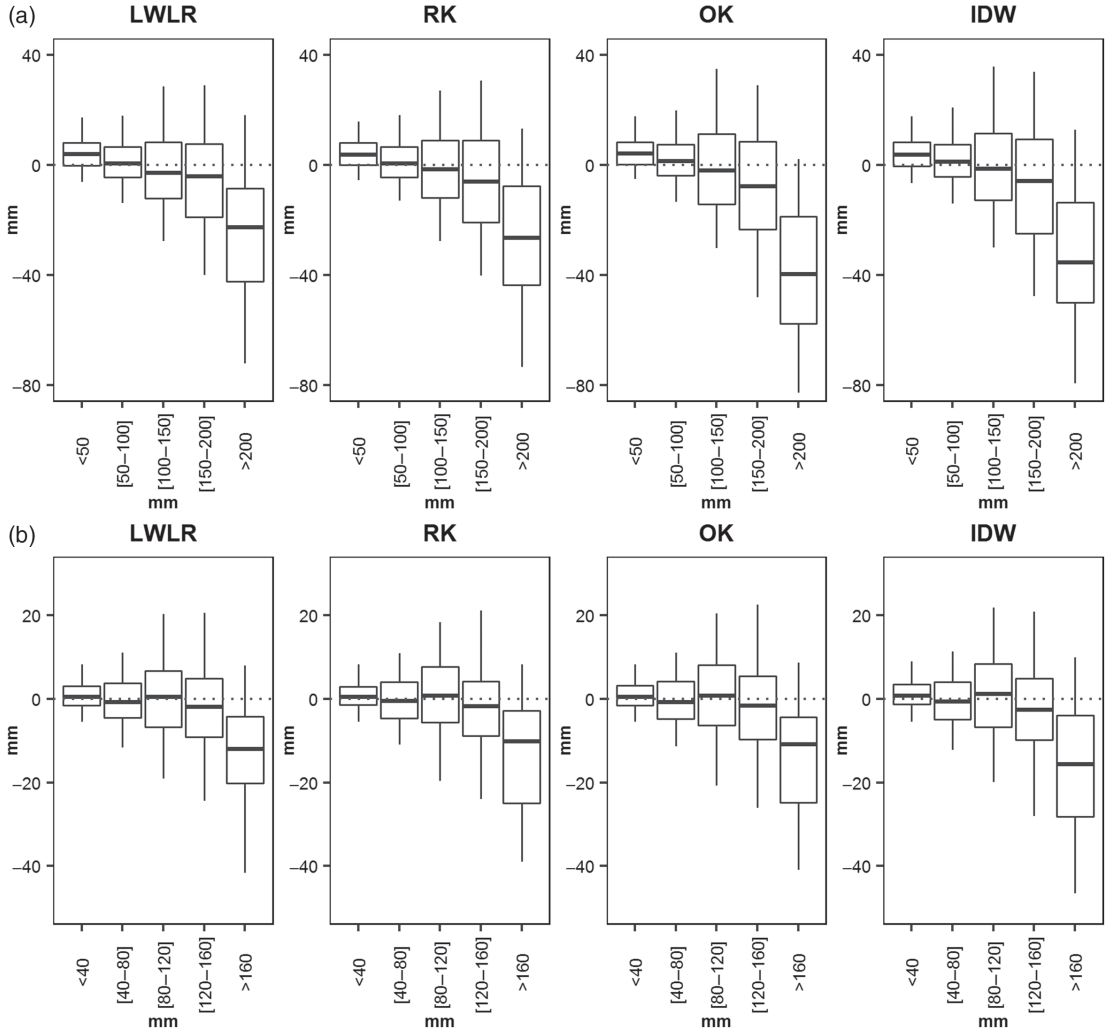

Figure 6. (a) January and (b) July bias distribution of the reconstructed normals by the four methods considering different ranges of station normals. The boxes range from the lower to the higher quartiles and are centred on the median; the whiskers represent the minimum and the maximum bias.

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,

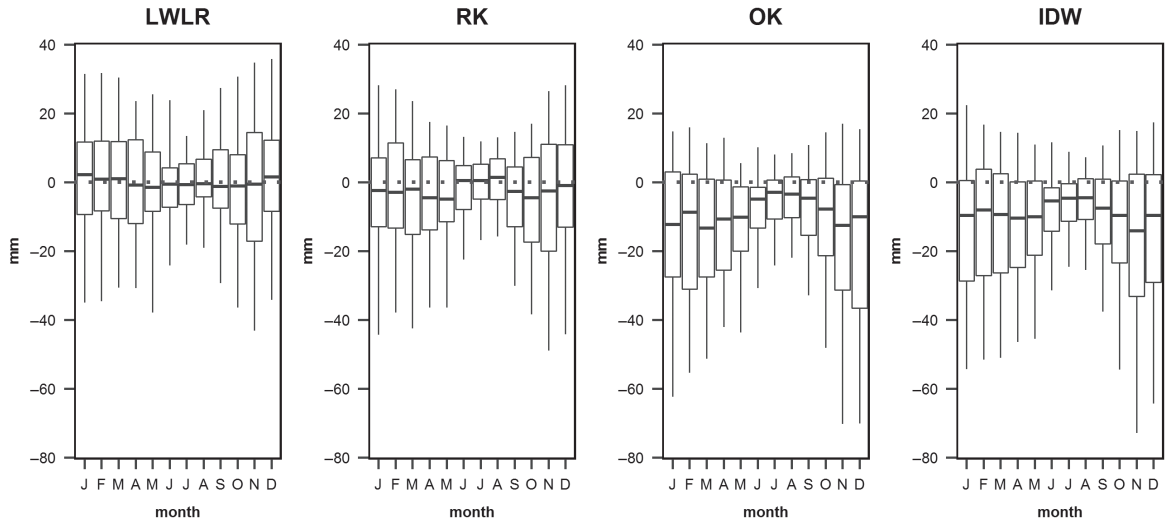

Figure 7. Monthly bias distribution of the reconstructed normals by the four methods for stations at higher elevation (at least 50 m) than the 10 nearest ones. The boxes range from the lower to the higher quartiles and are centred on the median; the whiskers represent the minimum and the maximum bias.

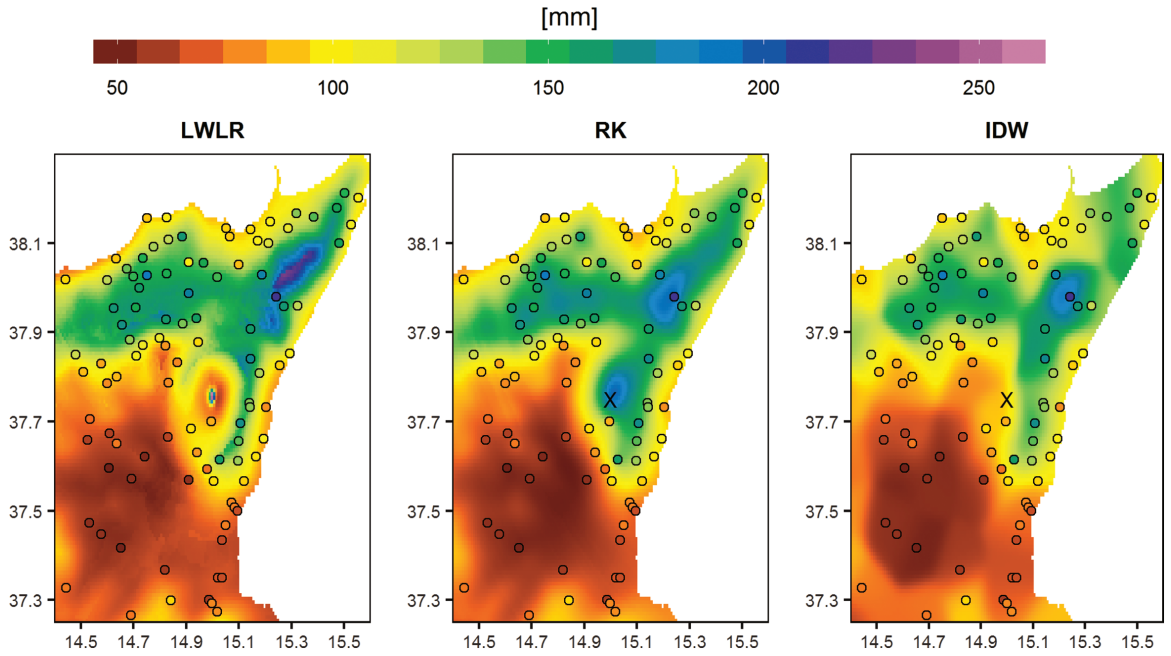

Figure 8. A detail of the January precipitation climatology over Mount Etna (Sicily) obtained by LWLR, RK and IDW. The points represent the available stations and the cross corresponds to the top of the volcano. OK climatology is not shown here as it provides a very similar precipitation distribution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

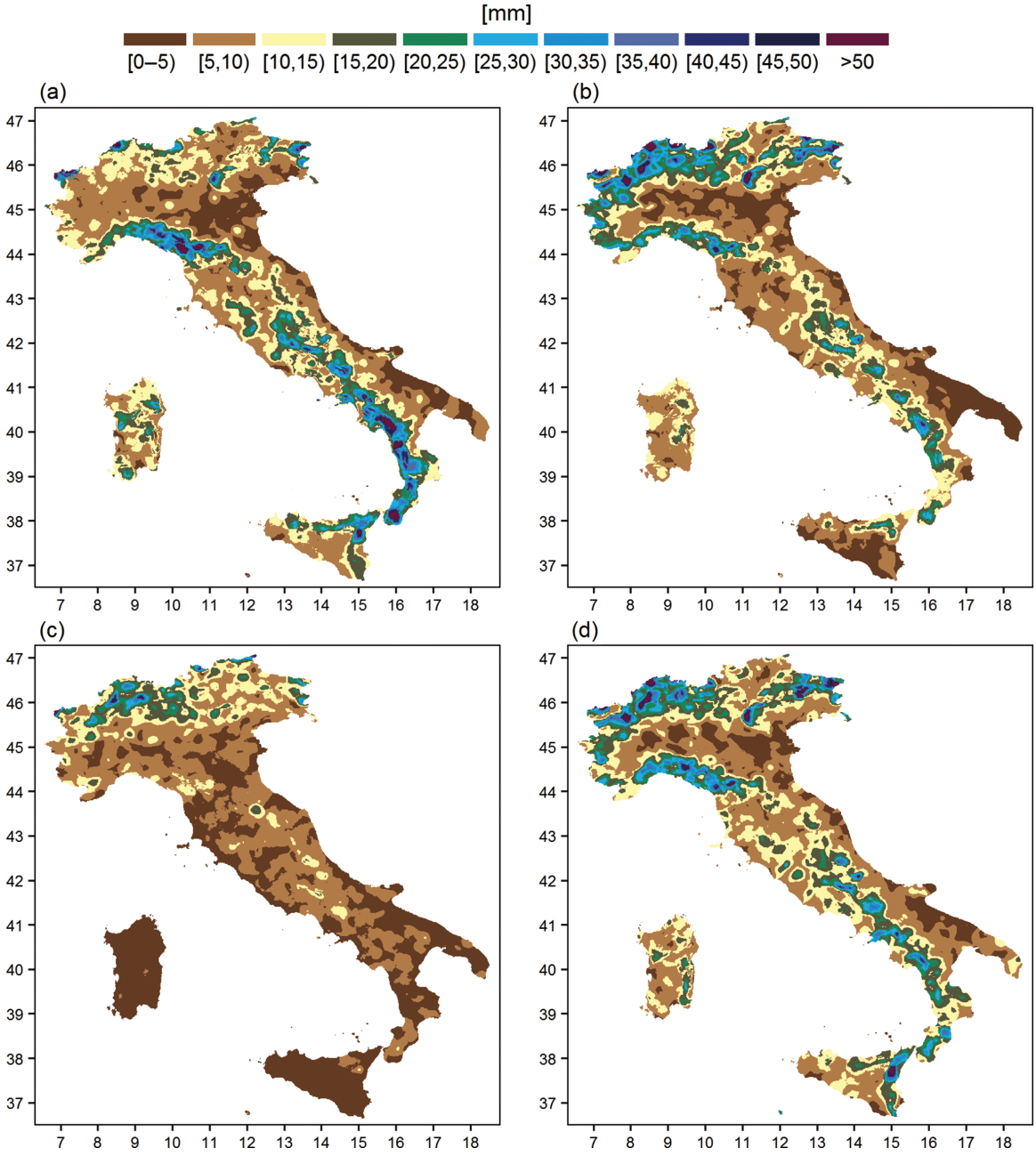

Figure 9. LWLR prediction interval half-width for (a) January, (b) April, (c) July and (d) October.

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

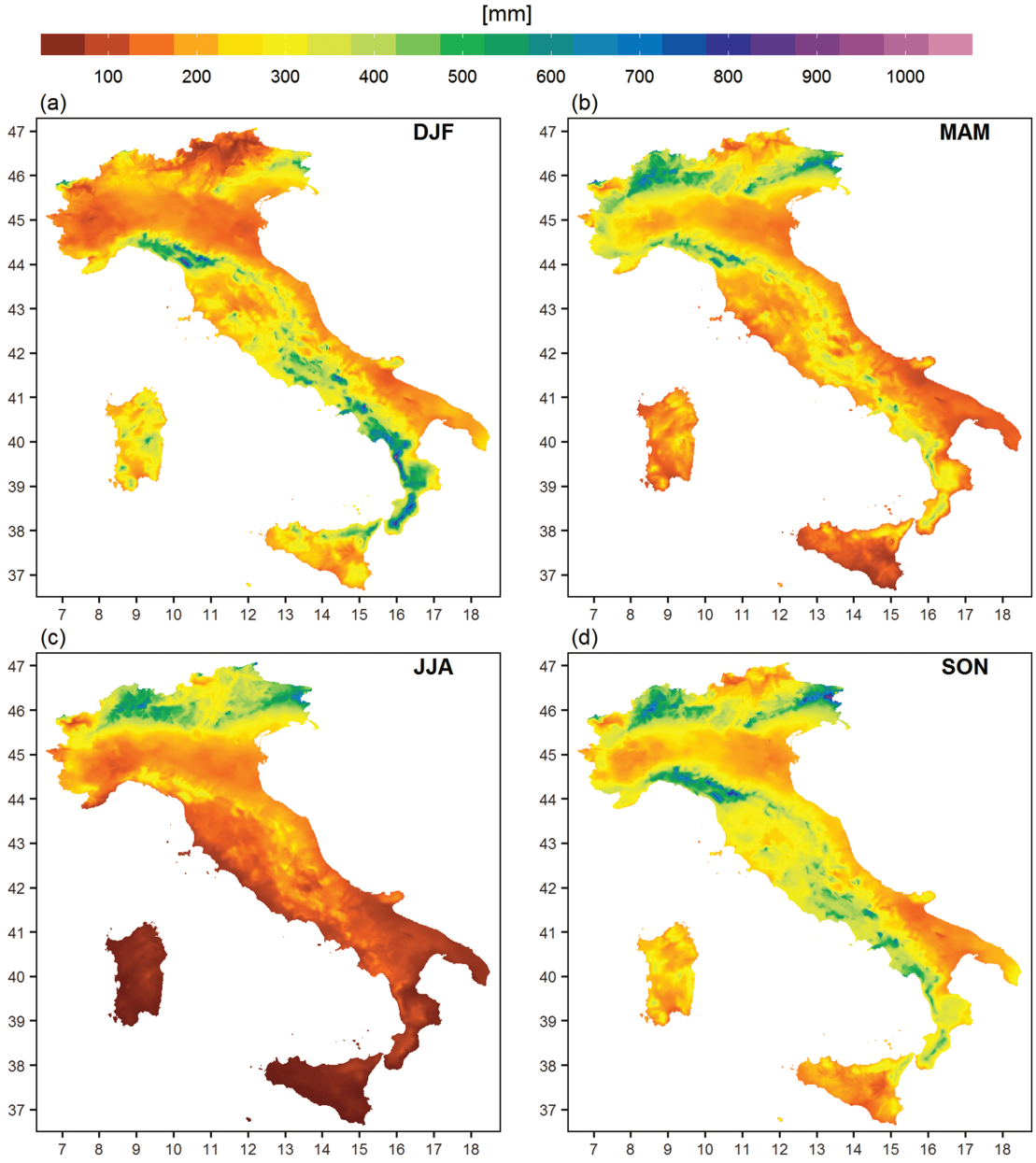

Figure 10. Seasonal LWLR precipitation climatologies.

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

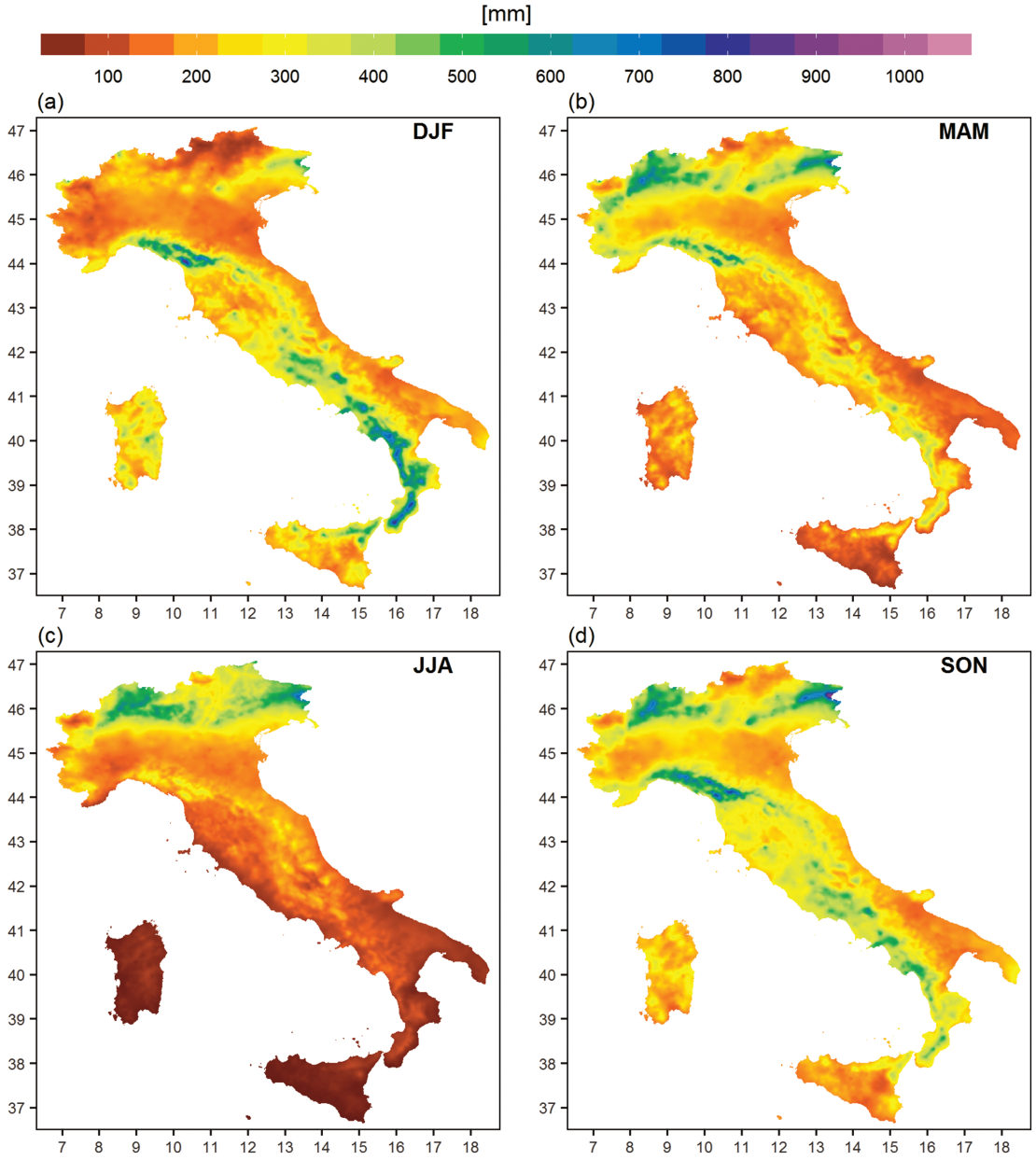

Figure 11. Seasonal RK precipitation climatologies.

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

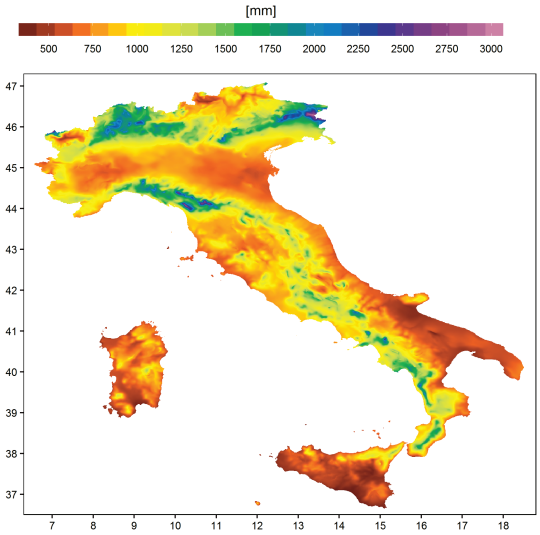

Figure 12. Annual LWLR precipitation climatology.

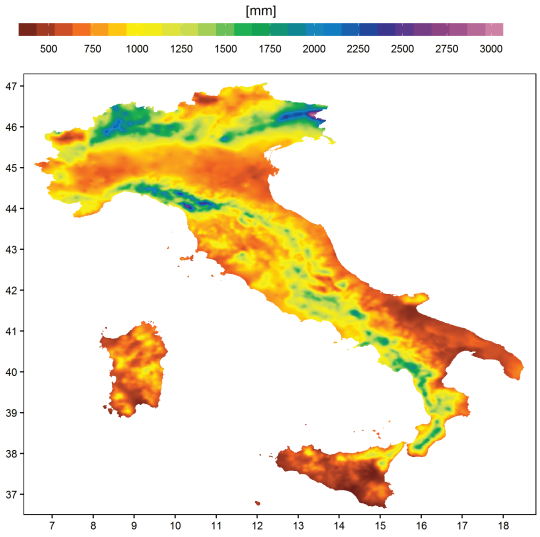

Figure 13. Annual RK precipitation climatology.

(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

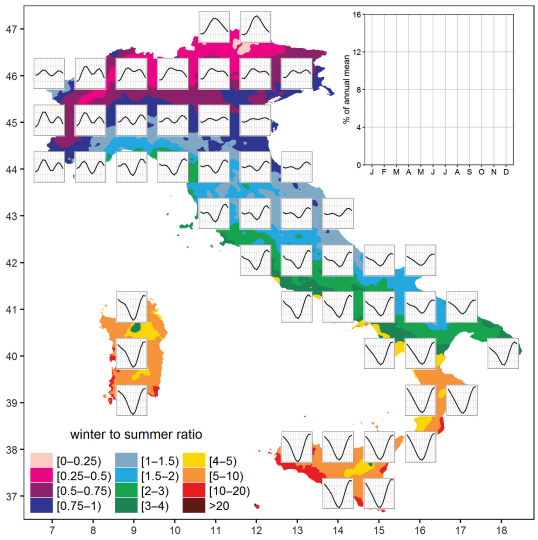

Figure 14. Distribution of winter (DJF) to summer (JJA) precipitation ratio over the domain and (superimposed) average yearly precipitation cycles over 1° sub-domains covering the whole study area. The inset box shows the range of the axes that is the same for all the plots.

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.

Acknowledgements

We thank all the institutions and projects whose data contributed to set up the 1961–1990 precipitation database (they are all listed in Table 1 which provides also information for data access) and the meteorological services from which we retrieved monthly normals (MeteoSwiss, MétéoFrance and ZAMG). Part of this work was supported by the Special Project HR-CIMA within the frame of the Project of National Interest NextData. We also thank Adriana Fassina, Diana Cricchio and Nicola Cortesi who contributed along the years to set up the database and Alessandro Delitala for providing metadata of the Sardinian stations. Finally, we acknowledge professor Gallus and professor Loewenstein for their help in improving the language of the manuscript and two anonymous reviewers for their useful comments and suggestions.

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.

Chapter 3

Spatio-temporal projection of precipitation over a complex high-mountain region

The presented work focuses on the reconstruction of a high-resolution dataset of long-term monthly precipitation series (1913-2015) for a mountainous area centred over the Forni Glacier (Central Alps). At this aim, the daily precipitation records of 30 rain-gauges available over the study area at different elevation bands were retrieved from both automatic and mechanical networks of regional and subregional services. The series were filled for daily gaps, checked for homogeneity and quality and converted to monthly scale. The newly recovered records were then integrated into the monthly database used to derive the Italian precipitation climatologies described in chapter 2. The 1961-1990 monthly climatologies over the study domain were computed at 30-arc second resolution and the 1913-2015 gridded dataset of precipitation series was then obtained by means of the anomaly method.
The study aimed at evaluating the ability to reconstruct the spatio-temporal distribution of precipitation for an area covered by a sparse network of in-situ observations without direct information at the highest elevations and over the glacialised surfaces. The accuracy of the results was thus computed by means of both the LOO validation on station sites and the comparison with the data collected by an Automatic Weather Station (AWS) operating on the Forni Glacier tongue since 2005 and not used into the interpolation procedure. The comparison with AWS pointed out the lowest agreement for winter when precipitation values were underestimated by the model.
The areal precipitation record over the period 1913-2015 obtained from the gridded dataset revealed no statistically significant long-term trends on both annual and seasonal scales, while a more relevant short-term variability was depicted.
The chapter is based on the peer-reviewed article published on *Advances in Meteorol-*

ogy:

Golzio, A., Crespi, A., Bollati, I. M., Senese, A., Diolaiuti, G. A., Pelfini, M., and Maugeri, M. (2018). High-Resolution Monthly Precipitation Fields (1913–2015) over a Complex Mountain Area Centred on the Forni Valley (Central Italian Alps), *Advances in Meteorology*, 2018, ID 9123814, pp. 17. doi:10.1155/2018/9123814

Hindawi
Advances in Meteorology
Volume 2018, Article ID 9123814, 17 pages
https://doi.org/10.1155/2018/9123814

Research Article

High-Resolution Monthly Precipitation Fields (1913–2015) over a Complex Mountain Area Centred on the Forni Valley (Central Italian Alps)

Golzio Alessio,[1] **Crespi Alice**,[2] **Irene Maria Bollati**,[1] **Senese Antonella**,[2] **Guglielmina Adele Diolaiuti**,[2] **Pelfini Manuela**,[1] **and Maugeri Maurizio**[2]

[1]*Earth Sciences Department, Università degli Studi di Milano, Via Mangiagalli 34, 20133 Milano, Italy*
[2]*Environmental Science and Policy Department, Università degli Studi di Milano, Via Celoria 2, 20133 Milano, Italy*

Correspondence should be addressed to Golzio Alessio; alessio.golzio@unimi.it

Received 26 October 2017; Revised 11 January 2018; Accepted 28 January 2018; Published 8 March 2018

Academic Editor: Harry D. Kambezidis

Mountain environments are extremely influenced by climate change but are also often affected by the lack of long and high-quality meteorological data, especially in glaciated areas, which limits the ability to investigate the acting processes at local scale. For this reason, we checked a method to reconstruct high-resolution spatial distribution and temporal evolution of precipitation. The study area is centred on the Forni Glacier area (Central Italian Alps), where an automatic weather station is present since 2005. We set up a model based on monthly homogenised precipitation series and we spatialised climatologies and anomalies on a 30-arc-second-resolution DEM, using Local Weighted Linear Regression (LWLR) and Regression Kriging (RK) of precipitation versus elevation, in order to test the most suitable approach for this complex terrain area. The comparison shows that LWLR has a better reconstruction ability for winter while RK slightly prevails during summer. The results of precipitation spatialisation were compared with station observations and with data collected at the weather station on Forni Glacier, which were not used to calibrate the model. A very good agreement between observed and modelled precipitation records was pointed out for most station sites. The agreement is lower, but encouraging, for Forni Glacier station data.

1. Introduction

Responses of the mountain environment to climate change represent one of the most studied topics in recent years [1–3]. The evident and fast landscape modifications occurring in the last decades testify the vulnerability of mountain areas to climate change (e.g., [4]). At higher altitudes cryosphere is shrinking and in the last 60 years Italian Alpine glaciers have lost about 30% of their volume [5], with a huge reduction of glacier surfaces [6] and a progressive widening of proglacial areas, which lead to new colonisation of glacier forelands [7, 8] and debris covered surfaces by vegetation [9–11] and animals [12, 13]. Such geomorphological changes involve not only the abiotic and biotic components but also the landscape fruition as they modify the local geoheritage [14], geodiversity [15], hazard, and risk scenarios [16–18].

As glacial and geomorphological processes are strictly climate related, accurate reconstruction and analysis of present and past climatic conditions are crucial. Qualitative reconstruction of the past climatic characteristics over long time scales is possible thanks to geomorphological and biological paleoclimatic indicators (e.g., typical features of glacial and periglacial environment, debris covered glaciers and rock glaciers, tree remnants under glacial deposits, and pollens) [19–22]. Quantitative reconstructions, instead, come from dendroclimatic analysis [23–25] or from meteorological observations that in Italy began to be collected regularly in the last decades of the XVIII century [26]. The longest records of meteorological data are however generally available for anthropized areas (in Italy they concern Milan, Padua, and Turin), with some excellences in high mountain environment, such as Capanna Margherita (Punta Gnifetti 4554 m, since 1899), Sonnblick (Austria, 3106 m, since 1886), and Jungfraujoch (Switzerland, 3466 m, since 1930). In the last decades many automatic weather stations (AWS) have been

positioned over glaciers and other high-elevated sites (e.g., at Forni Glacier in Valtellina (SO) [27]).

Therefore, mountain environments, especially glacier forelands, are extremely influenced by climate change but are also often affected by the lack of long and high-quality meteorological data, which limit the ability to deeply investigate the acting processes at local scale and their influence/impact on the biotic and abiotic environments as well as on the human component. This problem may be reduced, at least partially, by methodologies allowing us to extrapolate the existing observations over such areas. Many interpolation methods have been developed so far and allow reconstructing the high-resolution spatial distribution and temporal evolution of a meteorological variable over an extended area by integrating punctual station data and the topographic features of the surface, extracted from Digital Elevation Models (DEM). These methodologies allow us to obtain the climatic signal for a number of points much higher than the sparse station network and to get information even for those limited areas where no meteorological stations are present [28]. Results can be moreover integrated with dendroclimatic reconstructions, which can go back for hundreds to thousands years [29]. However, the ability of the model to reconstruct past climate series strongly depends on data distribution and availability as well as on orographic complexity of the area under investigation.

Temperature, precipitation, solar radiation, and wind are the main parameters that need to be widespread on glacier forelands. The present work is focused on precipitation as its distribution, characteristics, and trend have a great influence as triggering factors for hazard [3, 30]. They can represent limiting factors for biological community [31] and they can affect human fruition of the higher areas [32]. Moreover, the ability of quantifying the precipitation distribution (together with temperature) and its features over Alpine regions is extremely important to improve the estimation of mass balances over glaciers or hydrological balances over catchments.

Within this context, the paper aims at (I) presenting a methodology to reconstruct monthly precipitation time series (1913–2015) for any point of a 30-arc-second-resolution grid, covering a complex mountain area centred on the Forni Glacier (Central Alps) where an AWS is located since 2005; (II) analysing the spatial distribution of precipitation normals and precipitation temporal trends; (III) comparing the modelled precipitation series over Forni Glacier with the AWS records; and (IV) discussing the possible applications to multidisciplinary researchers focusing on glacier foreland changes.

2. Materials and Methods

2.1. Study Area and Data. The area selected for the present work is centred on the Forni Glacier, one of the largest Italian valley glaciers (11.34 km^2 [5]), and it covers about 2000 km^2 (grey shaded area in Figure 1).

In the last decades, this mountain area, as well as the similar ones all over the Alps, has suffered a rapid and deep modification due to climate change [33]. The most evident result is a generalized shrinkage of glacier bodies and, consequently, the expansion of glacier forelands with changes in the morphology of the upper valley portions. A deeply known moraine system (see for details [21]) characterises the Forni Valley allowing us to precisely date the glacier advancing phases since the Little Ice Age [34, 35] and to reconstruct the progressive glacier retreat and the following tree colonisation (ecesis) [8]. The germination of new plants in the proglacial areas, together with the increase of the glacial debris coverage, due mainly to gravity processes and ice ablation, contributes significantly to the deep transformation of the local glacial landscape. The debris coverage over the Forni Glacier ablation tongue is estimated to be increased from 26.7% in 2003 to 48.1% in 2015 [36]. This variation is in turn responsible for local microclimatic changes also in relation to the type of lithology constituting the debris. Land cover change caused by glacier shrinkage can modify climate through different mechanisms, some directly perturbing the Earth radiation budget and some perturbing other processes [37]. In fact, land cover change can alter the surface energy and moisture budgets through changes in evaporation and the fluxes of latent and sensible heat, directly affecting precipitation and atmospheric circulation as well as temperature [37].

The study area includes the Upper Valtellina (11 stations), the western part of the Venosta Valley (9 stations), and the western part of Sole Valley (9 stations), located in the Central Italian Alps, and a portion of the Grisons Canton (1 station), on the Swiss side of the Alps. A great part of this region is included within Stelvio National Park; see Figure 5 and Table 4. This area is characterised by a significant orographic heterogeneity with elevation values ranging from approximately 1000 m a.s.l. of the valley bottom to the 3769 m a.s.l. of the highest peak (Mount Cevedale). A bigger area was also considered (black border square region in Figure 1) in order to strength model computation of precipitation fields on the inner area. The considered weather stations (grey shaded area in Figure 1) can be grouped according to their distribution with respect to the relief features. Two stations belong to the *high mountain environment*, whose lower limit roughly corresponds to the upper limit of vegetation [38]; 8 stations are located on slopes *at middle altitude* (below the treeline) and 6 more stations are located at *Passes and plateau*. The remaining 14 stations are positioned at the *valley bottoms* (Table 4).

The database considered in the present work is composed by 30 daily precipitation series, covering the study area, which were retrieved from several Italian regional and subregional sources (Meteotrentino for the autonomous province of Trento, the Hydrological Office of the autonomous province of Bolzano/Bozen, the weather service of Regional Agency for Environmental Protection ARPA, in Lombardia, and the Geological Monitoring Service, CMG, weather network also in Lombardia) and from MeteoSwiss. Details on data sources are listed in Table 1. The series cover different time intervals in the 1913–2015 period and their elevation ranges are not equally distributed (Figures 2(a) and 2(b)). Fourteen of those series have been obtained by merging records from the traditional networks of mechanical stations with records from the new networks of automatic stations. This merging

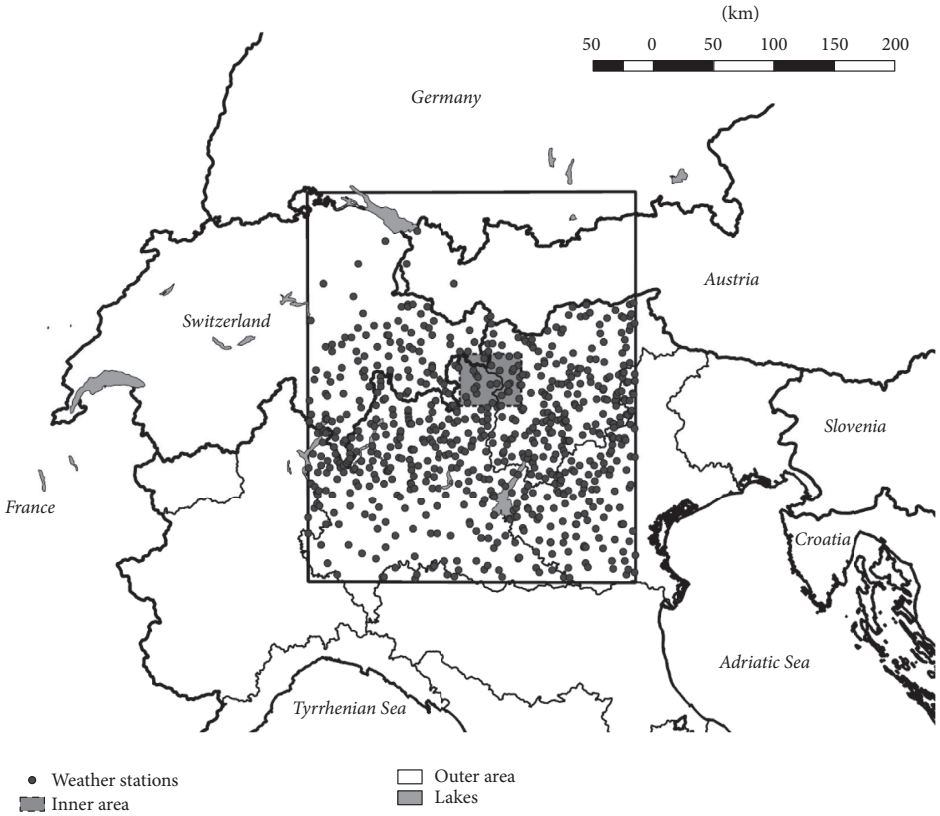

Figure 1: Study area domains: grey shaded the inner box area (longitude E 10°10.98′–10°49.14′ and latitude N 46°15.6′–46°37.34′) centred on the Forni Glacier, with the solid black square bordering the outer area (longitude E 8°36′–12°00′ and latitude N 45°00′–47°48′). Black dots indicate the weather stations considered.

Table 1: Meteorological networks considered in this work. References for datasets are available at Acknowledgments section.

Network	State	Region, province
ARPA Lombardia, Meteo	Italy	Lombardia, Sondrio
ARPA Lombardia, CMG	Italy	Lombardia, Sondrio
Meteotrentino	Italy	Trentino
Provincia autonoma di Bolzano, Bozen	Italy	Alto Adige
Meteo Svizzera, Meteo Swiss	Switzerland	Grigioni

was performed only if distance between the old and the new site was within 100 m and elevation difference within 30 m. Moreover, in some cases the two series had an overlapping period that supported their homogeneity assessment.

The daily data were subjected to a preliminary gross-check. Then, they were subjected to a gap-filling procedure in order to minimise the effect of sparse gaps in the daily records on the computation of monthly totals. For each station (test), all months with at least 15 available daily data were subjected to daily fulfilment procedure, where each daily reconstruction was performed by considering the daily records of a reference station. The reference station was selected considering distance, elevation gap, common days in the considered month, and correlation value with the test series. Correlation was computed by Pearson method and all series with correlation values less than 0.7 were discarded. Therefore, reference station was the nearest one showing the highest correlation and number of common data. The missing daily values were reconstructed using the multiplicative anomaly method [39]. More in detail, the average daily precipitation was computed for test and for reference stations on the basis of their common period, with these two averages, it was possible to compute the ratio that worked as scaling factor. The missing day in test series was reconstructed by rescaling the corresponding day in reference station using the aforementioned scaling factor.

On average, the percentage of daily data reconstructed for each station was about 1% (the maximum was about

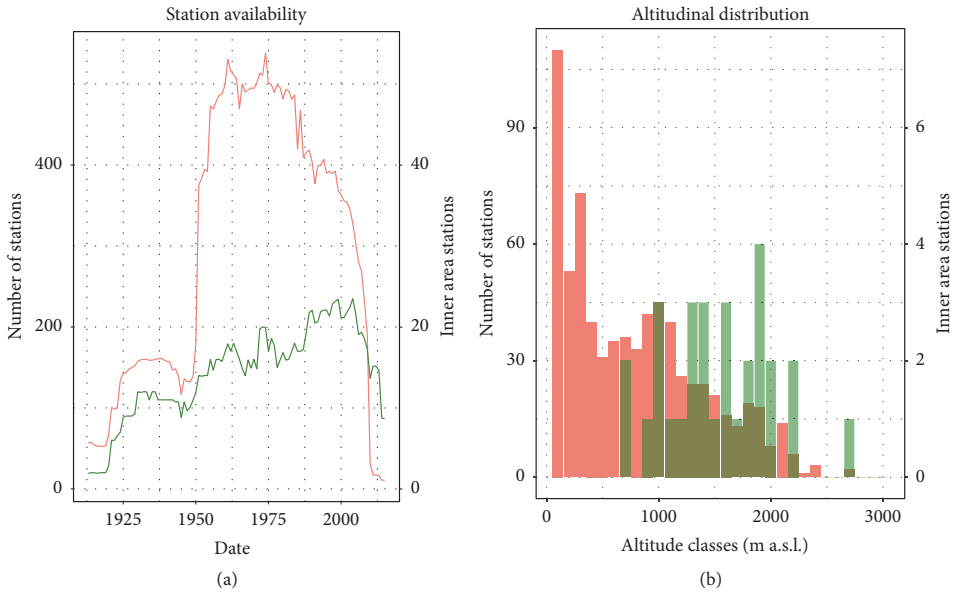

FIGURE 2: Availability of weather stations (a) from 1913 to 2015 and station distribution according to the altitude (b), with classes of 100 m. The inner study area is represented in dark green, and the outer area is represented in salmon.

5% at station Arnoga Valdidentro). The robustness of the filling procedure was evaluated by applying it to the available known records for each daily series and comparing the measured values to the simulated ones in terms of both daily precipitation values and number of wet days (Figure 3). Uncertainties were evaluated for the reconstruction of both absolute daily precipitation and number of wet days (days with precipitation greater than or equal to 1 mm) for each month. The uncertainties showed that for absolute daily cumulate precipitation root mean square error (RMSE) was around 2.5 mm and Mean Absolute Error (MAE) was 1.0 mm, with a total of 449534 simulated days. MAE was found to be more affected by the nonrainy days (precipitation lower than 1 mm), so we decomposed the computation of those errors into several rain classes. The chosen classes and the corresponding errors are summarised in Table 2 and highlight the increase in MAE and RMSE values with increasing precipitation, whereas the reconstruction of monthly wet days provided a quite good result: on a total of 14912 simulated months MAE was 1.5 days and RMSE was 2.1 days. These results are also reported as scatter plots in Figures 3(a) and 3(b).

After the daily fulfilment, monthly dataset was computed. Whenever missing daily data were still present, the corresponding monthly cumulate was not evaluated. The increment of daily data, even though it was little, had permitted to increase significantly the monthly precipitation availability: on average the benefit was around 15%, and the maximum increment was obtained for station Arnoga Valdidentro, that passed from 77 to 182 available months.

In order to investigate monthly series in terms of quality and homogeneity and to deal with the points at the borders of study area in the climatological reconstruction, they were integrated by several monthly precipitation series retrieved from a quality-checked database recently set up to compute the 1961–1990 monthly precipitation climatologies for Italy [40]. All the available stations in a wider area (latitude north 45°00′–47°48′ and longitude east 8°36′–12°00′) were included reaching a total of 734 series and a spatial density of about 9 stations/10^3 km^2.

Besides the 30 station series, we also considered, but did not include in the database used as input for the model, the daily observations collected by an AWS (named AWS1 Forni) that has been operating on the ablation tongue of the Forni Glacier since 2005 and monitors continuously the most important meteorological variables. It represents a precious and unique source of data directly measured on a glacier surface and they were used as validation of the model outputs.

2.2. Monthly Quality-Check and Homogenisation. The quality-check and homogenisation activities on the 30 monthly series in the domain were carried out in order to remove outliers and spurious values and to detect nonclimatic signals due to, for example, changes in station location, station malfunctions, changes in surroundings, sensors, and/or measurements techniques [41].

Quality-check was performed for all the 30 stations inside the domain, by following the method described in Crespi et al. [40]. It is based on the comparison between the measured series (test) and the simulated one reconstructed on the basis of ten neighbouring stations (reference) with enough data in common with the station under consideration.

The homogenisation algorithm is based on the Craddock test [42] and the Expert Judgment method [43, 44]. It is

Table 2: Reconstruction of absolute precipitation skill for precipitation classes subdivision.

Rain class (mm)	Days	MAE (mm)	RMSE (mm)
[0.0–1.0)	326043	0.22	0.96
[1.0–2.5)	32270	1.42	2.10
[2.5–5.0)	27178	2.13	2.86
[5.0–7.5)	17492	2.83	3.69
[7.5–10.0)	10898	3.45	4.54
[10.0–15.0)	14115	3.97	5.26
[15.0–20.0)	8056	4.67	6.22
[20.0–30.0)	7776	5.49	7.40
[30.0–40.0)	3180	6.55	8.87
[40.0–50.0)	1344	7.56	10.31
[50.0–75.0)	984	9.16	12.29
[75.0–100.0)	159	11.92	16.24
[100.0–125.0)	30	13.98	17.48
[125.0–150.0)	7	28.93	35.38
[150.0–200.0)	2	21.67	22.11

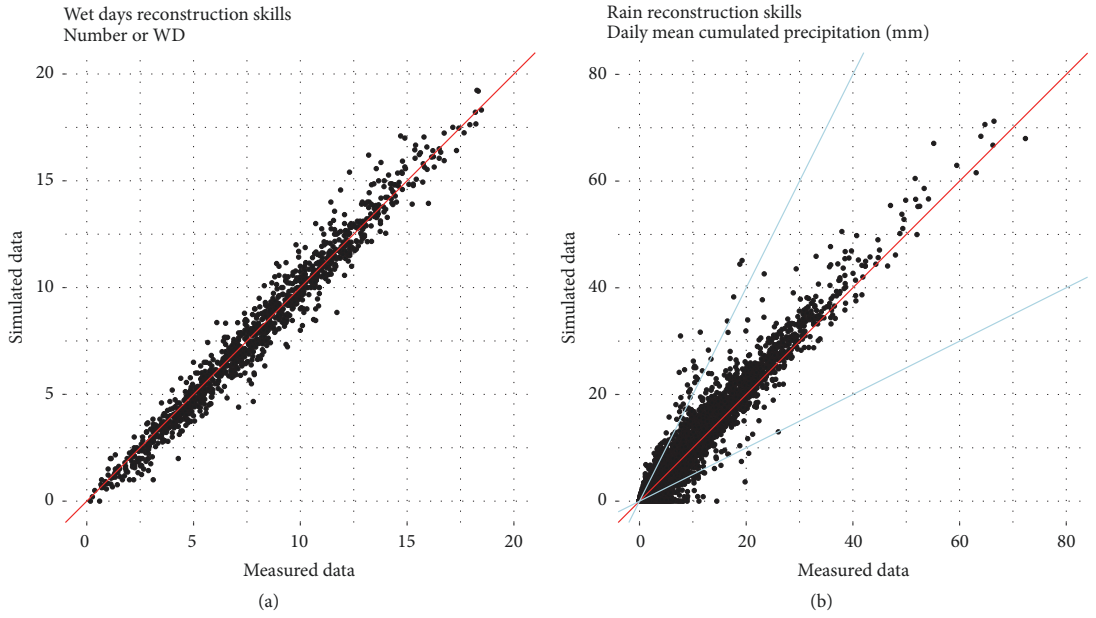

Figure 3: Scatter plot of simulated versus measured monthly wet days (a) and daily rainfall (b). The skill score is the red line (1 : 1), while blue lines represent the ratios 2 : 1 and 1 : 2.

based on the assumption that if two series are homogeneous their ratios should be constant in time. Therefore, for each test series, five reference stations are chosen based on distance, altitude criterion, and amount of common data. The monthly series of cumulative relative differences (Craddock series) between test and each reference were computed, and inhomogeneous periods were selected manually and corrected by the proper correcting factors; nine series out of 30 were homogenised for relative short periods. In Table 3, the homogenised series are listed together with the corrected periods, while in Figure 4 the annual series of Bormio station before and after the homogenisation is reported as an example.

After these procedures, the monthly normals in the period 1961–1990 were computed for each station. In order not to introduce biases in the normal computation due to unequal number of available data, missing months were reconstructed by following the procedure described in Crespi et al. [40] and based on the same method applied for quality-check.

Table 3: List of homogenised precipitation series and their homogenised periods.

Station name	Network	Homogenised periods
Bormio	ARPA Lombardia	1930–1959
Cancano (Valdidentro)	ARPA Lombardia	1951–1969
Solda	Provincia di Bolzano	1930–1965, 1966–1995
Careser (dam)	Meteotrentino	1975–1992
Cogolo Pont (Centrale)	Meteotrentino	1933–1940, 1990–1999
Passo del Tonale	Meteotrentino	1923–1930, 1959–1971
Mezzana	Meteotrentino	1990–2000, 2009–2016
Malga Mare (Centrale)	Meteotrentino	1974–2016
Peio	Meteotrentino	1930–1950, 1944, 1960–1972

Table 4: Inner area station characteristics: the ID column refers to Figure 5.

ID	Station name	Province	Lat N	Lon E	Elevation [m]	Positioning	Side
1	Grosotto	Sondrio	46.286	10.261	613	Valley bottom	
2	Premadio centrale	Sondrio	46.485	10.351	1305	Valley bottom	
3	Trafoi	Bolzano	46.550	10.508	1570	Mid mountain	E
4	Pezzo	Brescia	46.283	10.519	1557	Valley bottom	
5	Agumes Prato	Bolzano	46.621	10.582	907	Valley bottom	
6	Gioveretto dam	Bolzano	46.509	10.725	1851	Mid mountain	SE
7	Ganda	Bolzano	46.550	10.786	1257	Mid mountain	NW
8	Lago Verde	Bolzano	46.467	10.803	1488	Mid mountain	E
9	Rabbi Somrabbi	Trento	46.410	10.804	1352	Mid mountain	S
10	Santa Caterina Valfurva	Sondrio	46.415	10.498	1760	Valley bottom	
11	Forni	Sondrio	46.421	10.564	2118	Mid mountain	Bottom
12	Cancano	Sondrio	46.514	10.323	1855	Mid mountain	Bottom
13	Bormio	Sondrio	46.466	10.371	1225	Valley bottom	
14	Solda	Bolzano	46.514	10.594	1905	Valley bottom	
15	Lasa	Bolzano	46.615	10.693	865	Valley bottom	
16	Silandro	Bolzano	46.622	10.779	698	Valley bottom	
17	San Giovanni Martello	Bolzano	46.526	10.736	1616	Mid mountain	E
18	Pian Palù (dam)	Trento	46.337	10.614	1800	Mid mountain	NE
19	Careser (dam)	Trento	46.423	10.699	2600	High mountain	NW
20	Mezzana	Trento	46.313	10.796	905	Valley bottom	W-E
21	Passo del Tonale	Trento	46.263	10.597	1875	Mid mountain	Pass
22	Peio	Trento	46.364	10.678	1585	Mid mountain	SSE
23	Sondalo, Le Prese	Sondrio	46.350	10.355	950	Valley bottom	
24	Aquilone	Sondrio	46.401	10.355	1082	Valley bottom	
25	Grosio, Fusino dam	Sondrio	46.327	10.246	1220	Mid mountain	S
26	Valdisotto	Sondrio	46.384	10.330	2114	High mountain	E
27	Sta Maria Val Muestair	Grigioni	46.600	10.433	1383	Valley bottom	
28	Cogolo Pont	Trento	46.365	10.689	1190	Valley bottom	N-S
29	Passo del Tonale (old)	Trento	46.263	10.611	1795	Mid mountain	Pass
30	Malga Mare	Trento	46.414	10.681	1950	Mid mountain	SE

2.3. Interpolation Models and Anomaly Method. The 1961–1990 high-resolution precipitation climatologies were computed for each grid cell of a 30-arc-second-resolution DEM (GTOPO 30 DEM, U.S. Geological Survey) covering the whole study domain. This DEM has a root mean square error (RMSE) of about 18 m [45]. Since many climatological works (see, e.g., [46]) proved that precipitation distribution is strongly linked to orography, especially elevation, and the interpolation techniques taking into account terrain features are proved to provide better performances, the monthly normals were distributed applying a Local Weighted Linear Regression (LWLR) of precipitation versus elevation [47]. LWLR computed the precipitation normal p at each DEM cell (x, y) by considering its elevation (h):

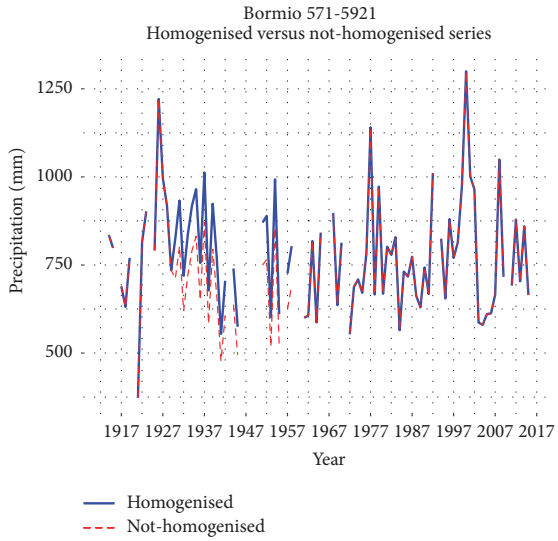

Figure 4: Bormio station annual cumulated precipitation: the homogenised series (blue solid line) is compared with the original series (red dashed line). The homogenised period 1930–1959 is clearly visible, when the original series underestimates precipitation.

$$p(x,y) = a(x,y) + b(x,y) \cdot h(x,y), \tag{1}$$

where $a(x, y)$ and $b(x, y)$ are the coefficients of the weighted linear precipitation-elevation regression performed at the considered cell by selecting the stations with the highest weights.

Station weights ($w_i^{\text{par}}(x, y)$) are expressed as the product of several weighting factors in the form of Gaussian functions (2) in which the distances and the level of similarity between the station cells and the considered grid cell in terms of orographic features (i.e., elevation, slope steepness, and slope orientation) are taken into account:

$$w_i^{\text{par}}(x,y) = e^{-((\Delta_i^{\text{par}}(x,y))^2/c^{\text{par}})}, \tag{2}$$

where $\Delta_i^{\text{par}}(x, y)$ is the difference of the parameter values between the station i and the grid cell and c^{par} is the coefficient regulating weight decrease.

Orographic features of each station were extracted from the closest DEM cell, while the most suitable decreasing coefficients to be used for the weighting factors were optimised month-by-month following the procedure described in Crespi et al. [40].

In order to assess the suitability of the chosen interpolation model, we compared it with the widely used Regression Kriging (RK) approach. This comparison allowed us to test model robustness and to assess the most suitable spatial scale at which the precipitation-elevation relationship has to be studied on the domain.

In RK, a global precipitation-elevation regression is computed on all the stations in the domain and the station residuals are then interpolated onto high-resolution grid by ordinary kriging. Precipitation normal at cell (x, y) is finally obtained as follows:

$$p(x,y) = a + b \cdot h(x,y) + \mathbf{k}^{\mathrm{T}}(x,y) \cdot \boldsymbol{\varepsilon}, \tag{3}$$

where a and b are the global regression coefficients, $\mathbf{k}^{\mathrm{T}}(x, y)$ are the kriging weights, and $\boldsymbol{\varepsilon}$ are station residuals. Further details on RK approach can be found in, for example, Hengl et al. [48].

Moreover, in order to define the actual spatial scale at which the direct effects of elevation on precipitation are important [49], a Gaussian filter was applied to smooth 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 3 grid steps from the cell itself; the degree of smoothing was chosen after an optimisation procedure (see more details in Crespi et al. [40]). The DEM is depicted in Figure 5. Climatologies were therefore computed on this smoothed version of the DEM (3smDEM), which is also used to assign to the stations the orographic parameters required in the interpolation procedure.

After computing the distribution of 1961–1990 precipitation normals on the domain, the century records for the period 1913–2015 were obtained for each grid point by applying the anomaly method [50]. This methodology is based on the assumption that the spatiotemporal behaviour of meteorological variables can be described by the superimposition of a constant field (i.e., the climatologies) and the departures from them (i.e., the anomalies). At this aim, long and high-quality homogenised series are required to avoid reconstructing nonclimatic signals. Due to the higher spatial coherence of anomalies in respect to climatologies, interpolation approaches considering larger scales and neglecting local surface features are preferable [28]. Monthly precipitation series from all stations were converted into anomaly series as the ratio to the corresponding 1961–1990 monthly climatologies (normals). Afterward they were interpolated onto the 30-arc sec grid by averaging the anomalies of neighbouring stations (within 200 km from the grid point) by means of Gaussian weighting functions (in the form of (2)), taking into account their distance from the considered grid cell and elevation difference.

Due to the evolution of data coverage during the considered time interval, the weight decrease was regulated for each year according to station availability. While for elevation weight a constant halving coefficient was set, for distance weight it coincided with the mean radius from the grid point within at least 3 available stations that were included. This value ranged between 45 km at the beginning of the period to 10 km for the central decades when station coverage was maximum. High-resolution field of 1913–2015 monthly precipitation series in absolute values was finally obtained by multiplying gridded anomalies for 30-arc-second-resolution climatologies.

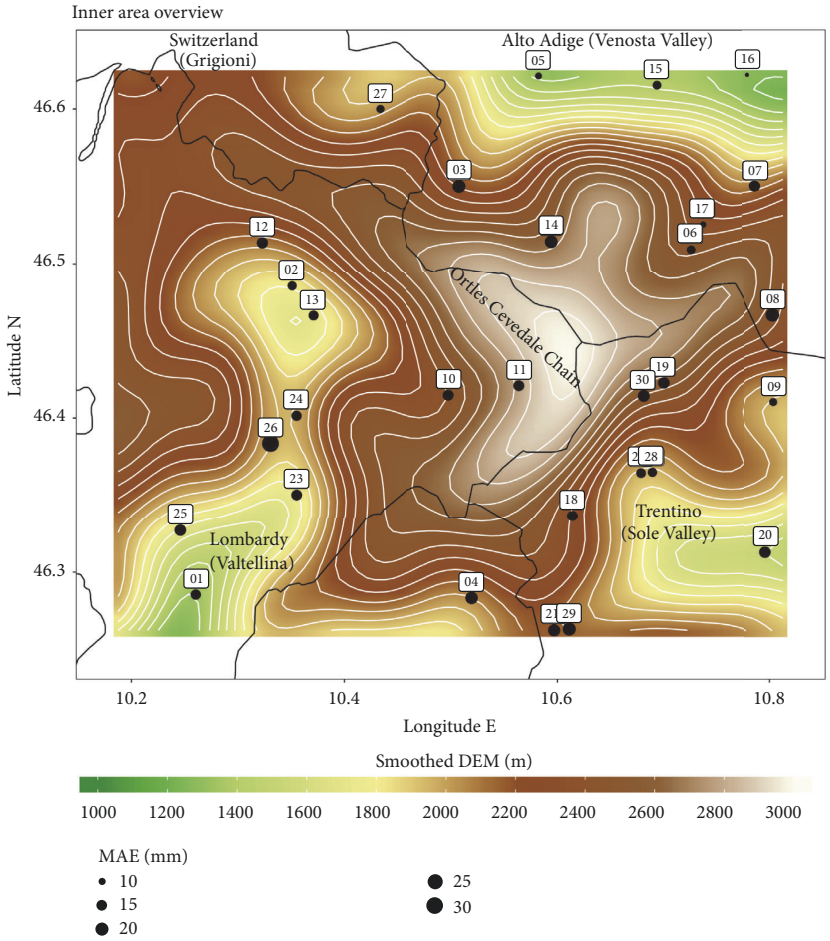

FIGURE 5: The inner area with stations, numbered following Table 4, the smoothed DEM with contour lines, and station MAEs as size of station points.

2.4. AWS1 Forni as Model Validator. The statistical model, as previously described, was set up using a net of weather stations ranging from valley bottom to the mountain top. Each station insists on a specific area and altitude gap, and its influence decreases progressively (see Section 2.3). Each station represents its surroundings, especially in such a complex terrain, and overlaps to the predominant climate signal a microclimate forcing. In order to assess the ability of the reconstruction approach to extrapolate the local microclimatic conditions of the area, especially over those points which are less monitored by in situ measurements, we compared the modelled series for a gird point over the Forni Glacier with the precipitation observations collected by AWS1 Forni. This station was not included in the database entering in the model, so its data were completely new and a very good start point for the validation.

The station is equipped with a not-heated rain gauge, which measures the summer (from June to October) liquid precipitation, and two sonic ranger sensors, which measure the snow depth. For the estimation of daily snow water equivalent (SWE) an approach based on hourly snow depth data and a fresh snow density of 140 kg/m^3 was applied [27, 51]. This method was validated in a previous study by comparing the estimated SWE values with data measured by means of manual snow pits and recorded by an automatic snow pillow [52].

Until October 2014, the AWS1 Forni was situated on the lower sector of the eastern tongue of the Forni Glacier. Because of the increased number of crevasses that made it very difficult and hazardous to reach the station site, the AWS was moved to an upper area. Unfortunately, due to the formation of ring faults that could compromise the stability of the station itself [53, 54], in November 2015 the AWS was moved on the central tongue. All the changes in AWS location are reported in Table 5. The actual AWS1 Forni position was set to the average of the three locations. The nearest DEM cell to this position is 2101 (Figure 6), which was used together with the other eight nearest points to compare the model with the station data.

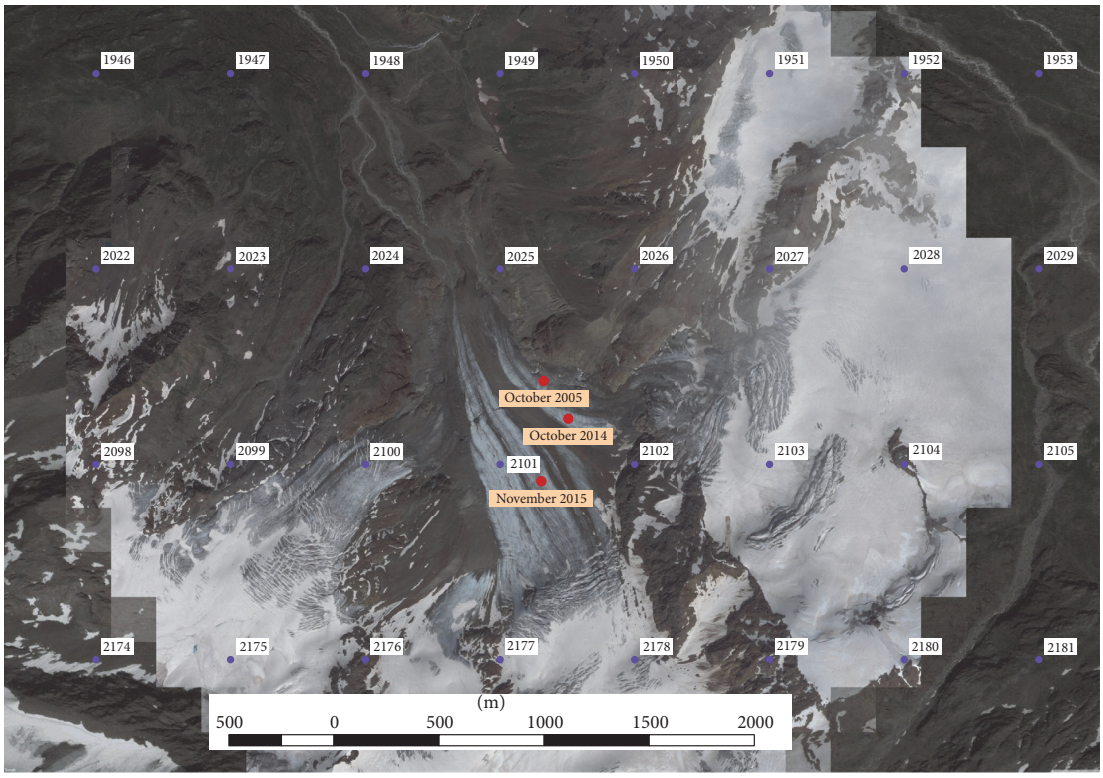

Figure 6: Inner area grid points (violet dots) and the locations of AWS1 Forni at Forni Glacier (red dots), superimposed to Google Earth© satellite images.

Table 5: Forni supraglacial weather station positions in the last years.

Beginning period	End period	Latitude N	Longitude E
October 2005	October 2014	46°23′57.8″	10°35′24.7″
October 2014	November 2015	46°23′52.0″	10°35′30.2″
November 2015	Today	46°23′42.4″	10°35′24.2″

3. Results and Discussion

3.1. Climatologies. The monthly climatologies (1961–1990 normals) were evaluated by LWLR for the inner domain reported in Figure 1 and model accuracy was evaluated in terms of Leave-One-Out (LOO) reconstruction of monthly normals of the 30 stations of this area, that is, by excluding the station meant to be reconstructed in order to avoid self-influence. Model uncertainties were estimated in terms of monthly MAE, RMSE, and BIAS. In order to assess the suitability of the chosen approach for the peculiar study area, model results were compared to those obtained by RK and the corresponding monthly errors are reported in Table 6. On average LWLR had shown lower values for MAE and RMSE with respect to RK, except for summer months when RK performances were comparable or even better than LWLR. The reason could be found in the peculiar LWLR extrapolation ability: the local precipitation-elevation regression tends to enhance precipitation amounts in summer when the highest station monthly normals occur. However, the annual evolution of climatologies for each considered station was well captured by both models.

Regarding the annual climatological map depicted in Figure 7, a very dry zone was reconstructed in Venosta Valley (northeastern part of the box), with less than 400 mm/yr, while the mountain chain of Ortles-Cevedale is more wet, with 1000–1100 mm/yr. Meanwhile the Upper Valtellina (Bormio, Valfurva, and Livigno) has an intermediate precipitation pattern (around 750–900 mm/yr). Moreover, in the western part of the area is visible a very high precipitation region (up to 1400–1500 mm/yr): this is related to the high precipitation measured across the Bernina Pass that has Stau

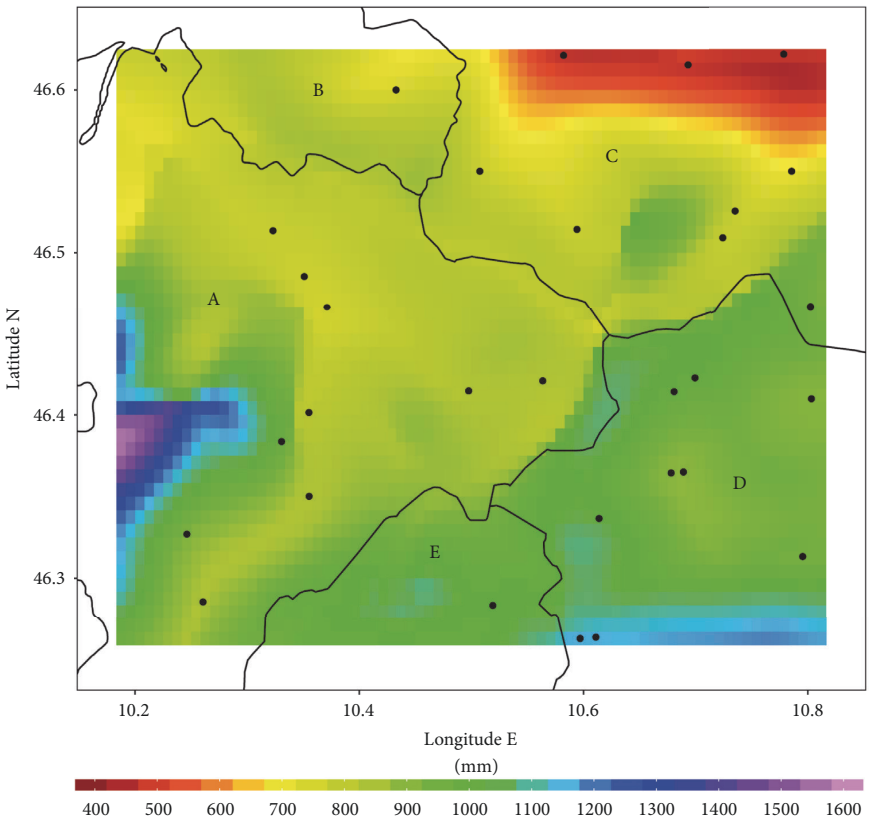

Figure 7: Annual climatology (1961–1990) evaluated with the LWLR method. Black dots indicate rain gauges; black lines indicate borders between regions. Letters indicate administrative regions: (A) Lombardy, Sondrio province; (B) Switzerland, Grisons canton; (C) Alto Adige Südtirol; (D) Trentino; (E) Lombardy, Brescia province.

Table 6: LWLR and RK leave-one-out uncertainties. BIAS were evaluated as difference between measured and simulated values.

Month		LWLR			RK	
[mm]	BIAS	MAE	RMSE	BIAS	MAE	RMSE
January	−0.48	4.6	5.8	−0.92	7.3	8.8
February	−1.52	5.2	6.3	−1.80	7.3	8.5
March	−1.20	5.8	7.2	−1.06	8.0	9.5
April	−0.94	8.2	10.3	−0.71	11.3	13.8
May	−0.81	9.9	12.7	−1.16	9.4	12.4
June	−0.78	7.7	9.5	−0.35	7.1	9.8
July	−0.16	7.6	9.9	0.13	7.7	9.6
August	−1.12	8.3	10.3	−0.57	8.2	10.6
September	−0.42	7.7	9.8	−0.58	8.1	11.2
October	−0.43	8.3	11.4	−0.91	8.8	11.8
November	−1.39	7.2	9.2	−1.12	7.9	10.1
December	−0.22	5.1	6.3	−0.46	6.5	7.8

conditions from both north and south sides. Considering an average on the entire area the rainfall is 880 mm/yr.

As pointed out by annual precipitation distribution, Ortles-Cevedale chain acts as a pluviometric shadow, as it stops the southerly wet air masses. This behaviour is even more evident observing the seasonal climatologies in Figure 8, especially during winter. Moreover the annual cycle points out the contrast between a very dry winter (even less than 100 mm/season, especially on Upper Valtellina and Venosta Valley) and a very wet summer (up to

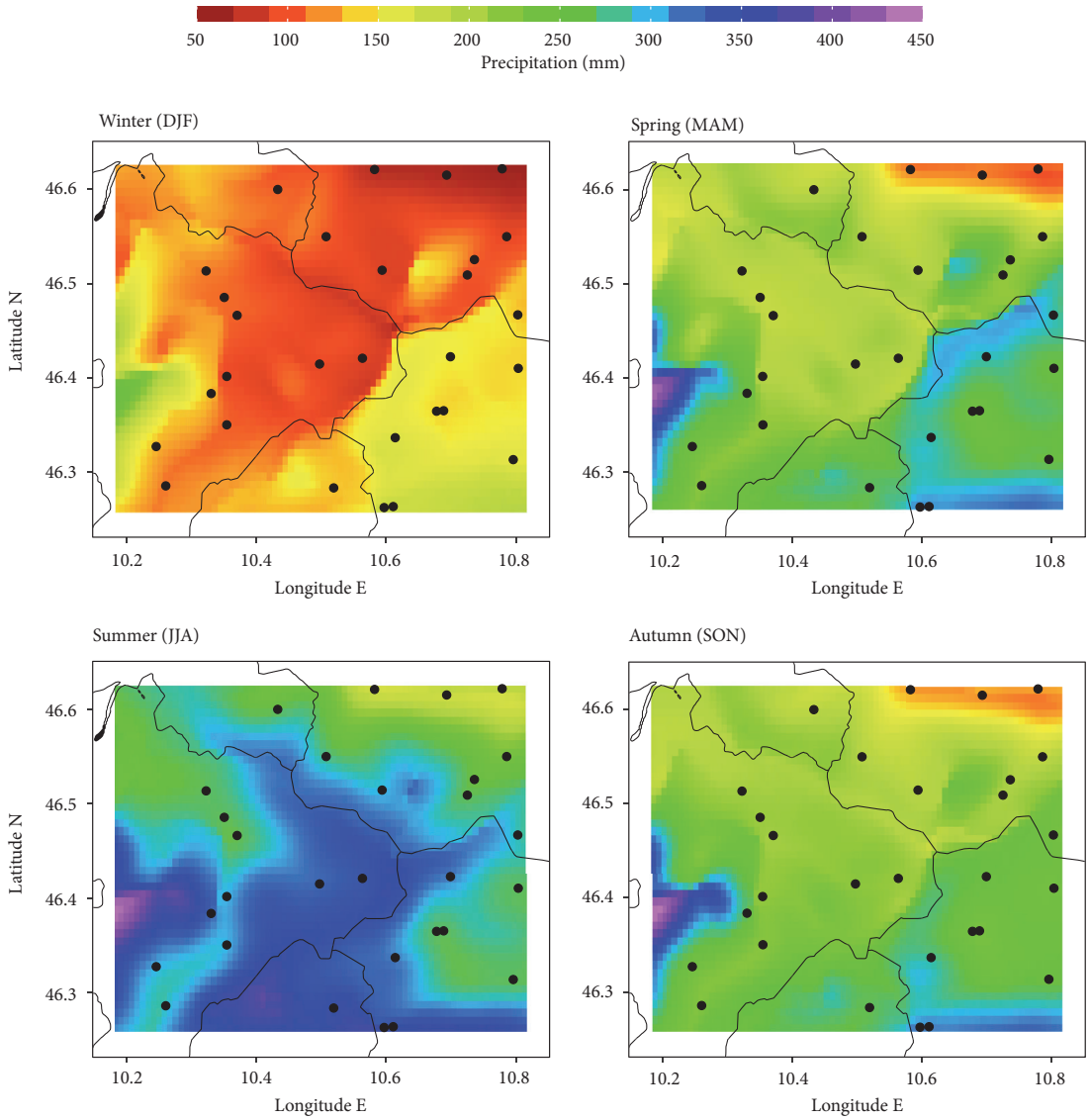

FIGURE 8: LWLR seasons climatologies. Black lines are borders of administrative regions as in Figure 6; black dots are weather stations.

400 mm/season, especially on Ortles-Cevedale chain). The relevant seasonality of the shadowing effect of the Ortles-Cevedale chain is a consequence of the fact that atmospheric circulation during rainy days exhibits strong differences along the year. Considering an average on the entire grid the amount of precipitation is 124 mm for the winter season, 228 mm for spring, 299 mm for summer, and 230 mm for autumn with minimum in February (40 mm) and maximum in August (107 mm). The lower winter values can be explained by the typical atmospheric circulation which is present in winter when precipitation occurs over North-Eastern Italy: it is characterised by southeasterly advection blowing wet air from the Adriatic Sea. Those wet air masses, blocked by the Central Alps, discharge their moisture over the South-Eastern side of Ortles-Cevedale chain. In the meanwhile, the North-Eastern part of the domain features analogous conditions due to the presence of Swiss Alps, which block frontal systems moving from the north and force them to lose humidity on the Switzerland side. The summer season, on a meteorological point of view, is characterised by a marked contrast between warm and moist air coming from the Po Plain and a relatively fresh continental air in the north side. Within this configuration, convective cells are promoted and convective precipitation is developed, in contrast with winter season where precipitation is generated mainly by orographic barring.

The same maps of Figures 7 and 8 were also built by the RK (not shown). While the annual cycle and the overall spatial distribution are comparable to those reconstructed by LWLR, pluviometric gradients were more smoothed out due to the global regression approach applied. LWLR local approach allows us to obtain a very detailed spatial distribution of the interpolated variable, and this could be particularly interesting for heterogeneous orographic regions, such as the considered one, but may lead to artefacts in case of poor station coverage over complex terrains. On the contrary, by considering elevation-precipitation relationship at a larger scale RK smooths more out the punctual climatological signal but provides more robust results in cases of low data availability.

3.2. Anomalies and Trends. After reconstructing the climatological field, the 1913–2015 gridded anomaly dataset was computed, and its suitability for the analysis of the time variability in precipitation series for a single grid point of interest or for the whole domain was assessed. Model ability in estimating anomaly series was evaluated in terms of BIAS, MAE, and RMSE values obtained by the reconstruction of station anomalies in LOO approach. Monthly MAE ranged between 0.28 for winter months and 0.14 for summer months.

The evolution of MAE during the entire period was also analysed in order to evaluate the reconstruction reliability. Due to the changes in quality and distribution of stations during the period, we obtained higher values of MAE, around 0.30, at the beginning and at the end of period, related to both lower data coverage and quality. In the middle part of our dataset MAE values were about 0.20. The evolution of BIAS was more uniform, especially since 1960, and approximately around zero.

The rather low LOO-errors of the simulated anomalies are due to rather high common variances between the anomaly records of the different stations: the scatter plot of correlation versus distance (Figure 11(b)) gives evidence of correlation coefficients over 0.7 even for the highest possible distances in the study area. The high common variance between the station anomaly records is also highlighted by Principal Component Analysis (PCA), the first eigenvector explaining more than 80% of the variance of the data set of inner area stations.

The high spatial coherence of the station anomaly records suggests investigating trends focusing only on an average inner area anomaly record. Short- and long-term trends for the whole domain were evaluated both on annual and seasonal scale. Long-term trends were computed using Theil-Sen test, and their statistical significance was proved by Mann-Kendall estimator [55–57]. Even if slope was slightly negative for winter and spring series, suggesting a centennial reduction of precipitation over the area, and positive for summer, autumn, and annual scale, the long-term trend has shown a statistical significance below the 95% confidence level. Short-term trends could be highlighted using an 11-year Gaussian window filter with a 3-year standard deviation (Figure 9) which smooths out random variability and evidences more clearly the decadal fluctuations. In order to better describe the complex time evolution of climatic signal, which cannot

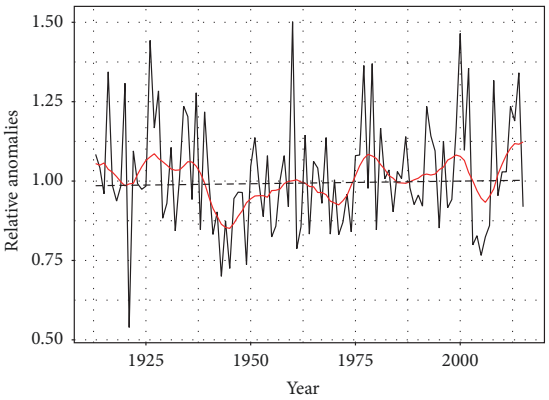

Figure 9: Annual rainfall anomalies. The red solid line represents the short-term trend (11-year Gaussian filter, 3-year standard deviation) and the dashed black line the Theil-Sen trend for the period 1913–2015.

be fully captured by the Theil-Sen slope on 1913–2015 period, a running-trend analysis was also performed on annual and monthly scale. Theil-Sen slopes and significance levels were estimated within windows whose widths range from 20 years up to the entire series length, running from the beginning to the end of the series [26]. Figure 10 shows the results of the running trend performed on annual precipitation anomalies, where colours represent the values of trend while size of pixels describes its significance. As already pointed out by long-term trend analysis, trend values and significance are mostly negligible over the longest time intervals, while more evident trends occurred considering shorter time windows. Years between 1925 and 1945 featured a clear negative trend, while annual precipitation amount was found to increase in both 1940–1960 and 1960–1980 periods.

3.3. Cross-Validation and Comparison at AWS1 Forni. The inner area stations reconstructed precipitation records—obtained superimposing the climatologies and the anomaly records—were also validated by means of the LOO approach. The MAE turned out to be about 15 mm/month. The lowest value was obtained for Silandro (ID 16 in Table 4) with less than 9 mm/month, while the highest value was found for Valdisotto (ID 26 in Table 4) with over 30 mm/month. The MAE errors of all inner area stations are represented in Figure 5.

We also computed the linear regression between the simulated and measured values at all inner area station sites (Figure 11(a)) and we found

$$y_{\text{sim}} = 0.96\, x_{\text{obs}} + 1.36\,\text{mm} \quad \left(R^2 = 0.84\right). \tag{4}$$

The highest common variance ($R^2 = 0.96$) between the reconstructed and the observed record was found for Rabbi Somrabbi (ID 9 in Table 4); the lowest ($R^2 = 0.66$) was found for Valdisotto.

A seasonal analysis was also performed using two macro-seasons: fall-winter (from September to February) and

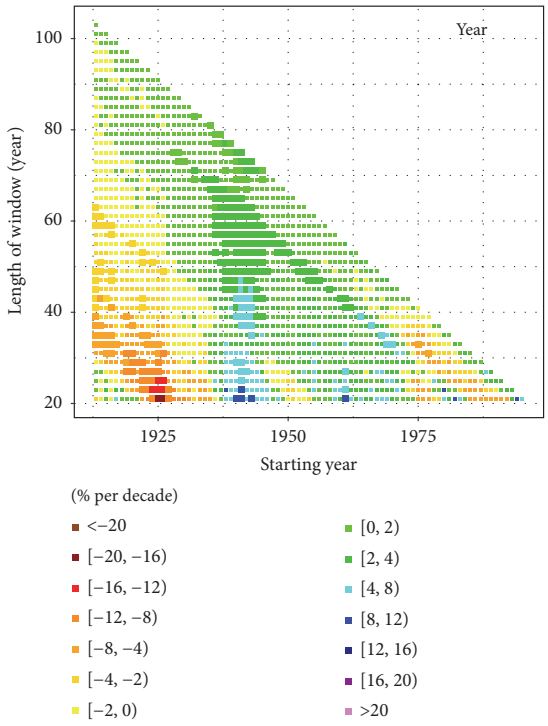

FIGURE 10: Running trends of annual precipitation anomalies. Colours represent the value of trends while pixel size refers to trend significance level. Trend evaluation is performed on windows whose widths range from 20 years up to the end of series starting from each year in x-axis.

spring-summer (from March to August). The results were similar to the all-data regression. They are reported below:

$$y_{\text{sim}} = 0.96\, x_{\text{obs}} + 1.23\ \text{mm} \quad \left(R^2 = 0.86\right) \text{ fall-winter}$$
$$y_{\text{sim}} = 0.96\, x_{\text{obs}} + 1.54\ \text{mm} \quad \left(R^2 = 0.80\right) \text{ spring-summer.} \tag{5}$$

Spring-summer reconstructions were generally less accurate than fall-winter ones and this could be related to the lower spatial coherence of precipitation in the summer period.

Finally, the model performances were further assessed by comparing modelled rainfall series and the measured values collected by the AWS on Forni Glacier, not included in the input dataset and located in a very remote and peculiar area. The precipitation series from the AWS on Forni Glacier counts 3563 days, from 1 October 2005 to 30 September 2016, and the number of complete months is 97 (no missing days). All trustworthy monthly data between 2005 and 2015 were compared with the modelled series of grid point 2101, which is the nearest to the average position of the AWS, and the computed errors are 33 mm/month for the MAE and 42 mm/month for the RMSE. The comparison of monthly data is shown in Figure 12(a). Considering an average of cell 2101 and its eight nearest neighbours, the errors raise to 38 mm/month of MAE and 48 mm/month of RMSE; this could be due to the inclusion of lower values far from the peak precipitation area of the watershed. These rather high errors depend indeed on the difficulties in capturing the exact position of the transition between the wet areas influenced by southeasterly currents and the very dry Forni Valley; however, they could also be strongly influenced by the high uncertainties of the SWE evaluation method. In fact, comparing the SWE values with the ones measured for two winter seasons by a snow pillow, a RMSE value of 45 mm w.e. was found [52]. In addition, even if the snow surface over the Forni Glacier seemed to be homogenous, a large spatial variability of snow depth was observed carrying out numerous measurements through a snow weighting tube (Enel-Valtecne©) around the AWS in February 2015 with a standard deviation of 29 cm of snow; see, for example, Senese et al. [52]. This is due to the high roughness of the ice surface that causes different snow accumulation rate over the glacier surface. Consequently, the error between the modelled series of grid point 2101 and the ones observed at the AWS1 Forni could be enhanced even by the high spatial variability of the snow depth in high mountain areas.

A deeper temporal analysis highlighted a high seasonality between simulated and measured precipitation (Figure 13); in fact the winter period is underestimated by the model; meanwhile the summer period is quite well depicted. The uncertainties follow the same pattern: December to February RMSE (DJF, meteorological winter) is 59 mm/month, and MAE is 54 mm/month, instead June to August RMSE (JJA, meteorological summer) is 24 mm/month, and MAE is 18 mm/month. The divergence between simulated and measured data could be explained by the peculiar precipitation distribution over the area. The orographic effect of Ortles-Cevedale chain is captured in the reconstructed field (Figure 8), but discrepancies with AWS1 Forni data suggest that the actual position of regime transition could not be fully resolved by the model together with the chosen DEM grid spacing. In fact, LWLR places the maximum rainfall (snowfall) over the crest, but especially for snowfalls, the wind could transport snowflakes for several kilometres, and it could create deposit far away from the precipitation maximum. Other possible sources of errors could be the conversion of snow height into water equivalent (see Section 2.4), for which a constant snow density was used, and the measurement itself of snow depth. In fact, these measurements could be overestimated due to the wind that acts on snow cover, changing its distribution and accumulating snow under the sonic ranger.

In order to investigate wind effect, the wind patterns were compared to days with big discrepancies between simulated and measured data over all the period. Wind records were retrieved from the anemometer installed at AWS1 Forni.

Daily differences were calculated as BIAS between observed and simulated rain, and minimum threshold of 20 mm was chosen. However, as shown in Figure 12(b), many points lay under 5 m/s of wind speed, and only eight

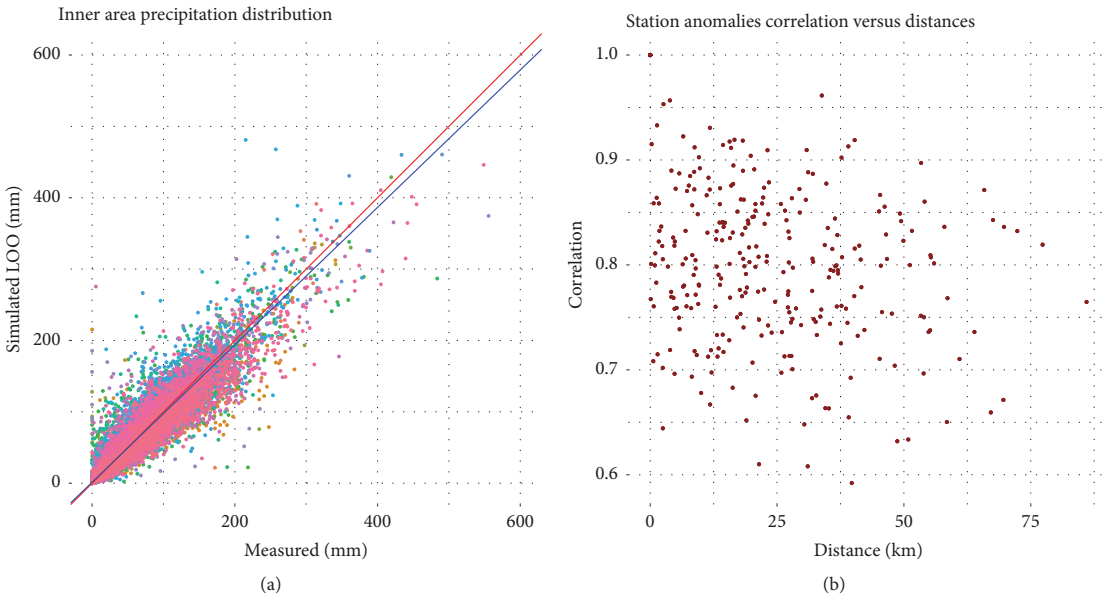

Figure 11: (a) Scatter plot of monthly simulated LOO precipitation versus the measured precipitations for each station. (b) Scatter plot of stations correlation versus distance.

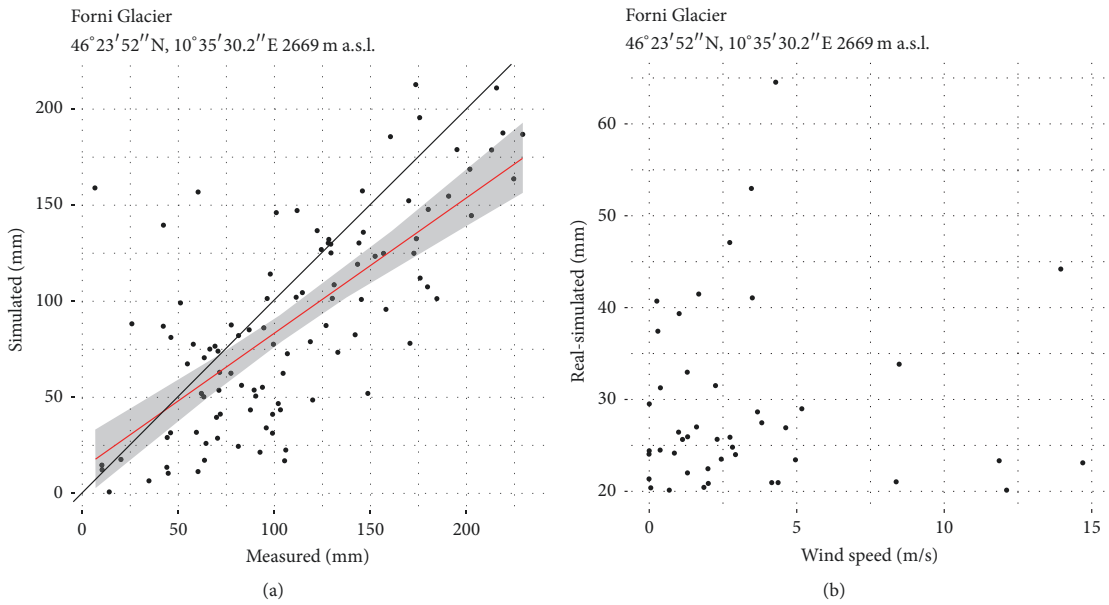

Figure 12: (a) Scatter plot of measured precipitations versus simulated precipitation at grid point 2101. The red line represents the linear regression of data, meanwhile the black line represents the skill score 1 : 1. (b) Scatter plot of monthly model discrepancies versus wind speed. Only discrepancies greater than 20 mm are shown.

underestimated days have a mean wind speed greater than 5 m/s. Only the 16% of points exceeded 5 m/s threshold, and they provided no evidence of a significant relationship occurring between precipitation underestimation and wind speed.

4. Conclusions

A statistical approach to project precipitation data on a high-resolution grid and based on anomaly method was applied over a complex mountainous terrain in the Central Italian

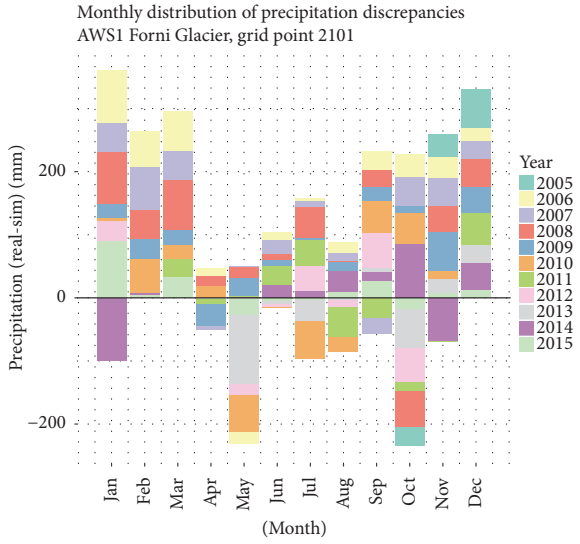

Figure 13: Monthly distribution of discrepancies between simulated and measured precipitation at Forni Glacier for the period 2005–2015. Positive values indicate an underestimation by the model.

Alps. The database counts up to 734 rain gauges, and the area where the climatologies were computed includes 30 quality-checked and homogenised stations. The 1961–1990 precipitation climatologies were spatialised over the study domain by LWLR starting from the assumption that a strong relationship occurs at local level between precipitation and orographic features. LOO model errors ranged between 4 mm and 10 mm, in terms of mean monthly MAE. These performances were compared to those obtained by global RK that had a worse winter and better summer performances, with a MAE ranging between 6 mm and 11 mm. Those results allowed us to assess the goodness of the proposed method for the reconstruction of monthly precipitation time series at a specific point of the high-resolution grid. We also appreciated the different spatialisation properties of the two approaches. In fact, LWLR can better represent orographic features than RK that, instead, smooths out such details; on the contrary, the first one is more susceptible to outliers and/or overestimations (especially during summer) than the second one, in particular over those areas less covered by stations.

Regarding the reconstruction ability of 1913–2015 anomalies, monthly MAE along the whole period ranged from 0.14 to 0.28.

We verified the absence of statistically significant long-term trends for the considered area; however short-term trends were present. In particular, three positive tendencies were found (1940–1960, 1960–1980, and the last decade) and one negative during the 1920s. The spatialised climatologies revealed a very dry area in the Northern part of our study domain, corresponding to Venosta Valley; by observing seasonal climatologies, a two-pronged behaviour of precipitation for Upper Valtellina was found: a very dry winter, where mountain barring mostly acts, opposed to a very wet summer, when convective precipitation is prevalent.

The cross-validation performed for the 30 inner area stations showed that the key factor for reconstructing unbiased precipitation fields is correctly capturing the spatial variability of the monthly precipitation normals. It is therefore very important to exploit the existing data as much as possible in order to have the best spatial resolution of these data and to recover all possible records even if they cover only rather short periods. On the contrary, the high spatial coherence of the anomalies allows capturing the temporal evolution of precipitation over the investigated area with a much lower number of stations. In this case, however, homogeneity is a key issue and the availability of a reasonable number of records turns out to be very important for homogenisation procedures.

The spatialised fields were also compared with the in situ measurements of the supraglacial automatic weather station located since 2005 at Forni Glacier, which was not included into the model database. We have obtained that the model had well depicted the main pluviometric features of the area, but it faced greater difficulties during winter, when the Upper Valtellina is drier with respect to Sole Valley in Trentino. The observed uncertainties could be explained with an underestimation of the model in this watershed zone for a misplacing of the transition zone, or with an overestimation of winter snow measurements and/or distribution.

In conclusion the applied approach reveals many advantageous aspects; in fact, from 30 in situ measurements it allowed reconstructing at high resolution the precipitation signal over an area of about 2000 km^2. Therefore, this approach could be particularly helpful if applied to a rather complex mountainous terrain, where direct measurements are poor and quite short, in order to understand and study fragile environments in the frame of climate change. However further researches are necessary to better investigate the role of precipitation, temperature, and wind on the biological component of glacier surfaces and forelands, especially if we consider that they represent a changing habitat for life.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors wish to thank the Regional Agency for Environment Protection of Lombardy (data series is available on the web at http://www.arpalombardia.it/siti/arpalombardia/meteo/richiesta-dati-misurati/Pagine/RichiestaDatiMisurati.aspx), the Geological Monitoring Service of Lombardy, the Weather Service of Trento Province (https://www.meteotrentino.it/dati-meteo/info-dati.aspx?ID=3), the Weather Service of Bolzano Province (http://dati.retecivica.bz.it/it/dataset/misure-meteo-e-idrografiche), and the Federal Weather Service of Switzerland (https://gate.meteoswiss.ch/idaweb/login.do) for supplying weather data. The AWS1 Forni has been developed by UNIMI, is hosted in the Stelvio

Park-ERSAF, and is presently managed by UNIMI DESP with the financial support of Sanpellegrino SpA (Levissima). The acquired data are not only used for the scientific research of the UNIMI staff but also shared with the WMO (World Meteorological Organization) in the framework of some international programs (i.e., SPICE and CRYONET).

References

[1] I. Bollati, M. Pellegrini, E. Reynard, and M. Pelfini, "Water driven processes and landforms evolution rates in mountain geomorphosites: examples from Swiss Alps," *Catena*, vol. 158, pp. 321–339, 2017.

[2] P. Whelan and A. J. Bach, "Retreating Glaciers, Incipient Soils, Emerging Forests: 100 Years of Landscape Change on Mount Baker, Washington, USA," *Annals of the American Association of Geographers*, vol. 107, no. 2, pp. 336–349, 2017.

[3] R. Tichavsky, K. Silham, and R. Tolasz, "Tree ring based chronology of hydro-geomorphic processes as a fundament for identification of hydro-meteorological triggers in the Hruby Jesenik MounMount (Central Europe)," *Science of the Total Environment*, vol. 579, pp. 1904–1917, 2017.

[4] M. Stoffel, B. Wyzga, and R. Marston, "Floods in mountain environments: A synthesis," *Geomorphology*, vol. 272, pp. 1–9, 2016.

[5] C. Smiraglia, G. Diolaiuti, and R. Azzoni, Eds., EvK2CNR, Bergamo, p. 400, 2015.

[6] C. D'Agata, D. Bocchiola, D. Maragno, C. Smiraglia, and G. A. Diolaiuti, "Glacier shrinkage driven by climate change during half a century (1954-2007) in the Ortles-Cevedale group (Stelvio National Park, Lombardy, Italian Alps)," *Theoretical and Applied Climatology*, vol. 116, no. 1-2, pp. 169–190, 2014.

[7] A. Mondoni, S. Pedrini, G. Bernareggi et al., "Climate warming could increase recruitment success in glacier foreland plants," *Annals of Botany*, vol. 116, no. 6, pp. 907–916, 2015.

[8] V. Garavaglia, M. Pelfini, and I. Bollati, "The influence of climate change on glacier geomorphosites: The case of two Italian glaciers (Miage Glacier, Forni Glacier)investigated through dendrochronology," *Geomorphologie: Relief, Processus, Environnement*, no. 2, pp. 153–164, 2010.

[9] M. Pelfini, G. Diolaiuti, G. Leonelli et al., "The influence of glacier surface processes on the short-term evolution of supraglacial tree vegetation: The case study of the Miage Glacier, Italian Alps," *The Holocene*, vol. 22, no. 8, pp. 847–856, 2012.

[10] M. Pelfini and G. Leonelli, "First results of the participatory approach for monitoring supraglacial vegetation in Italy," *Geografia Fisica e Dinamica Quaternaria*, vol. 37, no. 1, pp. 23–27, 2014.

[11] I. Bollati, G. Leonelli, L. Vezzola, and M. Pelfini, "The role of Ecological Value in Geomorphosite assessment for the Debris-Covered Miage Glacier (Western Italian Alps) based on a review of 2.5 centuries of scientific study," *Geoheritage*, vol. 7, no. 2, pp. 119–135, 2015.

[12] M. Gobbi, M. Isaia, and F. De Bernardi, "Arthstudy colonisation of a debris-covered glacier," *The Holocene*, vol. 21, no. 2, pp. 343–349, 2011.

[13] R. S. Azzoni, A. Franzetti, D. Fontaneto, A. Zullini, and R. Ambrosini, "Nematodes and rotifers on two Alpine debris-covered glaciers," *Italian Journal of Zoology*, vol. 82, no. 4, pp. 616–623, 2015.

[14] M. Pelfini and M. Gobbi, "Enhancement of the ecological value of Forni Glacier (Central Alps) as a possible geomorphosite: New data from arthropod communities," *Geografia Fisica e Dinamica Quaternaria*, vol. 28, no. 2, pp. 211–217, 2005.

[15] J. Hjort and M. Luoto, "Geodiversity of high-latitude landscapes in norhern Finland," *Geomorphology*, vol. 115, no. 1-2, pp. 109–116, 2010.

[16] V. Garavaglia, M. Pelfini, A. Bini, L. Arzuffi, and M. Bozzoni, "Recent evolution of debris-flow fans in thecentral Swiss Alps and associated risk assessment: Two examples in Roseg Valley," *Physical Geography*, vol. 30, no. 2, pp. 105–129, 2009.

[17] H. Purdie, C. Gomez, and S. Espiner, "Glacier recession and the changing rockfall hazard: Implications for glacier tourism," *New Zealand Geographer*, vol. 71, no. 3, pp. 189–202, 2015.

[18] I. Bollati, B. Crosa Lenz, E. Zanoletti, and M. Pelfini, "Geomorphological mapping for the valorization of the alpine environment. A methodological proposal tested in the Loana Valley (Sesia Val Grande Geopark, Western Italian Alps)," *Journal of Mountain Science*, vol. 14, no. 6, pp. 1023–1038, 2017.

[19] P. Palma, M. Oliva, and C. Garcia-Hernandez, "Spatial characterisation of glacial and periglacial landforms in the highlands of Sierra Nevada (Spain)," *Science of the Total Environment*, vol. 584, pp. 1256–1267, 2017.

[20] O. Peyron, N. Combourieu-Nebout, D. Brayshaw et al., "Precipitation changes in the Mediterranean basin during the Holocene from terrestrial and marine pollen records: A model-data comparison," *Climate of the Past*, vol. 13, no. 3, pp. 249–265, 2017.

[21] M. Pelfini, G. Leonelli, L. Trombino et al., "New data on glacier fluctuations during the climatic transition at ~4,000 cal. year BP from a buried log in the Forni Glacier forefield (Italian Alps)," *Rendiconti Lincei*, vol. 25, no. 4, pp. 427–437, 2014.

[22] O. Solomina, R. Bradley, and V. Jomelli, "Glacier fluctuations during the past, 2000, years," *Quaternary science reviews149*, pp. 61–90, 2016.

[23] K. Nicolussi, M. Kaufmann, T. M. Melvin, J. van der Plicht, P. Schießling, and A. Thurner, "A 9111 year long conifer tree-ring chronology for the European Alps: A base for environmental and climatic investigations," *The Holocene*, vol. 19, no. 6, pp. 909–920, 2009.

[24] M. Opala, K. Migala, and P. Owczarek, "Two centuries-long dendroclimatic reconstruction based on Low Arctic Betula pubescens from Tromso Region, Northern Norway," *Polish Polar Research*, vol. 37, no. 4, pp. 457–476, 2016.

[25] G. Leonelli, A. Coppola, C. Baroni et al., "Multispecies dendroclimatic reconstructions of summer temperature in the European Alps enhanced by trees highly sensitive to temperature," *Climatic Change*, vol. 137, no. 1-2, pp. 275–291, 2016.

[26] M. Brunetti, M. Maugeri, F. Monti, and T. Nanni, "Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series," *International Journal of Climatology*, vol. 26, no. 3, pp. 345–381, 2006.

[27] M. Citterio, G. Diolaiuti, C. Smiraglia, G. Verza, and E. Meraldi, "Initial results from the Automatic Weather Station (AWS) on the ablation tongue of Forni Glacier (Upper Valtellina, Italy)," *Geografia Fisica e Dinamica Quaternaria*, vol. 30, no. 2, pp. 141–151, 2007.

[28] M. Brunetti, G. Lentini, M. Maugeri, T. Nanni, C. Simolo, and J. Spinoni, "Projecting North Eastern Italy temperature and precipitation secular records onto a high-resolution grid," *Physics and Chemistry of the Earth*, vol. 40-41, pp. 9–22, 2012.

[29] S. Shi, J. Li, J. Shi, Y. Zhao, and G. Huang, "Three centuries of winter temperature change on the southeastern Tibetan Plateau

and its relationship with the Atlantic Multidecadal Oscillation," *Climate Dynamics*, vol. 49, no. 4, pp. 1305–1319, 2017.

[30] F. Marra, E. Destro, E. I. Nikolopoulos et al., "Impact of rainfall spatial aggregation on the identification of debris flow occurrence thresholds," *Hydrology and Earth System Sciences*, vol. 21, no. 9, pp. 4525–4532, 2017.

[31] C. Papagiannopoulou, D. G. Miralles, W. A. Dorigo, N. E. C. Verhoest, M. Depoorter, and W. Waegeman, "Thresholds anomalies caused by antecedent precipitation in most of the world," *Environmental Research Letters*, vol. 12, no. 7, Article ID 074016, 2017.

[32] M. Pelfini, P. Brandolini, A. Carton et al., *Geomorphosites, Assessment, mapping and management*, Chapter Geo-tourist and geomorphological risk/impact, Pfeil Verlag, München, Germany, 2009.

[33] M. Beniston, "Climatic change in mountain regions: a review of possible impacts," *Climatic Change*, vol. 59, no. 1-2, pp. 5–31, 2003.

[34] M. Pelfini, "Morfogenesi glaciale in alcune valli del Gruppo Ortles-Cevedale," *Memorie - Società Geologica Italiana*, pp. 381–390, 1989.

[35] M. Pelfini, *Le fluttuazioni glaciali oloceniche nel Gruppo Ortles-Cevedale (settore lombardo)*, Università degli studi di Milano, 1992.

[36] R. Azzoni, D. Fugazza, A. Zerboni et al., "High-resolution remote sensing data to reconstruct the recent evolution of supraglacial debris in the Central Alps (Stelvio Park, Italy)," *Progress in Physical Geography*, vol. 42, no. 1, pp. 3–23, 2018.

[37] P. Forster, V. Ramaswamy, P. Artaxo et al., "Changes in Atmospheric Constituents and in Radiative Forcing," in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth assessment report of IPCC*, S. Solomon, D. Qin, M. Manning et al., Eds., Cambridge University Press, 2007.

[38] A. Masseroli, G. Leonelli, I. Bollati, L. Trombino, and M. Pelfini, "The influence of geomorphological processes on the treeline position in Upper Valtellina (Central Italian Alps)," *Geografia Fisica e Dinamica Quaternaria*, vol. 39, no. 2, pp. 171–182, 2016.

[39] T. D. Mitchell and P. D. Jones, "An improved method of constructing a database of monthly climate observations and associated high-resolution grids," in *Proceedings of the An improved method of constructing a database of monthly climate observations and associated high-resolution grids*, vol. 25, pp. 693–712, 2005.

[40] A. Crespi, M. Brunetti, G. Lentini, and M. Maugeri, "1961-1990 high-resolution monthly precipitation climatologies for Italy," *International Journal of Climatology*, 2017.

[41] I. Auer, R. Böhm, A. Jurkovic et al., "HISTALP - Historical instrumental climatological surface time series of the Greater Alpine Region," *International Journal of Climatology*, vol. 27, no. 1, pp. 17–46, 2007.

[42] J. M. Craddock, "Methods of comparing annual rainfall records for climatic purposes," *Weather*, vol. 34, no. 9, pp. 332–346, 1979.

[43] P. Jones, S. Raper, R. Bradley, H. Diaz, P. Kelly, and T. Wigley, "Northern Hemisphere Surface air temperature variations: 1851-1984," *Journal of Climate and Applied Meteorology*, vol. 25, pp. 161–179, 1986.

[44] E. Aguilar, I. Auer, M. Brunet, T. C. Peterson, and J. Wieringa, *Guidelines on climate metadata and homogenization*, World Meteorological Organisation, Geneva, Switzerland, 2003.

[45] USGS., "GTOPO30 Documentation," Tech. Rep., United States Geological Survey, 1996.

[46] C. Frei and C. Schär, "A precipitation climatology of the Alps from high-resolution rain-gauge observations," *International Journal of Climatology*, vol. 18, no. 8, pp. 873–900, 1998.

[47] C. Daly, "Guidelines for assessing the suitability of spatial climate data sets," *International Journal of Climatology*, vol. 26, no. 6, pp. 707–721, 2006.

[48] T. Hengl, G. B. M. Hauvelink, and D. G. Rossiter, "About regression-kriging: from equations to case studies," *Computers & Geosciences*, vol. 33, no. 10, pp. 1301–1315, 2007.

[49] C. Daly, M. Halbleib, J. I. Smith et al., "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States," *International Journal of Climatology*, vol. 28, no. 15, pp. 2031–2064, 2008.

[50] M. New, M. Hulme, and P. Jones, "Representing twentieth-century space-time climate variability. Part II: development of 1901–96 monthly grids of terrestrial surface climate," *Journal of Climate*, vol. 13, no. 13, pp. 2217–2238, 2000.

[51] A. Senese, M. Maugeri, E. Vuillermoz, C. Smiraglia, and G. Diolaiuti, "Using daily air temperature thresholds to evaluate snow melting occurrence and amount on Alpine glaciers by T-index models: The case study of the Forni Glacier (Italy)," *The Cryosphere*, vol. 8, no. 5, pp. 1921–1933, 2014.

[52] A. Senese, M. Maugeri, E. Meraldi et al., "Snow data intercomparison on remote and glacierized high elevation areas (Forni Glacier, Italy)," *The Cryosphere Discussion*, https://doi.org/10.5194/tc-2017-124.

[53] R. Azzoni, D. Fugazza, M. Zennaro et al., "Recent structural evolution of Forni Glacier tongue (Ortles-Cevedale Group, Central Italian Alps)," *Journal of Maps*, vol. 13, no. 2, pp. 870–878, 2017.

[54] D. Fugazza, M. Scaioni, M. Corti et al., "Combination of UAV and terrestrial photogrammetry to assess rapid glacier evolution and conditions of glacier hazards," *Natural Hazards and Earth System Sciences*, https://doi.org/10.5194/nhess-2017-198.

[55] H. Theil, "A rank-invariant method of linear and polynomial regression analysis," in *Proceedings of Koninklijke Nederlandse Akademie von Wetenschappen A*, vol. 53, pp. 1397–1412, 1950.

[56] M. G. Kendall, *Rank Correlation Methods*, Griffin, London, UK, 1970.

[57] P. K. Sen, "Estimates of the regression coefficient based on Kendall's tau," *Journal of the American Statistical Association*, vol. 63, pp. 1379–1389, 1968.

Chapter 4

Reconstruction of a multi-secular record of monthly total precipitation over the Adda river basin (Central Italian Alps): trends, variability and runoff comparison

The chapter describes the results of the research activities focusing on the reconstruction of the long-term evolution of monthly precipitation onto a high-resolution grid covering the upper Adda river basin (4508 km^2), which is one of the main water Alpine reservoirs in Northern Italy. The activities were managed in collaboration with the Department of Civil, Environmental, Architectural Engineering and Mathematics of Università degli Studi di Brescia (Italy) which dealed with the measurements of basin runoff.

A database of monthly precipitation spanning more than two centuries (since '700 to present) was set up for a wide area centred on the basin by exploiting all the available data sources and trying to maximise the data coverage for the high-elevation regions. Some of the most remote series were retrieved from hardcopy yearbooks and digitisation projects while the merging of mechanical data with the most recent automatic records provided by the regional and subregional services allowed to update most series to present. Quality-check and homogenisation procedures were performed on the resulting database, which includes more than 30 series starting before the 1850s. By applying the anomaly method, a 30-arc second resolution grid of 1800-2016 monthly precipitation series over the domain was obtained. It allowed to provide the areal precipitation record of upper Adda river basin, to evaluate its variability and trends and to perform the comparison with the recontructed secular record of annual runoff

from the catchment, which is one of the longest runoff series in Italy. The comparison highlighted a significant long-term decrease in runoff coefficient which is likely to be determined by the increase of evapotranspiration rates driven by both warming and changes in land use and coverage.
The evolution of the reconstruction accuracy over the whole study period was evaluated and rather low errors were observed even for the earliest decades thanks to the large spatial coherence of anomalies which allows to provide a reliable precipitation field even in the periods of lower data availability.
An interesting application of the obtained dataset is the reconstruction and analysis of the spatial structure of specific past events of intense precipitation occurred over the basin, therefore some of these episodes are reported as case studies.
More specifically, the chapter is composed by two sections:

- a preliminary study on the construction of a gridded dataset of 1845-2016 monthly precipitation records over the upper Adda basin in order to compare the total areal precipitation evolution with the available runoff series spanning the same period. The work provides a first evaluation of the suitability of the retrieved database to project the secular precipitation signal onto the high-resolution grid and a preliminary assessment of the annual long-term trends in both precipitation and runoff. In addition, the extreme precipitation event of November 2002 is reconstructed and it is used to evaluate the sensitivity of the depicted field to the available station distribution.
  The section is based on the conference proceeding published in open-access version on the peer-reviewed journal *Advances in Science and Research*:

  Crespi, A., Brunetti, M., Maugeri, M., Ranzi, R., and Tomirotti, M. (2018). 1845–2016 gridded dataset of monthly precipitation over the upper Adda river basin: a comparison with runoff series, Advance in Science and Research, 15, 173-181. doi:10.5194/asr-15-173-2018

- a more comprehensive study in which the 30-arc second resolution dataset of 1800-2016 monthly precipitation series over the upper Adda basin is presented and a detailed evaluation on the reconstruction accuracy in relation to the variation in data coverage is provided. The robustness of both the modelled 1800-2016 records at station sites and the total areal monthly precipitation series of the catchement is discussed. Long and short-term trends and variability of annual and seasonal precipitation are analysed and the comparison with 1845-2016 runoff in terms of hydrological year is presented. The evolution of annual runoff coefficients are discussed and the role of evapontranspiration in the observed long-term change of hydrological cycle is estimated. The section is based on the

manuscript submitted to the peer-reviewed journal *Water Resources Research*. It is a part of a more extensive study performed on the evolution of water resources for the Adda basin. The reconstructed areal precipitation series was in fact used to analyse more in detail the runoff trends and variability, the response time of the basin and the natural and anthropogenic forcings influencing the amount of water disposal. The manuscript describing these results together with the analysis of runoff variability at daily scale is going to be submitted to *Water Resources Research*.

The research activities were also presented at international and national conferences, published as conference proceedings and reported on a national journal in Italian language in order to widespread the results on a larger public:

- Crespi, A., Brunetti, M., Mattea, E., Maugeri, M., Ranzi, R., and Tomirotti, M. (2017). 1845-2016 gridded database of monthly precipitation series for Adda basin (Italy): long-term variability and trend in rainfall-runoff comparison. EMS Annual Meeting: European Conference for Applied Meteorology and Climatology, Dublin (Ireland), Vol. 14, EMS2017-274.

- Ranzi, R., Goatelli, F., Castioni, C., Tomirotti, M., Crespi, A., Mattea, E., Brunetti, M., and Maugeri, M. (2017). Long term statistics (1845-2014) of daily runoff maxima, monthly rainfall and runoff in the Adda basin (Italian Alps) under natural and anthropogenic changes. Geophysical Research Abstracts, Wien (Austria), Vol. 19, EGU2017-15469.

- Ranzi, R., Tomirotti, M., Michailidi, E.M., Brunetti, M., Crespi, A., and Maugeri, M. (2018). Detection of rainfall and runoff trends of the Adda river in Lecco (1845-2014) at different time scales. 13th International Conference on Hydroinformatics, Palermo (Italy), pp. 7.

- Ranzi, R., Bacchi, B., Tomirotti, M., Castioni, C., Brunetti, M., Crespi, A., and Maugeri, M. (2018). Analisi delle tendenze di lungo termine nel regime degli afflussi meteorici e dei deflussi dell'Adda a Lecco (1845-2014), L'Acqua, 2, 51-60.

4.1 A preliminary study on the secular total monthly precipitation record recostruction over the upper Adda river basin

Adv. Sci. Res., 15, 173–181, 2018
https://doi.org/10.5194/asr-15-173-2018

Advances in Science & Research
Open Access Proceedings

1845–2016 gridded dataset of monthly precipitation over the upper Adda river basin: a comparison with runoff series

Alice Crespi[1], **Michele Brunetti**[2], **Maurizio Maugeri**[1,2], **Roberto Ranzi**[3], **and Massimo Tomirotti**[3]

[1]Department of Environmental Science and Policy, Università degli Studi di Milano, Milan, 20133, Italy
[2]Institute of Atmospheric Sciences and Climate, ISAC-CNR, Bologna, 40129, Italy
[3]DICATAM, Università degli Studi di Brescia, Brescia, 25123, Italy

Correspondence: Alice Crespi (alice.crespi@unimi.it)

Received: 17 January 2018 – Revised: 9 April 2018 – Accepted: 13 June 2018 – Published: 1 August 2018

Abstract. A new high-resolution gridded dataset of 1845–2016 monthly precipitation series for the upper Adda river basin was computed starting from a network of high-quality and homogenised station records covering Adda basin and neighbouring areas and spanning more than two centuries. The long-term signal was reconstructed by a procedure based on the anomaly method and consisting in the superimposition of two fields which were computed independently: 1961–1990 monthly climatologies and gridded anomalies. Model accuracy was evaluated by means of station series reconstruction in leave-one-out approach and monthly relative mean absolute errors were found to range between 14 % in summer and 24 % in winter. Except for the period before the 1870s when station coverage is rather low, reconstruction errors are quite stable. The 1845–2016 monthly areal precipitation series integrated over Adda basin was finally computed. The robustness of this series was evaluated and it was investigated for long-term trend. While no significant trend emerged for precipitation, the analysis performed on 1845–2016 annual runoff values recorded at Lake Como outlet highlighted a negative trend. Runoff decrease is supposed to be mostly due to an increasing role of evapotranspiration linked to temperature increase, which is only partially compensated by the increase in glacier melting rate. In order to test the applicability of the gridded database for the reconstruction of extreme past events, the episode with the highest precipitation in Adda basin series (November 2002) was considered and the corresponding gridded fields of monthly anomalies and precipitation values were evaluated both with actual station density and with station densities corresponding to 1922 and 1882. Even considering 1882 station density, the main spatial patterns are well depicted proving the suitability of anomaly method to deal also with sparse station networks.

1 Introduction

The availability of long-term observational series of meteorological variables is crucial to study the evolution of past climate and to detect climate variability and change. Projecting climate normals onto high-resolution grids covering specific regions of interest allows to complete the information about temporal variability deriving from the secular sparse station networks and to extract long-term series in absolute values for any point of the study area. These gridded datasets provide a very useful tool to investigate more accurately possible changes in the climatic variables, to perform areal trend analyses and to support climate change impact studies. In particular, at basin scale gridded precipitation can be combined with hydrological data to derive relevant information about interconnections occurring among climate, natural environment and human activities as well as about their changes over time.

Several observational gridded precipitation datasets covering the pan-Alpine region have been recently produced at different spatial and time resolutions, such as the 10 min resolution monthly dataset proposed by Efthymiadis at al. (2006) for the 1800–2003 period, the 1970–2008 daily dataset at 5 km resolution described in Isotta et al. (2014) and the most recent 1901–2008 monthly precipitation dataset presented by

Masson and Frei (2016) with a grid-spacing of 5 km. One of the methodologies which have been currently applied to interpolate station network observations onto high-resolution grid is based on the so-called "anomaly method" (e.g. New et al., 2000; Mitchell and Jones, 2005) whose main assumption is that any climatic signal can be described by the superimposition of a spatial field of normal values for a certain reference period (i.e. the climatologies), and a spatio-temporal field of anomalies from them. The two fields can be computed separately thanks to their different features. Climatologies are strictly linked to orography and require a dense data coverage as well as interpolation models taking into account the spatial variability of the rainfall-orography relationship, while anomalies, which exhibit higher spatial coherence, can be computed starting from a less dense database but providing long and homogeneous time series properly covering the entire study period. Homogeneity and good quality of data are mandatory to avoid non-climatic signals to affect the long-term anomaly reconstruction.

We are currently working on the application of this methodology to produce a gridded monthly precipitation dataset for an area centred on the upper part of Adda river basin and including neighbouring regions (8.8–10.7° E and 45.6–46.7° N, the blue box in Fig. 1). Adda river is one of the main left-side tributaries of Po river, its upper catchment covers about 4508 km^2 and it is located in Northern Italy, and to a lesser extent in Switzerland, over an area characterised by a great orographic complexity. The upper part of Adda river basin is a very interesting area both because it is surrounded by a relevant number of secular precipitation records and because it has an uninterrupted runoff daily record starting from 1845, which was recently recovered (Ranzi et al., 2017). Our goal is to obtain a secular precipitation record for any point of a 30 arcsec resolution grid ($\sim$ 1 km) covering this region which could be investigated in trend analysis and compared to hydrological data. The resulting dataset represents an update and improvement of already existing monthly fields for the region since it is constructed by considering a rain-gauge database including newly digitised data relative to Adda basin, or close sites, and most series updated to the most recent period. In this paper, we present the data and the applied methods and we show some preliminary analyses highlighting the potential of the gridded dataset of secular precipitation records.

2 Materials and Methods

2.1 Data and Study area

The database we used to compute the secular monthly precipitation records for the upper part of Adda river basin is composed by more than 2500 precipitation series located over a wide region extending over Central-Northern Italy (latitude above 43.7° N) and surrounding countries as shown in Fig. 1. This very large area is considered to fully exploit available precipitation records in the early period when there are no data from stations located within the blue box region (Fig. 1). Data were retrieved from both daily and monthly archives of several national and international sources. The main data providers were Regional Agency for Environmental Protection (ARPA) and Geological Monitoring Service (CMG) in Lombardy, MeteoTrentino for Autonomous Province of Trento, Autonomous Province of Bozen, MeteoSwiss and the former National Hydrographic Service for the most ancient data. Moreover, these data were integrated by those included in databases available from previous works, such as Brunetti et al. (2006), Auer et al. (2007), Brugnara et al. (2012) and Crespi et al. (2018). Data collection was focused both on maximising the spatial coverage over the study domain and closer areas, especially where strong orographic gradients occur, and on increasing the availability of long and homogeneous precipitation series over the entire considered period. At this aim, precipitation data collected by the recent automatic weather station networks were recovered and integrated to those from the mechanical ones, and the most ancient data for some Italian records were digitised from hardcopy historical archives (Servizio Idrografico, 1920, 1957). Except for the series which were retrieved from already homogenised and quality-checked archives, all monthly precipitation series underwent a quality-check procedure (for details see Crespi et al., 2018) aiming at removing spurious records or digitisation oversights and at detecting the presence of relevant inhomogeneous periods. In addition, the longest records showing possible breaks during quality-check activity were considered for homogenisation by applying a procedure based on Craddock test (Craddock, 1979). 125 monthly series, 57 of which located inside the study domain, of an average length of 85 years were homogenised by identifying more than 300 breaks. The applied corrections lead to an overall decrease in reconstruction errors for homogenised series of about 10 % in comparison with those obtained by considering original versions.

The resulting database covers more than two centuries from 1750 to 2016 and 338 series are located inside the study domain. Here we consider however only the 1845–2016 period. While spatial distribution of stations over Adda basin is relatively dense (about one station per 44 km^2), the corresponding data availability is highly variable over the study period and no stations have data before 1873 (dashed line in Fig. 2). However, the inclusion of outer stations with available data in the most ancient years allows to extend the computation of precipitation fields over Adda basin back in the past and to improve the robustness of reconstructed values for the earliest period. These stations, especially the ones at the greatest distance from the basin, do not influence the reconstruction over decades of dense data coverage on the study domain and they mostly contribute for the first decades of the study period. In fact, the availability of records in the past is very good if we focus on the entire area in Fig. 1 (solid line in Fig. 2). From 1750 to the first half of the 19th cen-

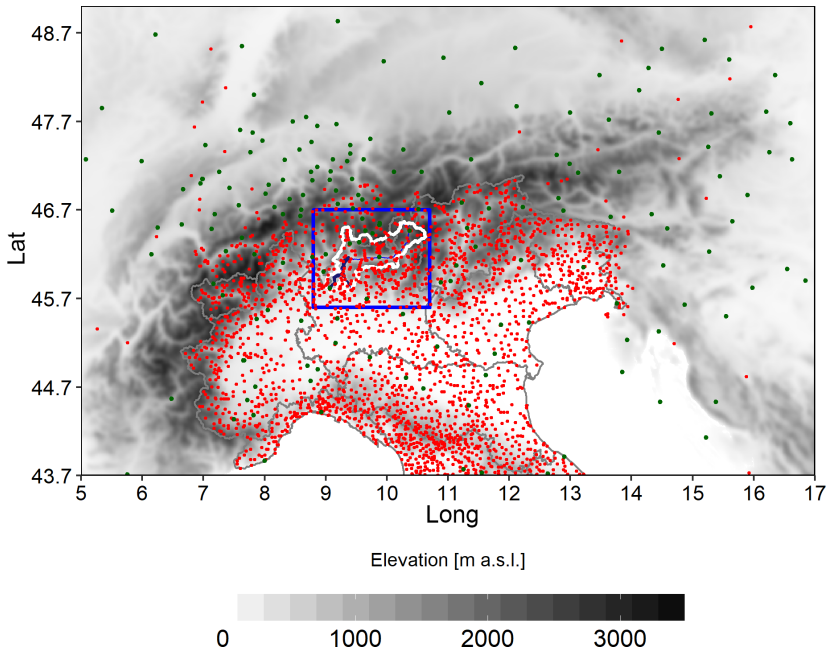

Figure 1. Distribution of the available monthly precipitation series. Study domain is contained in the blue square where the upper Adda river basin area is bordered in white. Green dots represent the stations with the longest records (more than 120 years of data).

tury the number of active stations gradually increases reaching about 50 observation sites. Since the second half of the 19th century this increase is more evident until 1921 when more than 600 sites were operating thanks to the contribution of the new meteorological network managed by Italian Hydrographic Service. After the negative consequences of the Second World War on the management of station sites, data availability reaches its maximum from 1951 to the last decades of the 20th century with more than 1500 series. Over recent years, the decrease in station availability is largely due to the fragmentation of Italian meteorological network under regional and sub-regional managements together with a general disposal of mechanical stations which in some cases were not substituted by the automatic ones. The database contains 9 series with more than 200 years of data, 31 series with more than 160 years and 120 series more than 120 years long, 14 out of these located within the study domain (blue box in Fig. 1). However, no one of these 14 stations has data before 1861. Looking at the variation in mean inter-station distance (i.e. the distance between each station and the closest one) over the whole period, we could retrieve further information about the development in data coverage and the spatial scales that can be effectively resolved in reconstruction analysis. The mean inter-station distance decreases gradually from $\sim$ 75 km (30 stations over the whole area) in 1845 to $\sim$ 25 km in 1900 when 276 stations are available (25 out of them located inside the study area). Its value reduces further during the 20th century and it is about 8 km during the period of maximum station availability (1951–2000).

As regards runoff, 1845–2016 monthly record for the station of Fortilizio at Lake Como outlet was reconstructed from data provided by Adda River Authority and National Hydrographic Service or digitised from historical hardcopy and it is the result of a procedure of merging and correction of data collecting at different stations. In particular, the sub-periods 1845–1922, 1931–1934 and 1944–1945 were derived from Malpensata and Malgrate water level gauges by applying a linear relationship computed on the basis of available overlapping periods with Fortilizio series. Moreover, since 1946, when a new dam in Olginate started lake regulation, runoff values at Fortilizio were derived from the daily series of Lavello station, located downstream of the dam and having runoff data in common with Fortilizio for the period 1946–1950. Thanks to this activity of data recover, a very long and uninterrupted record of runoff for the upper part of Adda river basin is now available for both climatological studies, as discussed in the present work, and operative purposes as decisional support tool in the management of lake regulation. Further details on the reconstruction and analysis of outflows and water levels at Fortilizio station can be retrieved in Ranzi et al. (2017, 2018).

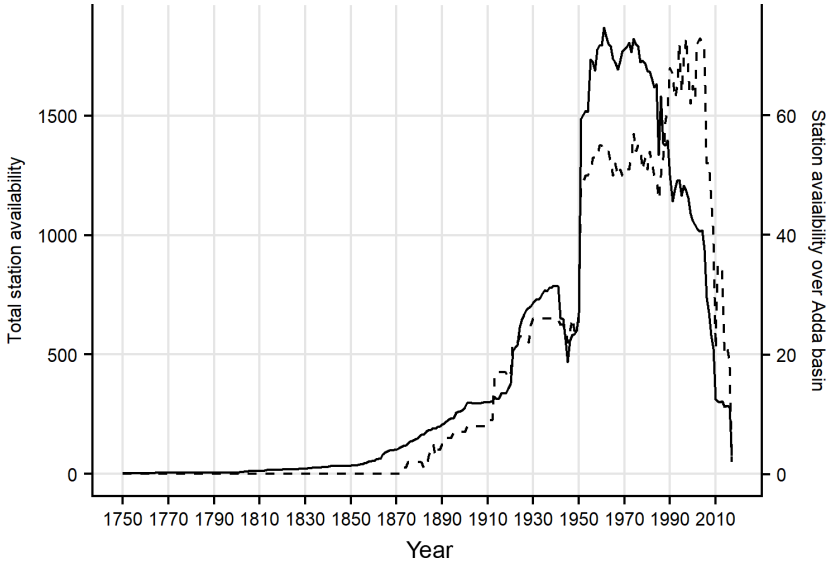

Figure 2. Availability of monthly precipitation records over the whole period spanned by the database considering all stations (solid line) and only stations located inside upper Adda river catchment (dashed line).

2.2 The anomaly method: gridding climatologies and anomalies

The first step of the anomaly method is the calculation of the station monthly normals that we estimated for the 1961–1990 period as it corresponds to the best data availability of our database. However, since a relevant fraction of stations has missing data in 1961–1990 (about 40 % of stations in the study area have no more than 50 % of available data in this period), before computing their 1961–1990 monthly normals, series were subjected to the gap filling procedure presented in Crespi et al. (2018). Station normals are then used both to transform the station records into anomaly records and as input data for the model applied to get the 30 arcsec monthly precipitation climatologies. Monthly climatological fields are obtained, as described in Crespi et al. (2018), for a smoothed version of 30 arcsec resolution digital elevation model by means of a Local Weighted Linear Regression (LWLR) of precipitation versus elevation applied at each grid cell:

$$p(x,y) = a(x,y) + b(x,y) \cdot h(x,y) \quad (1)$$

where $h(x,y)$ is cell elevation and $a(x,y)$ and $b(x,y)$ are the regression coefficients estimated at grid point. The precipitation-elevation relationship at each cell is defined by means of a weighted linear regression applied on neighbouring stations. In fact, in order to prioritise stations which are mostly representative of orographic conditions of target location, they enter in the regression with weights depending on their nearness and orographic similarity (elevation, slope steepness, slope orientation and sea distance) to the grid cell. More precisely, the weights are expressed as the product of Gaussian functions, one per geographical feature (par):

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

whose decreasing rate c^{par} is optimally defined for each grid point and month by means of a minimisation procedure of model errors. This approach is based on the assumption that precipitation distribution is strongly influenced by the orographic features of the domain, especially elevation, and that this link could vary along the year and at a very local level (Daly et al., 1994, 2002). It is therefore supposed to be suitable to handle with complex domains such as Adda river basin, where remarkable orographic gradients occur.

1845–2016 station monthly anomaly records, which are defined as the ratio of absolute values to the corresponding monthly normals, are then projected onto the same grid points by means of a weighted average of neighbouring stations. Station weights are in the form of Eq. (2) and they take into account distance and elevation difference from the grid cell. The decrease of distance weight is regulated year by year accordingly to the variation in data density over the reconstructed period. More precisely, for each year the halving distance of weight is set to the mean grid radius containing at least three valid station measurements. This procedure allows to use the fine-scale information available for periods of dense station coverage and to exploit also the data available from distant stations when data availability gets lower.

Finally, gridded anomalies and climatologies are superimposed by their product to obtain the 1845–2016 gridded

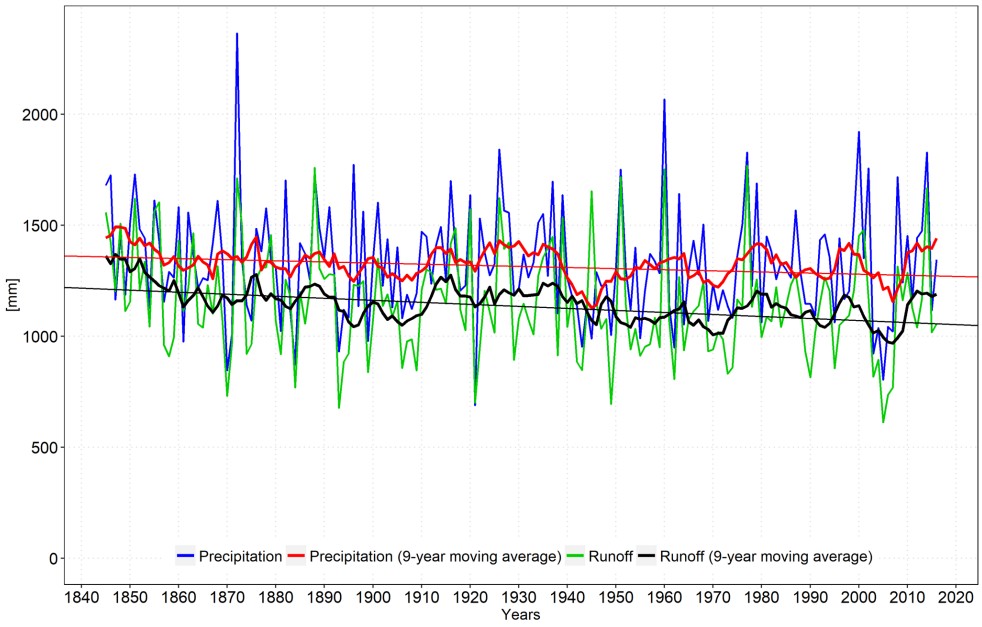

Figure 3. 1845–2016 total annual precipitation (blue line) and annual runoff (green line) records for upper Adda river catchment together with estimated Theil-Sen trends (red and black straight lines). Annual cumulates are computed considering solar year.

dataset of monthly precipitation records in absolute values covering the whole upper Adda river basin.

Robustness of climatology and anomaly models was assessed by reconstructing station monthly normals and anomaly series in leave-one-out approach (LOO), i.e. by excluding the station under reconstruction in order to avoid self-influence, and comparing them to measured values. All the 338 stations within the study domain are considered as validation subset. LOO evaluation was however possible only since 1861 because there are no stations inside the considered area before this year. Relative Mean Absolute Error (MAE) for the reconstruction of station anomaly series was found to range, as monthly average, between 14 % in summer and 24 % in winter, whereas the mean common variance between observed and reconstructed series turned out to be 0.86. The time evolution of MAE, evaluated year by year over all reconstructed stations, presents higher and more unstable values in the beginning period, when network is sparser and few validation stations are available. However the error turns out to decrease rapidly after the 1870s and it remains quite constant around 20 % over the rest of the spanned period (figure not shown).

3 Results and Conclusions

One of the most interesting applications of the dataset of high-resolution gridded precipitation records we obtained for the upper part of Adda river catchment is the reconstruction of an areal series of monthly precipitation from 1845 to 2016, obtained integrating precipitation over the catchment domain. This series provides the total amount of precipitation falling into the catchment area at any time step, information not available using the original sparse station data without applying the interpolation procedure. In order to evaluate the effects of the variability in station coverage on the robustness of this series, we took as reference the precipitation estimates for 1951–2000 period, which corresponds to the years with the highest data density, and we iteratively reconstructed them by varying station availability. More precisely, 1951–2000 areal precipitation record was iteratively evaluated by using only the available data in each 5-year subsequent period from 1846 to 1950. All the simulated annual records show a remarkable agreement with reference series and correlation value rapidly increases from 0.72 when the first 5-year availability is considered (1846–1850) to more than 0.9 using the available stations after 1860.

Considering solar year, annual cumulated values of 1845–2016 areal precipitation series are here presented in Fig. 3, together with the corresponding yearly runoff values at Lake Como outlet, spanning the same period and referring to the same domain. Looking at Fig. 3, the evolution of the two variables over the reconstructed period shows a great agreement and the correlation value between their annual records over the whole period is 0.79 suggesting the robustness of the independently reconstructed datasets.

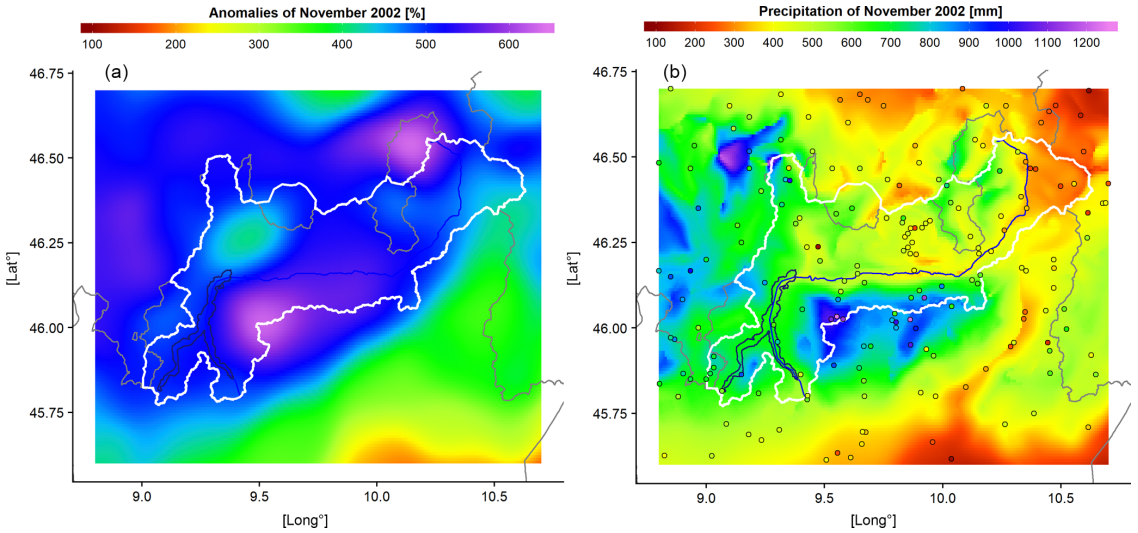

Figure 4. Monthly anomaly and precipitation distribution over the domain in November 2002. Points on the second panel represent station locations with available data and they are coloured in accordance to measured precipitation.

Both catchment precipitation and runoff yearly records were subjected to trend analysis by means of Theil-Sen test and Mann-Kendall estimator in order to assess and quantify slopes and statistical significances of precipitation and runoff trends over 1845–2016 period. Setting significance level to 0.05, a negative but not statistically significant trend was found for precipitation record with a Theil-Sen slope of $-5.1\,\mathrm{mm\,decade^{-1}}$, while a negative statistically significant tendency was pointed out for runoff with a Theil-Sen slope of $-9.2\,\mathrm{mm\,decade^{-1}}$. We considered also the series of annual runoff to precipitation ratio for which a significant negative trend was found with a decrease of about 0.04 % per year. The different behaviour of precipitation and runoff can be possibly explained arguing an increase of evapotranspiration losses due to the temperature increase, only partially compensated by an increasing rate in glacier melting. However, other not negligible factors could be ascribed as possible causes of the significant decline of runoff from 1845 to date. In particular, the growth of forest coverage occurred in the last decades over the region and due to natural afforestation, as discussed for instance in Ranzi et al. (2017), could have contributed to increase the evapotranspiration rate (Birkinshaw et al., 2014). Also the changes in regulation of water uptake for irrigation purposes over the last century could have a significant impact on runoff regime. However further and detailed information about the evolution of land-use and human activities have to be collected in order to quantify their role in runoff variations.

Another important application of the high-resolution gridded precipitation dataset is the reconstruction of extreme past events. This task could be even more meaningful when dealing with daily data. However, most series in our database has daily resolution only after 1950, whereas before this year daily data are available only for a fraction of records. We limit therefore our analyses to the monthly resolution. In order to assess the ability of the applied methodology to capture the occurrence and distribution of extreme events, we consider as case study the month with the highest value in catchment precipitation record (November 2002 with an average of 576 mm) and represent the spatial distribution for both precipitation and anomalies over the domain in Fig. 4. As highlighted by the distribution of station points, the considered month is characterised by a dense data coverage, with more than 150 available stations inside Adda basin, which allows to depict in detail the spatial pattern of the exceptional precipitation event occurring over Adda river catchment during the central days of the month.

In order to further assess the robustness of single event reconstruction also when a lower data coverage is available, the anomaly and precipitation distributions for November 2002 are evaluated by reducing the number of series entering in the reconstruction and selecting only the stations available in 1882 and in 1922 (or the closest ones in case of records ending before 2002), i.e. 120 and 80 years before 2002, respectively (Fig. 5). As pointed out in Fig. 5, in both periods station availability is significant lower than in 2002 with 22 stations in 1922 and only one (Como station) in 1882 located inside Adda basin. Despite data coverage discrepancy, the comparison with the anomaly distribution obtained from all stations operating in 2002 (Fig. 4) highlights a remarkable agreement with the anomaly field depicted by means of 1922 station distribution (Fig. 5b). Even though the detailed structure of event is not completely captured, the leading spatial features are still evident even when the lowest station cov-

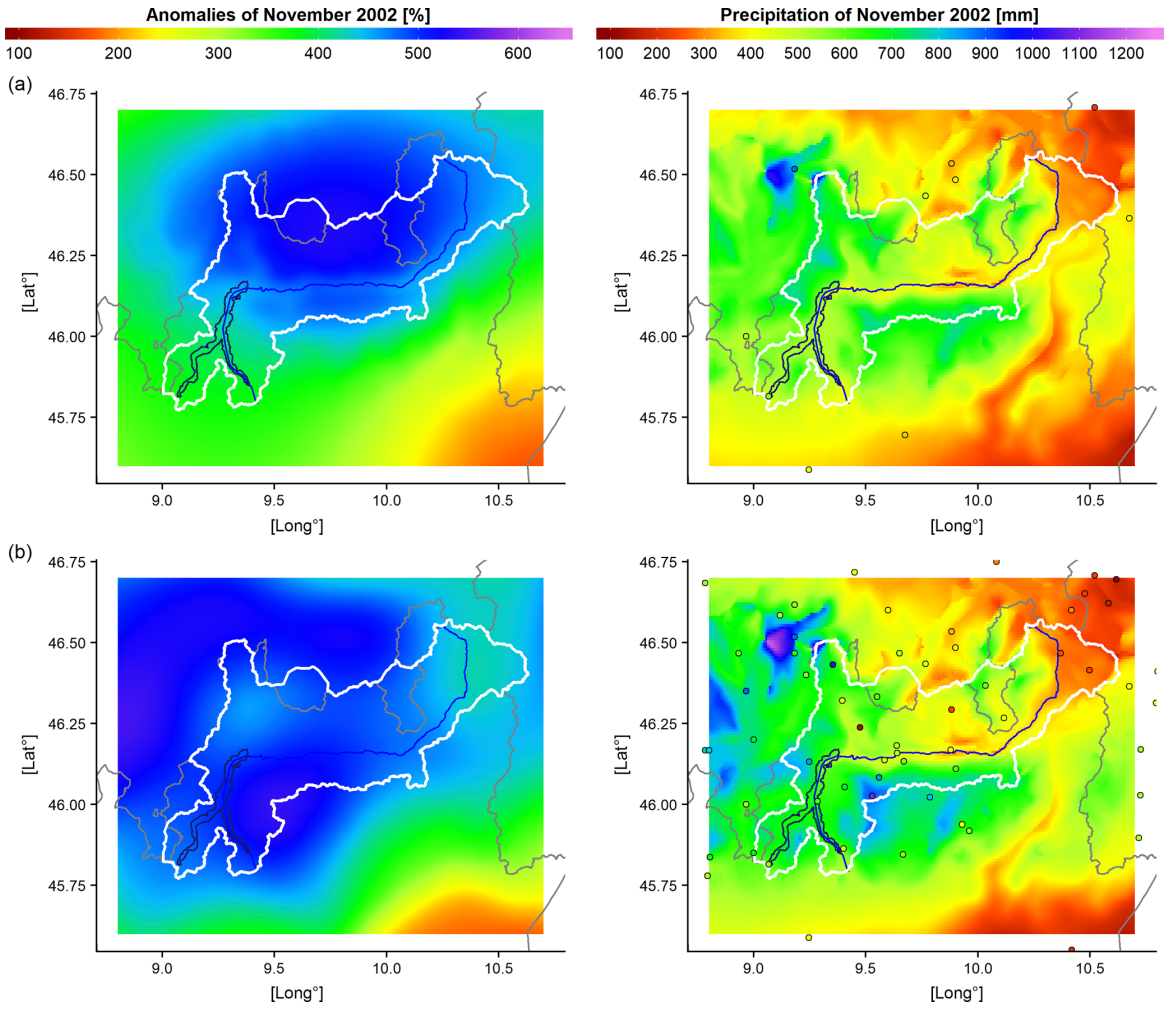

Figure 5. Monthly anomaly and precipitation distribution over the domain in November 2002 obtained by considering station availability in **(a)** 1882 and **(b)** 1922. Points represent station locations with available data and they are coloured in accordance to measured precipitation.

erage of 1882 is considered (Fig. 5a). Moreover, the catchment precipitation for November 2002 estimated by 1882 and 1922 station density turns out to be only 10 and 6 %, respectively, lower than the one obtained with 2002 station availability. This case study shows therefore that thanks to the high spatial coherence of anomalies, the main features of a meteorological event could be reconstructed even when only a few number of stations is available and unevenly located over the study area. This test confirms also the reliability of the computed areal precipitation record for Adda basin which was discussed above and it provides further information about the impact of variations in station coverage on the variability of modelled values.

The analyses we performed show the great potential of the database and of the methodology applied for the construction of 1845–2016 gridded precipitation records. However, further analyses are required to investigate more in detail the error of grid point and catchment records, especially for the periods with lower data availability, as well as to improve trend evaluation by applying this analysis over different time windows for both precipitation and runoff series. The results could be useful to better explain the different behaviour of precipitation and runoff over upper Adda catchment and to identify the main forcing factors affecting the changes and interactions of these two variables. These analyses are in progress and they will be discussed in forthcoming papers.

Data availability. The observation database of monthly precipitation for the study area was set up by the data retrieved from: ISPRA (Istituto Superiore per la Protezione e Ricerca Ambientale) archive (http://www.scia.isprambiente.it/home_new_eng.asp, last access: June 2018), ARPA Lombardia (http://www.arpalombardia.it/siti/arpalombardia/meteo/richiesta-dati-misurati/Pagine/RichiestaDatiMisurati.aspx, last access: June 2018), HISTALP archive (http://www.zamg.ac.at/histalp/, last access: June 2018), the database of CNR-ISAC (Italian National Research Council-Institute of Atmospheric Sciences and Climate) – Milan University (contact the authors of this paper for data requests), MeteoTrentino (https://www.meteotrentino.it/index.html#!/content?menuItemDesktop=141, last access: June 2018) and MeteoSwiss (https://gate.meteoswiss.ch/idaweb/, last access: June 2018). The precipitation climatologies over the study area are retrieved from the gridded dataset of 1961–1990 monthly climatologies which is freely available on ISAC-CNR website (http://www.isac.cnr.it/climstor/CLIMATE_DATA/, last access: June 2018).

Author contributions. AC, MM and MB managed the reconstruction of the gridded dataset of secular monthly precipitation series over the Adda river basin. RR and MT provided the 1845–2016 runoff record.

Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue "17th EMS Annual Meeting: European Conference for Applied Meteorology and Climatology 2017". It is a result of the EMS Annual Meeting: European Conference for Applied Meteorology and Climatology 2017, Dublin, Ireland, 4–8 September 2017.

Acknowledgements. We sincerely thank all data providers who contributed to set up the database we used for the present study. We also thank the reviewers for their very useful suggestions and comments which allowed to improve the quality of the manuscript.

Edited by: Ole Einar Tveito
Reviewed by: Melita Perčec Tadić and one anonymous referee

References

Auer, I., Bohm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schoner, W., Ungersbock, M., Matulla, C., Briffa, K., Jones, P., Efthymiadis, D., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisselin, J.-M., Begert, M., Muller-Westermeier, G., Kveton, V., Bochnicek, O., Stastny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic-Capka, M., Zaninovic, K., Mejstorovic, Z., and Nielplova, E.: HISTALP – historical instrumental climatological surface time series of the Greater Alpine Region, Int. J. Climatol., 27, 17-46, https://doi.org/10.1002/joc.1377, 2007.

Birkinshaw, S. J., Bathurst, J. C., and Robinson, M.: 45 years of non-stationary hydrology over a forest plantation growth cycle, Coalburn catchment, Northern England, J. Hydrol., 519, 559–573, https://doi.org/10.1016/j.jhydrol.2014.07.050, 2014.

Brugnara, Y., Brunetti, M., Maugeri, M., Nanni, T., and Simolo, C.: High-resolution analysis of daily precipitation trends in the central Alps over the last century, Int. J. Climatol., 32, 1406–1422, https://doi.org/10.1002/joc.2363, 2012.

Brunetti, M., Maugeri, M., Motti, F., and Nanni, T.: Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series, Int. J. Climatol., 26, 345–381, https://doi.org/10.1002/joc.1251, 2006.

Craddock, J.: Methods of comparing annual rainfall records for climatic purposes, Weather, 34, 332–346, https://doi.org/10.1002/j.1477-8696.1979.tb03465.x, 1979.

Crespi, A., Brunetti, M., Lentini, G., and Maugeri, M.: 1961–1990 high-resolution monthly precipitation climatologies for Italy, Int. J. Climatol., 38, 878–895, https://doi.org/10.1002/joc.5217, 2018.

Daly, C., Neilson, R. P., and Phillips, D. L.: A statistical-topographic model for mapping climatological precipitation over Mountanious Terrain, J. Appl. Meteorol., 33, 140–158, https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2, 1994.

Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P.: A knowledge based approach to the statistical mapping of climate, Clim. Res., 22, 99–113, 2002.

Efthymiadis, D., Jones, P. D., Briffa, K. R., Auer, I., Böhm, R., Schöner, W., Frei, C., and Schmidli, J.: Construction of a 10-min-gridded precipitation data set for the Greater Alpine Region for 1800–2003, J. Geophys. Res., 111, D01105, https://doi.org/10.1029/2005JD006120, 2006.

Isotta, F. A., Frei, C., Weilguni, V., Tadić, M. P., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertačnik, G.: 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, https://doi.org/10.1002/joc.3794, 2014.

Masson, D. and Frei, C.: Long-term variations and trends of mesoscale precipitation in the Alps: recalculation and update for 1901–2008, Int. J. Climatol., 36, 492–500, https://doi.org/10.1002/joc.4343, 2016.

Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol., 25, 693–712, https://doi.org/10.1002/joc.1181, 2005.

New, M., Hulme, M., and Jones, P.: Representing Twentieth-Century space-time climate variability. Part II: development of 1901–96 monthly grids of terrestrial surface climate, J. Climate, 13, 2217–2238, https://doi.org/10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2, 2000.

Ranzi, R., Caronna P., and Tomirotti, M.: Impact of climatic and land use changes on riverflows in the Southern Alps, in: Sustainable Water Resources Planning and Management Under Climate Change, edited by: Kolokytha, E., Oishi, S., and Teegavarapu, R. S. V., Springer Science+Business Media, Singapore, 61–83, https://doi.org/10.1007/978-981-10-2051-3, 2017.

Ranzi, R., Bacchi, B., Tomirotti, M., Castioni, C., Brunetti, M., Crespi, A., and Maugeri, M.: Analisi delle tendenze di lungo termine nel regime degli afflussi meteorici e dei deflussi dell'Adda a Lecco (1845–2014), L'Acqua, 2, 51–60, 2018.

Servizio Idrografico: Osservazioni pluviometriche raccolte a tutto l'anno 1915. Volume II: Bacino Imbrifero del Po. Fascicolo I. Ministero dei Lavori Pubblici – Consiglio Superiore delle Acque, Roma, 1920 (in Italian).

Servizio Idrografico – 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, 1957 (in Italian).

4.2 A multi-century gridded dataset of monthly precipitation records for the Adda river basin (Central Alps) based on in-situ observations

4.2.1 Introduction

The temperature increase of global warming period has been particularly strong in the Alpine region, which has recorded a significantly higher temperature trend than Earth's average one. This relevant temperature increase is documented by one of the best networks of secular meteorological stations in the world (see e.g. Auer et al., 2007) which gives evidence of a temperature increase of about 1.5 °C in the last 150 years, recorded both at low-elevation and high-elevation areas (Böhm et al., 2001; Brunetti et al., 2009). Beside temperature trend, also precipitation variability and changes in its seasonality over the Alpine region were documented in the scientific literature (e.g. Schmidli et al., 2002; Casty et al., 2005; Brunetti et al., 2006a).

Global climate change poses a grave threat to the Alpine hydrological system, causing a significant pressure on the key role that Alps have for water storage and supply with potential impacts on economic activities, such as agriculture, energy and industrial production, also over wide portions of surrounding valleys and plains (EEA, 2009; Viviroli et al., 2006).

In order to better understand this problem, it is crucial i) to investigate the spatio-temporal evolution of precipitation over Alpine areas on secular time scales, ii) to assess long-term areal precipitation records for the main Alpine catchments and iii) to compare them with the corresponding runoff records. The comparison with runoff records allows in fact to highlight the effect of temperature induced changes in the occurrence of solid and liquid precipitation and its seasonality, in the snow water equivalent (SWE) as well as in the contributions of melting glaciers and evapotranspiration on the water disposal. High-resolution datasets of monthly precipitation have been recently produced over the greater Alpine region (GAR) for example by Efthymiadis at al. (2006) for the period 1800–2003 at 10-min resolution and by Masson and Frei (2016) for the period 1901–2008 at 5 km resolution, while Isotta et al. (2014) provided a dataset at 5 km grid spacing of 1971–2008 daily precipitation series. High-resolution analyses of precipitation were also provided over smaller Alpine domains, such as by Golzio et al. (2018) and Brugnara et al. (2012) for the central European Alps, by Gyalistras (2003) for Switzerland and by Durand et al. (2009) for French Alps. Among the gridding methods which have been proposed so far, the so-called "anomaly method" is one of the most applied approaches (e.g. New et al., 2000; Mitchell and Jones, 2005). All these gridding procedures aim at projecting the in-situ observations onto the cells of a regu-

lar grid, allowing to evaluate the climatic signal for a number of points which can be of several orders of magnitude larger than the number of the available rain-gauges. Gridded datasets find application in a wide range of analyses, including the validation of regional climate models, the study of extreme events, the river runoff assessment and the evaluation of long-term variability and trend (Daly, 2006). The long-term analyses performed on both monthly and daily scales revealed different trends in total precipitation and in statistical indexes over different subregions of the Alps (e.g. Brunetti et al., 2006a; Scherrer et al., 2016; Pavan et al., 2018). In addition to the spatial variability, the significance and features of the trends were found to be dependent also on the period considered for the evaluation, suggesting the need of continuous update for both present and past years (see e.g. Brugnara et al., 2012).

In this framework, we applied the anomaly method to reconstruct a 30-arc second resolution dataset of 1800–2016 monthly precipitation records for an area centred over the upper part of Adda river basin, which represents an important water reservoir in Italian Central Alps for a wide portion of Northern Italy. This area is covered by a relevant number of long meteorological observations and it has one of the longest runoff records in Italy which has been recently recovered at daily resolution over the period 1845–2016 by Ranzi et al. (2017). It represents therefore a very interesting region to better investigate the issue of the response of Alpine water resources to climate variability and change.

The gridded precipitation dataset here presented constitutes an improvement to the already existing datasets covering the area: it is based on a very dense observation database spanning more than two centuries and containing quality checked and homogenised series retrieved from both national and extra-national sources, from new digitisation of ancient data and from the recent automatic station records, which allow to extend the reconstruction up to date and to improve the data coverage, especially at higher elevations. In the last decades in fact a relevant number of rain-gauges was established in mountainous areas where observations are requested for both hydropower production and activities of prevention of natural hazards, such as landslides, avalanches and floods, which could occur over these vulnerable environments. The availability of new long records allows to reduce the general difficulty in reconstructing the spatial distribution and temporal variability of precipitation over mountainous regions. The main limitations to the evaluation of a reliable climate signal for these areas are in fact the uneven coverage of observations decreasing towards higher elevations, the low availability of long secular station records, especially at daily resolution, and the highly heterogeneity of precipitation gradients due to the complex interactions between atmospheric circulation and the roughness of the surface (Haylock et al., 2008).

The paper aims at presenting the dataset, at assessing the robustness of the gridding

methods in relation to the development of the station network and at describing the multi-century record of areal precipitation extracted from the monthly gridded fields for the upper portion of the Adda river basin. This record is then analysed to investigate the trends and variability in precipitation regime over the basin and to perform the comparison with the available runoff record over the period 1845–2016. All the analyses were performed on monthly resolution only due to the low availability of daily precipitation data before 1951, which are still stored in hardcopy yearbooks. However, the monthly database allowed to extend the precipitation reconstruction back to 1800 when the study domain was covered by a very sparse station network. In addition, the availability of total areal precipitation values for the basin over the first half of the 19^{th} century could represent a valuable proxy-data which could help to estimate runoff even when this information was not yet collected.

4.2.2 Materials and Methods

The study area

The present study focuses on an area centred on the upper part of the Adda river catchment (45.6°–46.7°N and 8.8°–10.7°E, Figure 4.1b). This area will be called thereafter study domain, while the word study basin will be used to specifically refer to the upper portion of the Adda river catchment (the region bordered by the yellow line in Figure 4.1). The Adda river is one of the main tributaries of Po river and its drainage area covers little less than 8000 km^2, which are located for 94% in Italy and 6% in Switzerland. The study basin includes an area of 4508 km^2 mostly located over the southern Alpine ridge in Lombardy region and characterised by a very heterogeneous orography. It extends from the origin of the Adda river in the Rhaetian Alps to Lavello's dam at the Lake Como outlet, with the main valley (Valtellina) characterised by an East to West orientation. The region is predominated by the mountain environment and relevant height gradients ranging from about 200 m a.s.l. of Lake Como to more than 4000 m a.s.l. of Piz Bernina. The occurrence of several secondary valleys with narrow and variously oriented axes enhances the orographic complexity of the domain. In addition, the study basin includes some of the main Alpine glaciers, which are located in the groups of Bernina, Disgrazia and Ortles-Cevedale and directly contribute to the hydrological cycle. In the last three decades, the province of Sondrio, where all the Alpine glaciers of Adda basin are located except for the eastern side of Bernina whose contribution could be considered negligible, experienced a strong reduction of glacier coverage of about 30% and a volume loss of about 1.4 km^3 for the period 1981–2007 (D'Agata et al., 2018). This loss corresponds to a mean annual contribution to the basin runoff of about $5 \cdot 10^7$ m^3 of water. This amount is relatively small (~1%) in comparison with the yearly total water input due to precipitation, but it is mainly released

during summer representing a significant support to water related activities when the water demand is higher. The available projections of future glacier extent in the study basin under climate change scenarios suggest a loss up to 2/3 of the present volume by the second half of the 21^{th} century (Garavaglia et al., 2014). This strong reduction would lead to the disappearance of the water supply from glacier melting and to the strengthening of the impacts of summer drought periods on water disposal. The water resources of the study basin are in fact currently exploited for a wide range of activities on both valleys and surrounding plains, such as civil use, agriculture, industry and, thanks to the mountainous features of the area, hydropower production. In order to regulate the activities and the related water uptakes, a relevant number of artificial reservoirs have been built over the past decades. The largest artificial reservoirs are the lakes of Cancano and S. Giacomo, which can store about $200 \cdot 10^6$ m^3 of water primarily used for electricity production. The volume of artificial water storages integrates the capacity of the natural reservoirs included in the basin, the most important of which is Lake Como with a surface of 145 km^2 and a volume of about 25 km^3.

The observation database

The database used to compute the study domain gridded dataset of 1800–2016 monthly precipitation series was set up by starting from the wide collection of monthly precipitation records performed by Crespi et al. (2018) which is mainly based on the archives of the former Italian Hydrographic Service (Servizio Idrografico, SI). Two different approaches were then adopted for the study domain (Figure 4.1b and red bordered region in Figure 4.1a) and for the outer area (Figure 4.1a).
As regards the study domain, the activities of data collection and rescue focused on improving the data availability by:

- including in the database the Trentino-Alto Adige/Südtirol monthly records provided by the digitisation project "Before 1921" (https://before1921.wordpress.com). These new data allowed to extend into the past 37 series previously available only from 1921 and to add two new records to the archive;

- digitising new data from the yearbooks of the former SI (Servizio Idrografico, 1920; 1925; 1959). After this activity, 28 series previously available only from 1951 were extended in the past. The records to digitise were selected by prioritising the stations of the study domain with the longest time coverage and located at the highest elevations, where the low availability of past in-situ observations is more relevant;

- performing a new download of all the precipitation records available for the

study domain from MeteoTrentino (updated at April 2017) and from MeteoSwiss (updated at September 2017). In case of available homogenised version of Swiss station series, they were preferred to the original ones;

- collecting all precipitation records provided by the mechanical and automatic stations of the Regional Environmental Protection Agency (ARPA) and the Geological Monitoring Centre (CMG) of Lombardy. The records were integrated to the database and the series of automatic stations were merged with those of the mechanical ones allowing to overcome the general fragmentation of records occurring at most Italian sites since the 1980s, due to the transition from manual to automatic monitoring systems as well as the transition of the national rain-gauge network from SI to the Italian administrative regions. However, the merging activity was not trivial and required a careful study of the available metadata. In some cases, in fact, the mechanical stations were dismissed and the new automatic stations were established in completely different places so that their records could not be integrated with the previous series. In other cases, the new rain-gauges were kept rather close to the mechanical ones with just minor relocations allowing the merging of the records after applying a homogeneity evaluation. In a minor fraction of cases, finally, the new automatic stations were kept exactly at the same site of the previous mechanical ones.

The precipitation records of the stations included inside the study domain are indeed the mostly relevant ones for the goals and results of the work. However, no data are available over this area before the 1860s and the station coverage remains quite sparse over the following decades too. In order to improve the observation availability and to get reliable information about the climate evolution also for the ancient years, it was therefore necessary to include in the data collection activities a wider region around the domain. For this region, all the longest series available from the monthly homogenised archives of MeteoSwiss and HISTALP (Auer et al. 2007) were considered together with the homogenised versions of some of the longest Italian precipitation records starting in the 18^{th} century or in the first half of the 19^{th} century (Brunetti et al., 2006b). These stations allow the evaluation of the precipitation signal in the earliest period, while they provide a marginal contribution to the reconstruction in the following decades when the station coverage over the study domain increases.
The collected series were then checked for quality and homogeneity, except for the ones derived from already quality controlled and homogenised archives. The quality-check procedure was performed, as described in Crespi et al. (2018), by comparing each measured series with a simulated one by means of neighboring observations. High errors in the comparison allowed to detect and remove gross errors such as outliers, spurious sequences of null values and digitisation oversights. The series showing high devia-

tions from the reconstructions and those with less than 10 years of data were definitely discarded from the database. When more than one source were available for the same site, the record with the longest time coverage and the highest quality was retained.
The homogeneity was then evaluated by applying the Craddock test (Craddock, 1979) on more than 400 monthly series at least 20-year long which were not derived from already homogenised archives and, in case of relevant breaks, the homogenisation was performed. 125 series containing on average 85 years of data were homogenised by identifying more than 300 breaks. By comparing again the homogenised series with the simulated ones from surrounding station data, the reconstruction errors were significantly reduced in respect with the values obtained for the original records.
The final monthly database contains 338 precipitation series inside the study domain and 102 inside the study basin, while the stations outside the study domain with data before 1861 are 53, with 10 stations having data before 1810. 120 series contain more than 120 years of data and 14 out of them are located within the study domain. The time evolution of the data availability for the study domain and inside the study basin is shown in Figure 4.2. The first stations operating inside the study domain started in 1861, and in 1873 inside the study basin. The station density increased significantly from the beginning of the 20^{th} century thanks to the contribution of the new meteorological network managed by SI. After the difficulties due to the Second World War, the greatest station availability is reached, especially from 1951 to the end of the century when SI was closed and the Italian national network was fragmented into regional and subregional managements with the concurrent transition from the mechanical to automatic monitoring systems.
The inter-station distance, i.e. the mean distance between each station and the closest one, over the whole area spanned by the database decreases gradually from about 150 km in the first decade, when less than 10 stations are available, to about 80 km in 1850 when 33 rain-gauges were operating and to 25 km in 1900 with more than 250 records. During the 20^{th} century this value reduces further reaching about 8 km within the study domain in the decades of maximum data availability (1951–2000).
The rain-gauge density varies significantly along the years also on altitude ranges (Figure 4.3). Most of the ancient sites are located at low elevations and the distribution gradually becomes more homogeneous over all the elevation bands in the 20th century. The improvement in the data availability for the highest elevations is mainly related to the development of hydropower sites even if a relevant contribution, especially over the last decades, derived from the increasing requests of meteorological monitoring for the natural hazard prevention as well as for the touristic promotion and fruition of mountains. Despite of this increment, the database is still not completely representative of the study domain since a relevant fraction of the grid cells are located above 2000 m a.s.l. (30%) where the coverage of in-situ data is lower or even missing.

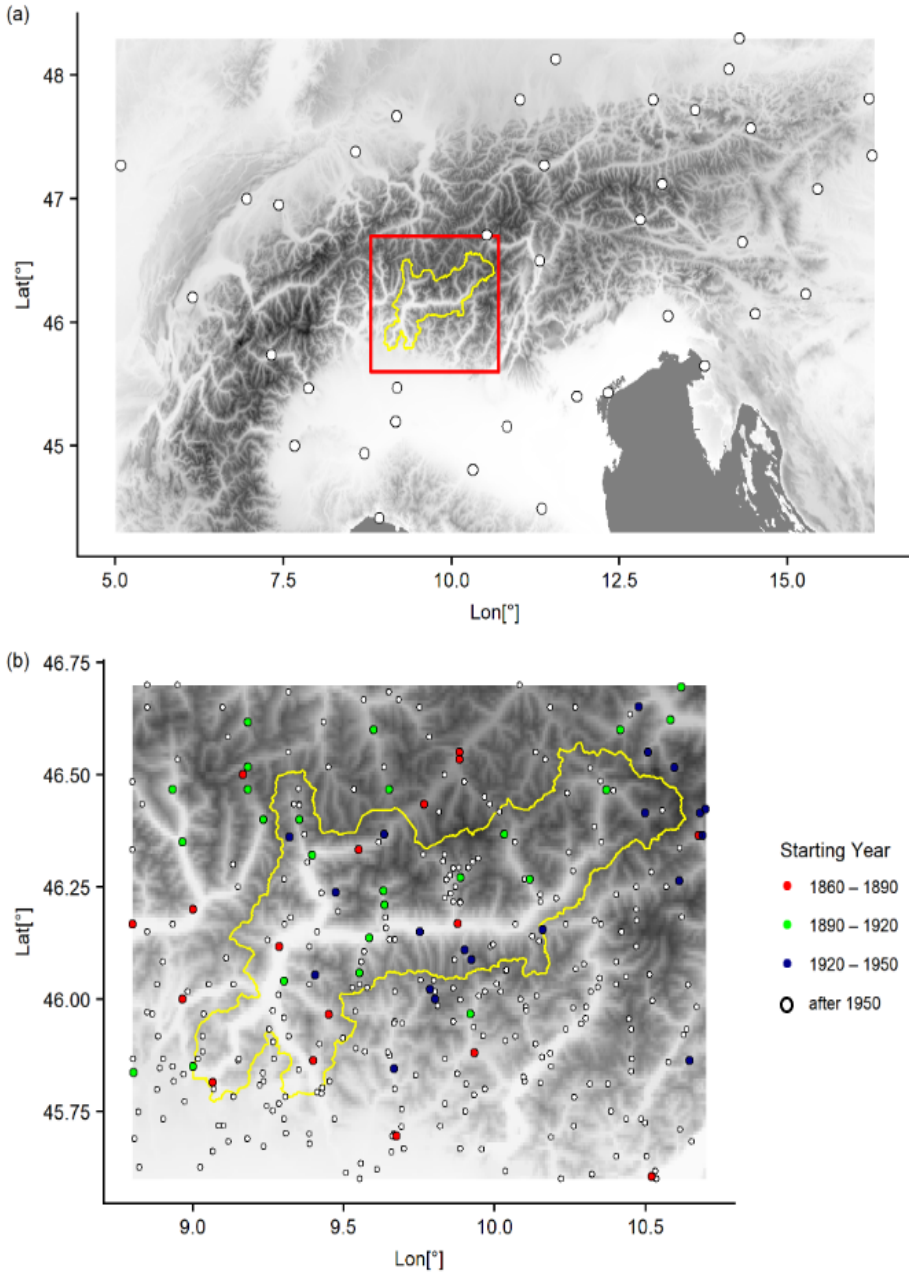

Figure 4.1: Spatial distribution of the available stations over a) the larger area surrounding the domain and b) inside the study region. In panel a) only the station series starting before 1861 are shown, in panel b) all the available records are reported together with the corresponding starting year.

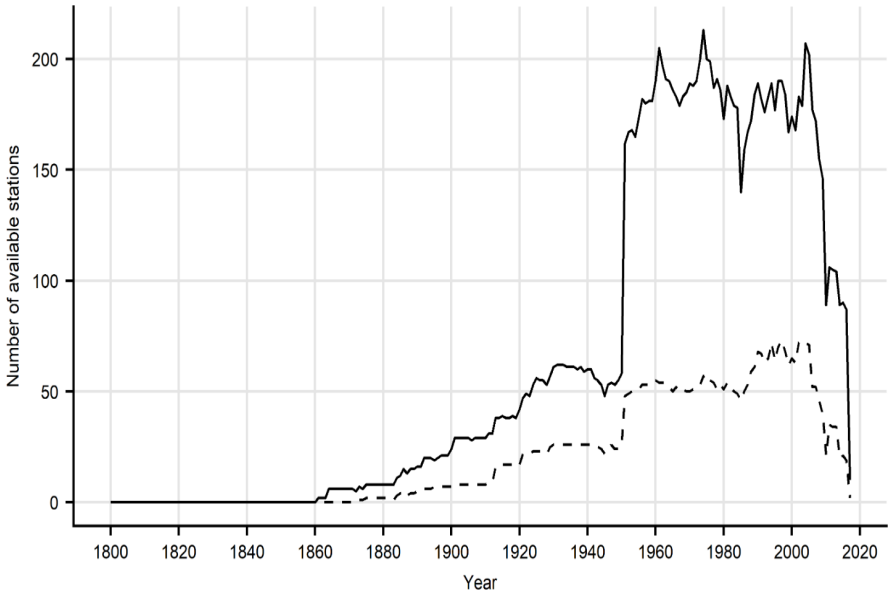

Figure 4.2: Availability of monthly precipitation of over the study domain (solid line) and the study basin (dashed line).

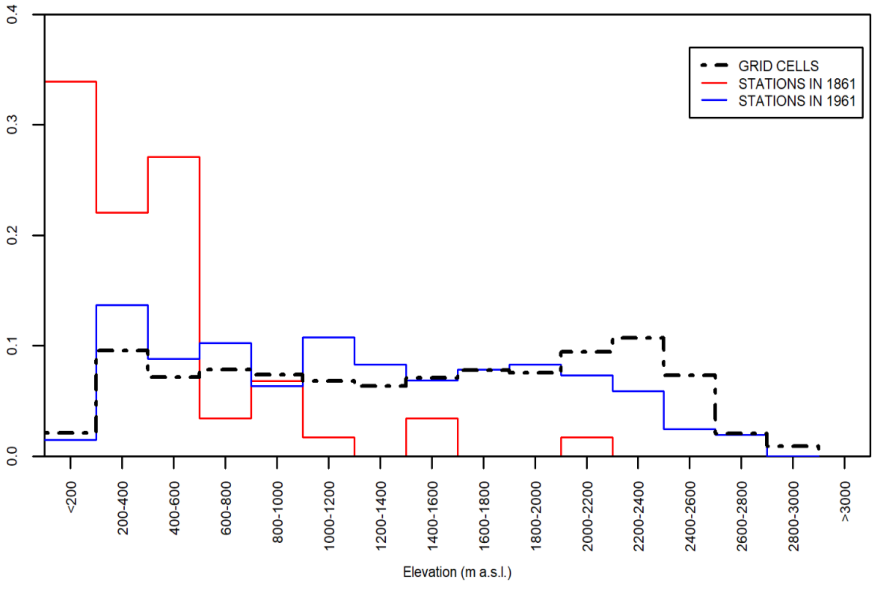

Figure 4.3: Distribution of stations on elevation intervals in the year of maximum data availability (1961, blue line) and 100 years before (red line) together with the elevation distribution of the grid cells covering the study domain (dashed line). In 1861 all the stations are considered, whereas in 1961 only those located inside the study domain are used.

The interpolation scheme: from rain-gauge network to regular grid

The gridded dataset of 1800–2016 monthly precipitation records over study domain was reconstructed on a 30-arc second resolution Digital Elevation Model (DEM) by means of the anomaly method. In this scheme the climatic signal of precipitation is reconstructed by superimposing the spatial fields of monthly climatologies and the spatio-temporal fields of relative anomalies, i.e. the deviations from the reference for a certain month. The two fields are computed separately and the final monthly estimate is obtained by their product. The 1961–1990 monthly precipitation climatologies were considered as the reference conditions. They were constructed by a local weighted linear precipitation-elevation regression (LWLR). In this method the monthly precipitation normal at each DEM cell is defined by applying the cell regression coefficients (slope and intercept) to the cell elevation (Daly et al., 2002; Crespi et al., 2018). The precipitation-elevation regression is computed giving more importance to the stations which are mostly representative of the orographic conditions of the target point. For this purpose, the stations enter in the regression with weights depending on their nearness and geographical similarity (e.g. elevation, slope steepness and slope orientation) to the cell to estimate. The weights are of Gaussian shape and their halving factors are locally defined month-by-month by means of a minimisation procedure of model errors (Crespi et al., 2018).

In order to perform the LWLR the 1961–1990 normals of all the monthly stations were computed after completing their missing records over the 30-year period (Crespi et al., 2018). The chosen period of reference corresponds to the interval of maximum data availability in the database and the resulting coverage of the study period is on average one station per 53 km^2. In order to take into account the actual spatial scales at which the interactions between atmospheric circulation and orography are expected to occur, we smoothed out the too fine terrain details of the DEM but retaining the original horizontal step of 30 arc seconds (Daly et al., 2002; Foresti et al., 2018). The smoothing was performed by assigning to each cell a weighted mean of the elevations of surrounding cells whose weights decay with distance as Gaussian functions with halving distance of 3 km. Also in this case, the halving distance was defined on the basis of the minimisation of model errors.

The relative anomaly fields were obtained by normalising the station records over the period 1800–2016 to their corresponding 1961–1990 normals and interpolating them over the 30-arc second resolution grid by means of a weighted averaging scheme. In this method, the relative anomaly for a certain month at each grid cell is defined by the weighted average of the surrounding station anomalies whose weights are expressed again as the product of Gaussian functions depending on radial and vertical distance from the target point.

The halving distance of the radial weight was defined year-by-year according to the evolution of data availability over the whole considered period. In particular, it corresponds to the mean radius of the circle centred on the grid cells and containing at least three stations with valid observations. The value rapidly decreases from about 200 km during the earliest years to about 20 km in 1900 and it decreases further to 10 km after the 1951 up to present. This approach allowed to exploit the fine-scale information provided by the period of dense data coverage and to include the available records on a larger area when the station distribution gets sparser.
The halving factor for the weight of elevation difference was set to 2250 m over the whole period on the basis of a minimisation error procedure. It is relatively high and so that the station weight are mainly defined by their nearness to the considered cell.
The 1800–2016 monthly precipitation records in absolute values over the grid were finally obtained by multiplying the 1961–1990 monthly climatologies times the interpolated series of monthly anomalies. As examples of the resulting fields, in Figure 4.4 the precipitation distribution for the months featuring the highest total areal precipitation on the study basin over each 40-year subperiods from 1861 to present are reported together with the available station records. In particular, for November 2002 the absolute maximum of the 1800-2016 total areal precipitation series on the study basin is estimated with 584 mm as monthly total. The figure allows also to highlight the gradually increase of data coverage over the study domain which passes from 8 stations for the firstly considered event in 1882 to almost 200 stations in 2002.

4.2.3 Results

Climatologies, anomalies and precipitation records

The accuracy of the modelled values was assessed by reconstructing in leave-one-out (LOO) approach the 1961–1990 monthly normals of the 338 stations inside the study domain, i.e. by excluding the observation under reconstruction in order to avoid self-influence. The model errors were evaluated in terms of mean error (BIAS), mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE). The error values are listed in Table 4.1. BIAS is almost null in all months even if a slight tendency to underestimation, especially in summer, is depicted. Despite the lowest MAE values, which are mainly due to the drier conditions, the model is affected by the greatest uncertainty in winter, when MAPE reaches the maximum in January (16%). The greater errors in winter could be partly ascribed to the difficulties in measuring the contribution from solid precipitation in high-elevated areas. The wind-induced undercatch is in fact one of the main causes of precipitation underestimation at mountain sites, which are mostly exposed to intense winds and snowfall events (Sevruk et al., 2009).

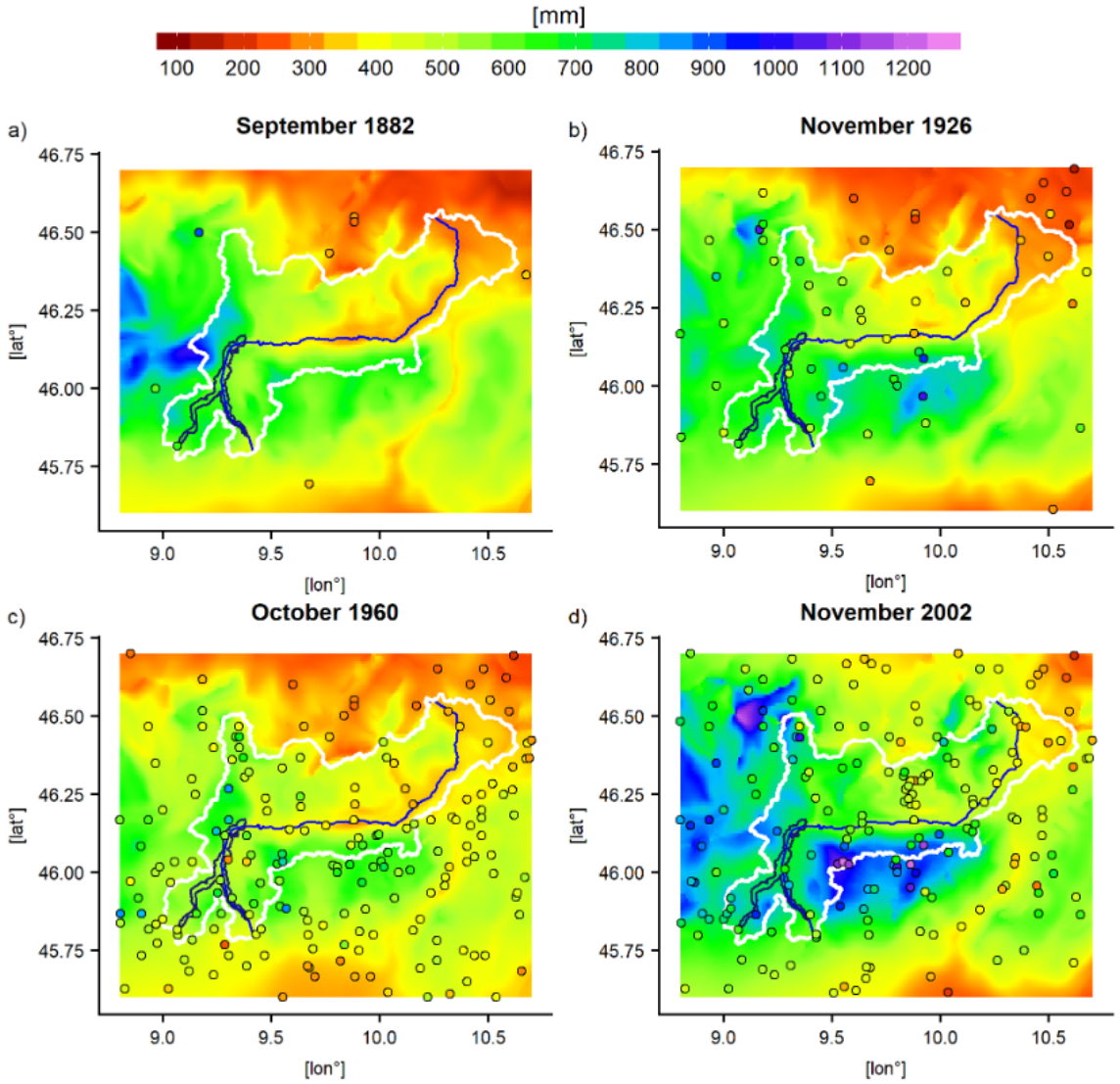

Figure 4.4: Monthly precipitation fields of the events with the greatest total areal precipitation on the study basin of the subperiods a) 1861–1900, b) 1901–1940, c) 1941–1980, d) 1981–2016. The points represent the monthly records at the available station sites.

MONTH	BIAS	MAE	MAPE [%]	RMSE
1	0.0	9.1	15.6	12.5
2	0.4	7.9	14.0	11.0
3	0.2	9.8	12.4	14.5
4	-0.1	13.7	12.1	20.5
5	-0.4	14.8	9.6	20.6
6	-0.3	12.6	9.3	16.6
7	-0.3	11.0	9.0	14.1
8	-0.5	11.1	8.0	14.8
9	-0.2	11.8	10.0	16.0
10	0.0	13.0	10.7	18.1
11	-0.2	13.6	12.3	20.3
12	0.2	8.0	14.3	11.6

Table 4.1: Monthly leave-one-out reconstruction errors of the 1961–1990 normals for the 338 stations included in the study domain. Except for MAPE, all the values are expressed in mm and BIAS is defined as the difference between simulation and observation.

As regards the annual precipitation climatology represented in Figure 4.5, the driest area is located over the North-Eastern part of the study domain along the Venosta Valley with less than 500 mm/year, while the highest precipitation values occur over the Western portion of the study domain over the Como Prealps and the Canton of Ticino with annual maxima exceeding 2000 mm, especially around the San Bernardino Pass. Other wet conditions are evident over the Orobian Prealps, South of the basin, with annual totals ranging between 1600 and 2000 mm, while the innermost part of Valtellina, all along the river course, is characterised by lower precipitation values, especially around Sondrio, where annual totals are below 800 mm/year. A similar regime is also evident along the course of Oglio river flowing South-East of the Adda basin into Lake Iseo where annual totals decrease of about 25% with respect to surrounding areas. The mean annual precipitation normals for the study domain and the study basin are 1314 mm and 1296 mm, respectively. The driest conditions occur during winter over the entire area, with less than 100 mm/season over wide portions of Valtellina, Grisons and Venosta Valley (Figure 4.6). Except for winter, the precipitation regime is almost invariant along the year, with slightly wetter conditions during summer when they are mainly driven by convective phenomena enhanced by moist and warm Mediterranean air at low levels and drier and colder continental currents at higher levels. Considering the average over the grid cells of the study basin only, the precipitation normal is 170 mm in winter (DJF), 350 mm in spring (MAM), 410 mm in summer (JJA) and 365 mm in autumn (SON). The yearly cycle of monthly precipitation over the study basin is shown in Figure 4.7 and it points out that the precipitation contributions are rather constant between April and November with a maximum in May.
The accuracy of the anomaly fields reconstruction was evaluated by simulating in LOO

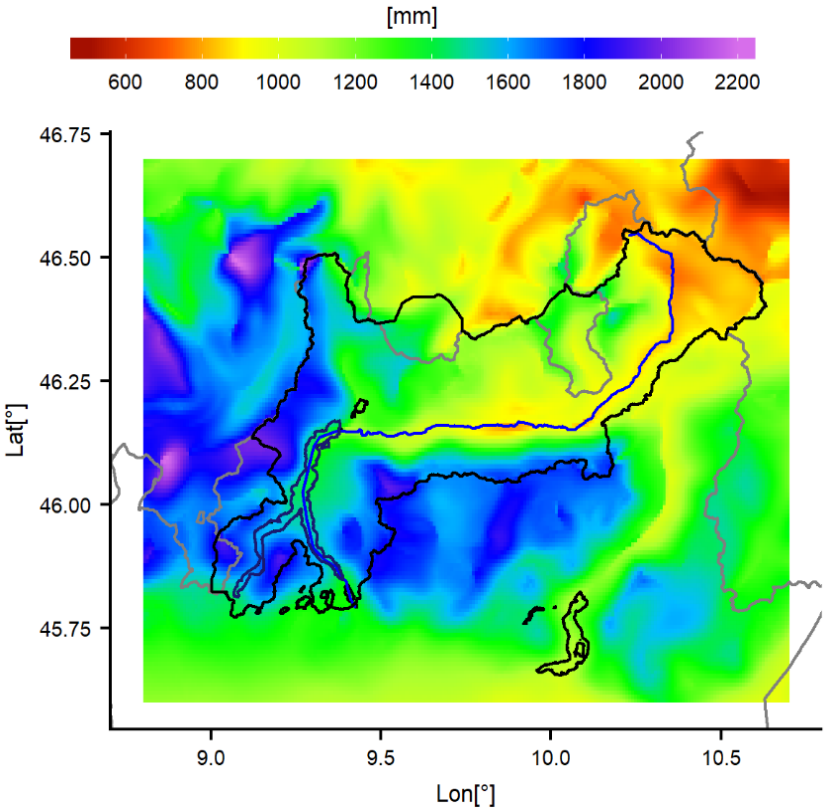

Figure 4.5: LWLR 1961–1990 annual precipitation climatology.

approach the monthly anomalies over the whole period of all the 338 stations within the study domain and the errors were assessed by the comparison of the modelled anomalies with the observed ones. Actually, the errors on station data reconstruction could be computed only from 1864, because no record is available for the study domain until 1861 and the data coverage before 1864, when most Swiss rain-gauges started operating, is still too low to provide a significant evaluation of the reconstruction ability at station sites. By considering all the available entries over the study period for each month, MAE ranges from 0.24 in December to 0.14 in May, confirming the greater difficulty in capturing winter precipitation already highlighted by the LOO simulation of station normals.

Figure 4.8 reports the time evolution along the 1864–2016 period of annual MAE and BIAS obtained over all the available station anomalies at each time step. The highest MAE values occur at the beginning of the considered interval with maxima around 0.30 when they are mostly affected by the uneven distribution and lower quality of data. Thanks to the improvements in station coverage and measurement accuracy, MAE gradually reduces over the following decades to about 0.20 and it decreases further from the second half of 20^{th} century during the period of best data availability when it is around 0.15. A slight MAE increment is evident at the end of the interval, which could be due to the decrease in the availability of updated data for the most re-

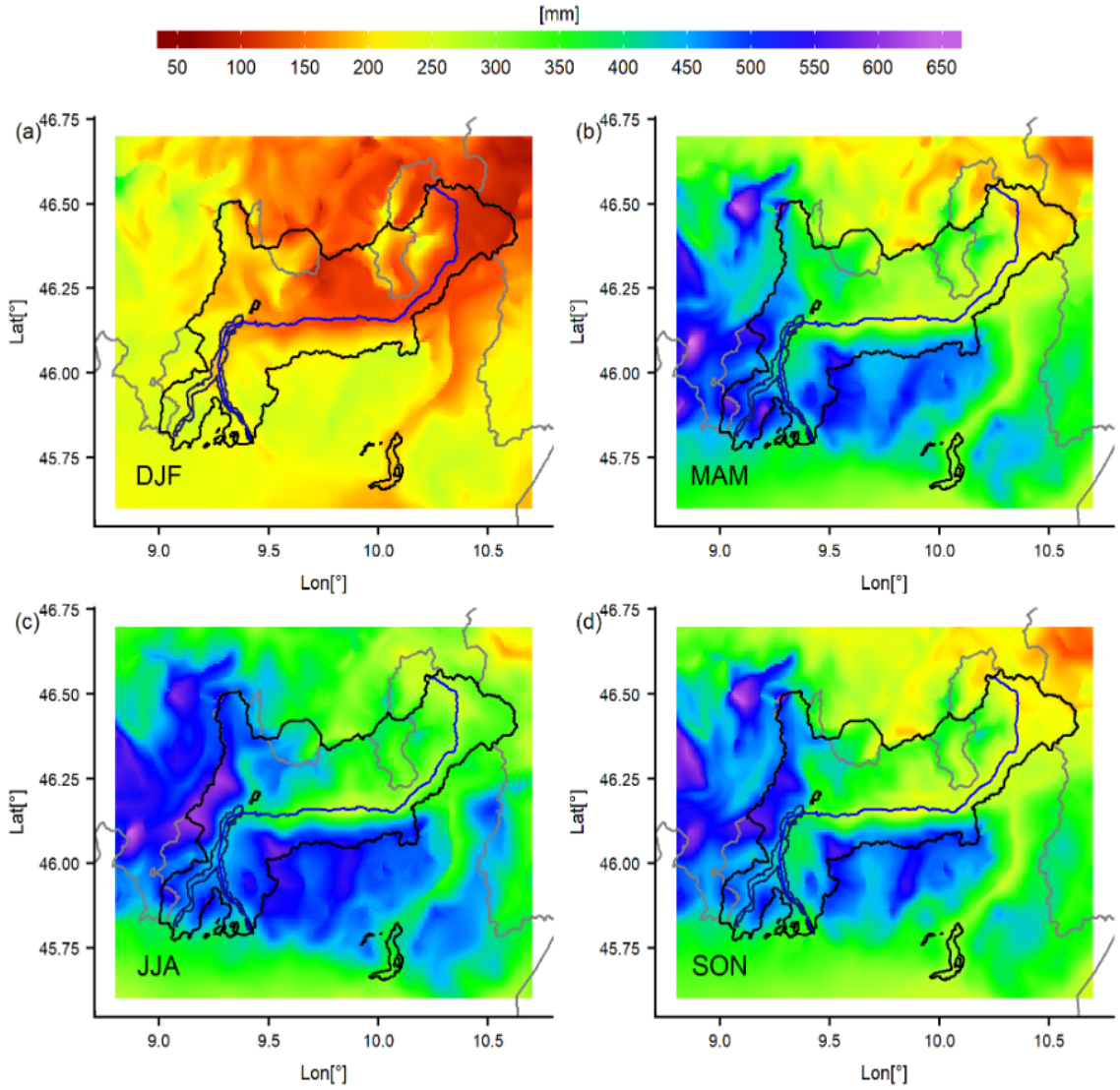

Figure 4.6: LWLR 1961–1990 seasonal precipitation climatologies.

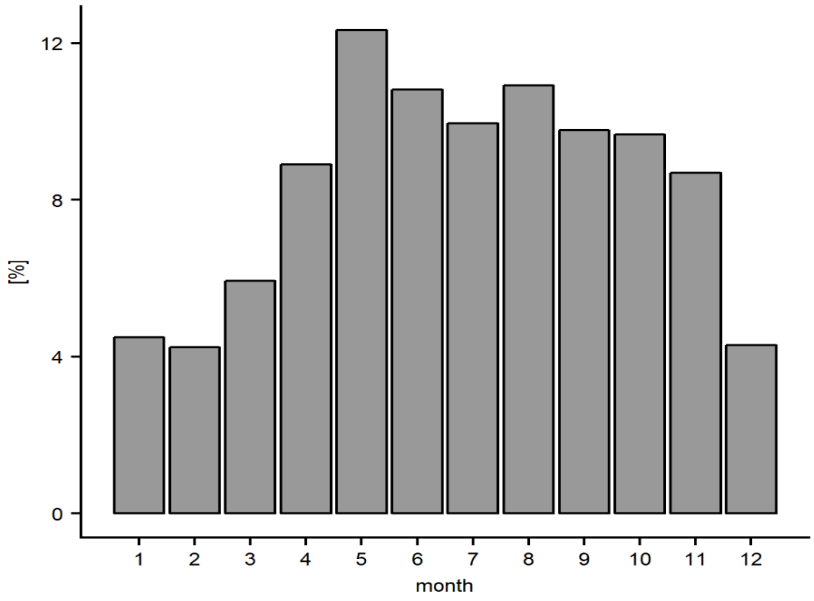

Figure 4.7: Yearly cycle of monthly precipitation normals over the study basin. The values are expressed as percentage of the annual precipitation climatology.

cent years. The evolution of BIAS is more regular and, except for the first two decades, the value is almost zero over the whole period.

In addition to reconstruction errors, the evolution of the correlation coefficients between simulated and measured station monthly anomalies was evaluated over subsequent 30-year intervals spanning the whole analysed period (Figure 4.9). It is worth noting that the median of correlation values is above 0.9 for all intervals and the agreement significantly increases from the second half of the 20^{th} century concurrently with the improvement in data coverage. The highest correlation coefficients occur during the 1981–2010 period when the integration of a relevant number of stations, especially at high-elevation, leads to a further reduction in the mean inter-station distance. On the contrary, besides the earliest period, the lowest correlation occurs in 1931–1960 period which probably reflect the overall temporary decrease in data availability and quality due to the difficulties in managing the meteorological network during the Second World War.

In order to evaluate the reconstruction accuracy of the gridding anomaly procedure and its evolution with the variability in station distribution, also over the early years when no records are available within the study domain, an iterative procedure was performed. The period 1801–1950 was divided in 30 consecutive 5-year intervals and the 1951–2000 monthly anomalies of all the stations in the study domain were iteratively reconstructed in LOO approach by considering only the records of the stations operating in each 5-year subperiod. The 1951–2000 period was chosen as reference as it corresponds to the years of best station coverage. The same halving distance coefficient already computed for each year whose data availability is considered (see section 4.2.2)

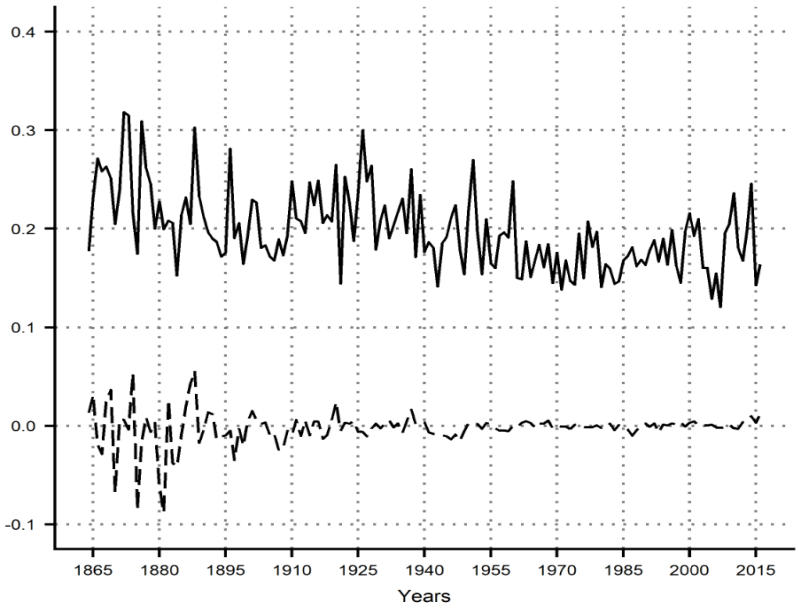

Figure 4.8: Evolution of annual MAE (solid line) and BIAS (dashed line) of leave-one-out reconstruction of the 1864–2016 station monthly anomalies available inside the study domain. The differences between simulated and measured station anomalies are considered.

was used and whenever a station record selected for the reconstruction did not cover the period 1951–2000 the anomalies of the nearest available station were used in order to prevent the reconstructions from being biased by the missing data. The distribution of MAE values computed by comparing the measured 1951–2000 monthly anomalies of each station in the study domain with the corresponding simulated series obtained by varying the data coverage is reported in Figure 4.10.

The errors reduce significantly from the 1861 availability, when the first stations started operating inside the study domain, and the MAE distributions remain almost stable with median values around 0.2 onwards. It is worth noting that even if the coverage before 1860 leads to larger errors, the distributions are similar, with medians around 0.4 in all cases, even for the lowest data coverage of the first years with very few stations entering in the interpolation.

The rather low errors obtained from the reconstruction of station anomalies over almost the whole validation period are mainly due to the large spatial coherence of anomalies which allow to reconstruct a reliable climatic signal even when the data density is lower. The distribution of the correlation between the anomaly records of all the station pairs within the study domain as function of their distance shows in fact correlation values above 0.7 even for distances greater than 70 km (Figure 4.11).

Finally, the accuracy in reconstructing the secular monthly precipitation records in absolute values was evaluated at station sites by converting the LOO monthly anomalies into millimeters by multiplying them times the corresponding LOO normals. The

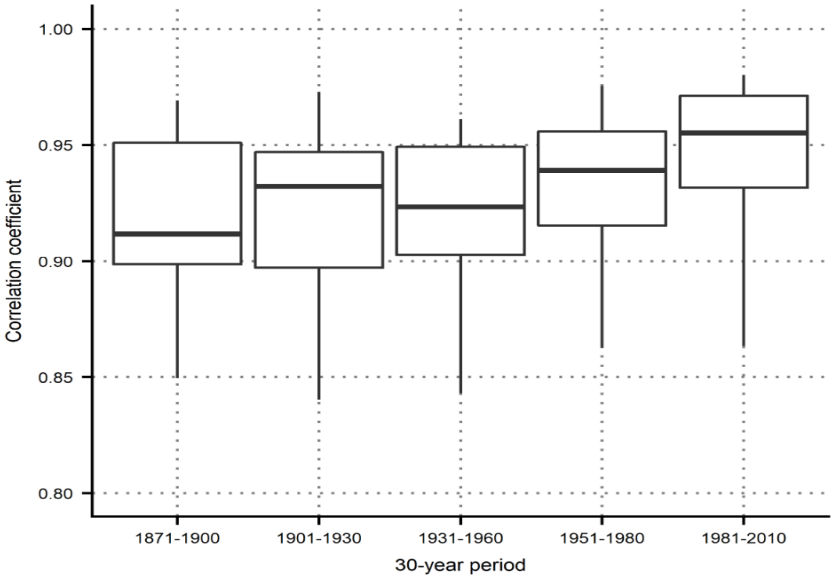

Figure 4.9: Distribution of correlation coefficients between simulated and measured monthly anomalies for the available stations inside the study domain over five subsequent 30-year periods. In each case, the correlation was computed only for the series with at least 80% of valid entries in the considered interval. The boxes represent the inter-quartile range of the distribution and the median is reported by the bold line; the whiskers represent the 5-95% quantile range.

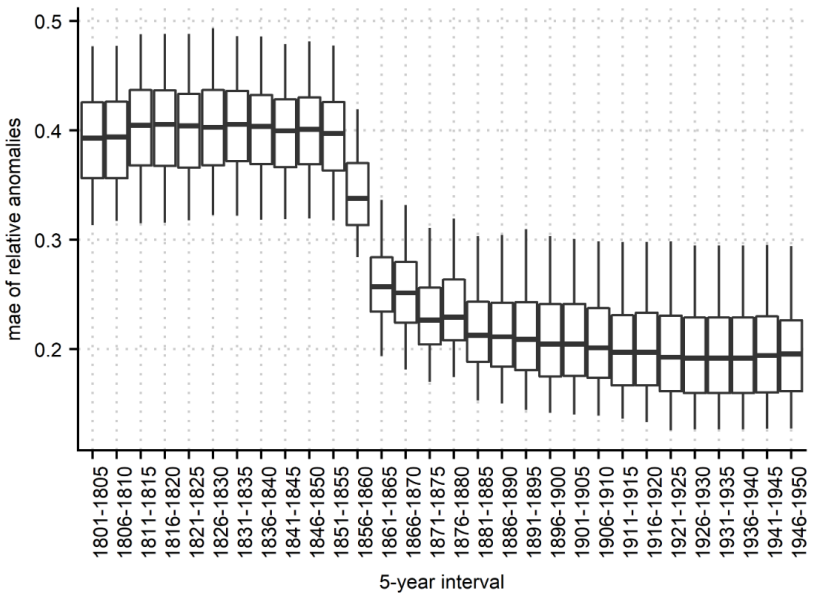

Figure 4.10: Distribution of MAE values obtained from the reconstruction of the 1951–2000 monthly anomalies of all the stations in the study domain by considering only the records from the rain-gauges available in each 5-year subperiod from 1801 to 1950. The boxes represent the inter-quartile range of the distribution and the median is reported by the bold line; the whiskers represent the 5-95% quantile range.

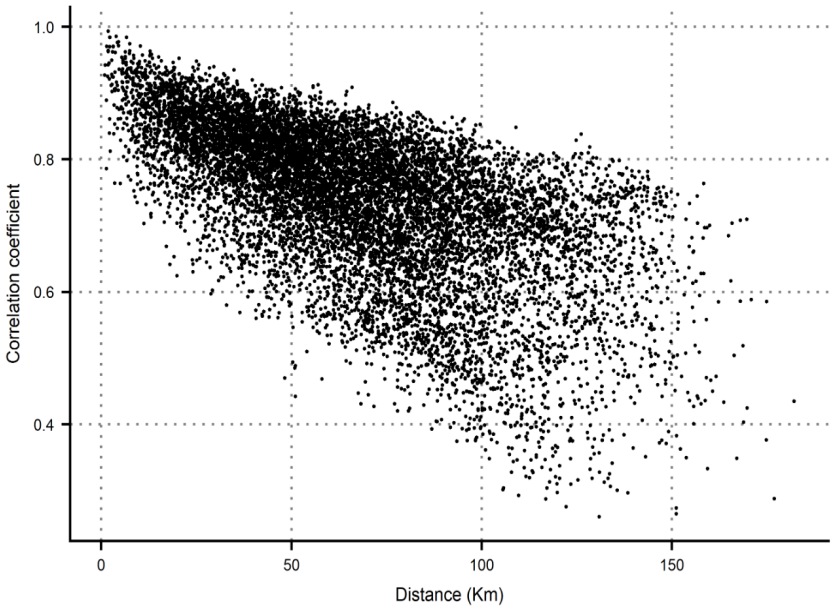

Figure 4.11: Distribution of correlation coefficients between all the station pairs within the study domain versus their distance. The correlation was evaluated only for the series with at least 30 years of valid entries.

largest fraction of MAPE values obtained by the comparison of the station simulations with observations over the 1864–2016 period ranges between 15.9% and 22.8% and the average value is 20.0%. The greatest errors, with MAPE values above 40%, are pointed out only for four stations all located in mountain environments. They show a systematical underestimation of monthly precipitation, which is particularly evident for the two records relative to the sites above 2000 m a.s.l. at San Bernardino Pass and Bernina Pass, respectively. The difficult reconstructions could be due to the fact that they are isolated sites on the tops of narrow valleys characterised by strong pluviometric gradients with very wet conditions over high-elevations and rapidly drier regimes over the bottoms, where most observations are located. This uneven data distribution in such complex conditions does not allow to fully capture the high-precipitation regime of the two sites in the LOO simulation. On the contrary, the precipitation series showing the greatest overestimations in reconstruction and MAPE above 30% are Trepalle, Fusine and Lago della Vacca. In this case, the low agreement with observations could be ascribed to both a possible underestimation of solid precipitation by rain-gauges and the low representativeness of the network for the areas in which they are located, featuring relevant altitude gradients and heterogenous climatic conditions with sharp transitions of pluviometric regime within very short distances.

The 1800–2016 monthly areal precipitation record of the study basin: record reconstruction and validation

The 1800–2016 monthly areal precipitation record of the upper part Adda river basin was obtained by summing for each time step the precipitation contributions from all its grid cells scaled for their areas:

$$q_t = \sum_j p_{j,t} \cdot S_j \cdot f_j \tag{4.1}$$

where q_t is the areal precipitation for the time step t, $p_{j,t}$ is the precipitation reconstructed for the time step t at the grid cell j and S_j and f_j are the cell area and the fraction of cell contained in the basin, respectively. In order to express the precipitation record in millimeters, the computed values were finally normalised by the basin total surface. The availability of the gridded precipitation dataset allowed to assess the study basin total precipitation at any time step over the whole considered period. It is worth underlining that this information can be retrieved by means of the anomaly method interpolation chain and not directly from the absolute station records only: if the monthly precipitation fields were estimated by simply interpolating the monthly records of available stations, the study basin areal precipitation would be significantly underestimated, at least before 1951. The underestimation results particularly high before the 1860s of about 30% and decreases to about 10% in the following years until 1900. After that the bias gradually decreases, even though it remains above 5% until 1951. This outcome reflects the fact that the station distribution before 1951 is biased towards low level areas, where also precipitation normals are generally lower than on higher elevation regions. The anomaly method integrates the spatial precipitation gradients contained in the climatological fields allowing to avoid the bias due to the uneven station coverage in the resulting areal precipitation records for the study basin. However, even though unbiased over the whole study period, the confidence on this record varies significantly over time. In order to evaluate the evolution of uncertainty in relation to the variation of the rain-gauge coverage over the study period, we used the same iteratively reconstruction procedure applied to station anomalies and described in section 3.1. Since the actual monthly total precipitation series over the basin is unknown, the estimated values over the 50-year of best data availability (1951–2000) were considered as reference as they are expected to be characterised by the highest reconstruction robustness. For each 5-year subperiod from 1801 to 1950, the 1951–2000 monthly precipitation record of the basin was reconstructed by computing the monthly anomaly fields only from the data of the available stations in the considered interval. At each iteration, the gridded dataset was rescaled to absolute values by means of the 1961–1990 climatologies and equation 4.1 was applied to get the areal precipitation es-

timates which were compared to the reference one. In particular, MAPE values were computed for each iteration as follows:

$$MAPE_i = \frac{1}{T} \cdot \sum_{t=1}^{T} \frac{|\tilde{p}_{t,i} - p_t|}{p_t} \cdot 100 \tag{4.2}$$

where $\tilde{p}_{i,t}$ is the simulated basin precipitation for time step t by considering the station availability in subperiod i, p_t is the reference precipitation value at time t, T is the total length of the series and t runs from January 1951 to December 2000.

The error evolution is very similar to that obtained from the validation on station anomalies as shown in Table 4.2. The MAPE computed on annual scale is about 10% for all the station coverages before 1855, it decreases to 7% by considering the data density of 1856–1860 and to 4% with the availability of 1861–1865 when the first stations inside the study domain started operating. The annual error reduces further to 2% with the station availability from 1881 and to 1% by taking into account the network coverage from 1921. The correlation coefficients between 1951–2000 simulated and reference precipitation records on annual scale show the same behavior as MAPE: the correlation for the reconstruction with the first data distribution is 0.72 and it remains substantially invariant until the station coverage of 1856–1860 for which the correlation increases to 0.87. By taking into account the station availability from 1861 the correlation turns out to be 0.96 and it increases further with the development of the network, especially from 1911, when the rain-gauge distribution allows to reach a correlation coefficient of 0.99 between simulated and reference values.

It is worth noting that the robustness of the reconstruction with the low observation availability in the first decades is relatively high and errors are constantly around 10% even when less than ten stations are used for the interpolation. In order to assess the meaningfulness of the information on the study basin we get from these few stations all outside the study domain, we extracted 1000 sets of 50 random values from a normal distribution with the same mean and standard deviation as the 1951–2000 annual precipitation series and the corresponding MAPEs were computed. The average error turns out be 20%, which proves the consistency of the estimated accuracy for the reconstructions performed by means of the station availability of the early period and gives evidence of the ability of the database to provide a reliable climate signal over the basin even for the most remote years by exploiting all the available observations.

Very similar outcomes are depicted also on seasonal scale. The lowest agreement is pointed out for winter with MAPE of about 30% for the data availability of the first decades, when all the rain-gauges were located at low elevation, and it decreases below 5% only after 1910. For all the other seasons, the reconstruction errors are below 5% already from the 1881, when only 4 stations were operating inside the study basin, and they generally reach 2% by using the data distribution after 1910.

	YEAR		DJF	MAM	JJA	SON
5-YEAR DATA AVAILABILITY	MAPE [%]	COR	MAPE [%]			
1801-1805	9.5	0.72	28.8	23.4	15.5	25.2
1821-1825	9.9	0.69	29.4	23.9	17.2	25.2
1841-1845	9.9	0.72	29.2	24.4	16.7	22.8
1856-1860	7.1	0.87	22.7	18.7	12.3	18.3
1861-1865	3.8	0.96	16.5	7.1	6.3	7.8
1881-1885	2.0	0.99	7.4	3.8	3.7	4.2
1901-1905	2.2	0.99	8.0	3.4	3.7	4.1
1911-1915	1.6	0.99	4.3	2.3	2.4	2.7
1921-1925	1.1	>0.99	3.4	1.6	1.7	1.9

Table 4.2: Annual and seasonal MAPE values of the reconstructed 1951–2000 total precipitation series by considering the data availability in some selected 5-year subperiods from 1801. For the annual simulations only, the correlation with the values obtained from the actual 1951– 2000 data availability is reported too.

Variability and trends of the secular precipitation record of Adda basin

The 1800–2016 annual and seasonal precipitation records for the study basin were analysed for long-term trends by performing the Theil-Sen test (Theil, 1950), while the significance of trends was assessed by the Mann-Kendall test (Kendall, 1938). The annual and seasonal series in absolute values are reported in Figures 4.12 and 4.13, together with the corresponding Theil-Sen fits and the 11-year Gaussian filters with 3-year standard deviation. Considering a confidence level of 0.05, a significant negative annual trend of -5.1 mm decade^{-1} was pointed out, while, on seasonal scale a negative tendency is depicted in all cases even though the trend turns out to be significant only for the autumn precipitation series with a loss of 4.2 mm decade^{-1}. The negative trend depicted for the whole period is likely to be largely driven by the wet conditions characterising the beginning of the 19^{th} century. In addition, as discussed in the previous section, the first period of the series is affected by a higher uncertainty which has to be taken into account in the evaluation of the whole trend. By excluding the first decades in fact the annual trend remains negative, but its significance decreases far above 0.05, suggesting that future analyses are needed to further investigate the actual occurring long-term tendency in annual total precipitation over the basin. On the contrary, the negative trend in autumn precipitation series remains significant even by excluding the first decades, suggesting a possible long-term change in the precipitation regime of the autumn season, which requires to be further analysed in future studies.
In order to investigate the variability of precipitation over the basin on a finer timescale, a running-trend analysis was performed on the 1800–2016 annual and seasonal records. More precisely, the Theil-Sen slopes and Mann-Kendall significance levels were computed on windows of increasing width from 20 years up to the entire period spanned by the series and running from the beginning to the end of the record. The

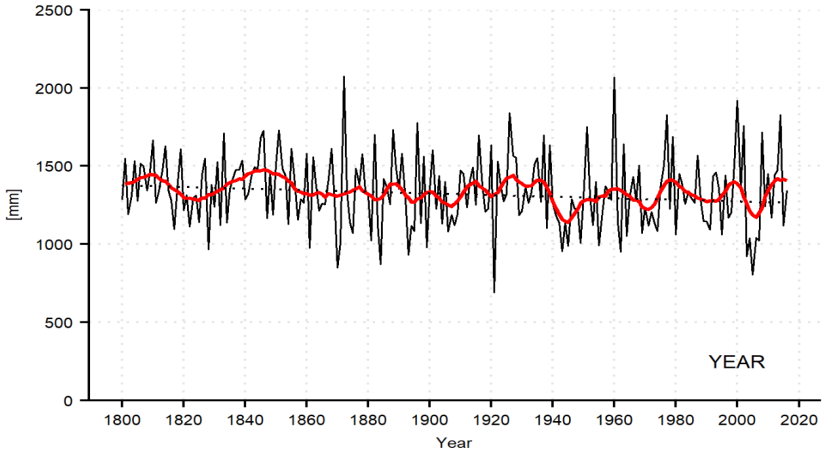

Figure 4.12: 1800–2016 annual series of total precipitation of upper Adda river basin (solid black line) together with a 11-year Gaussian filter with 3-year standard deviation (red solid line) and the Theil-Sen linear fit (dotted line).

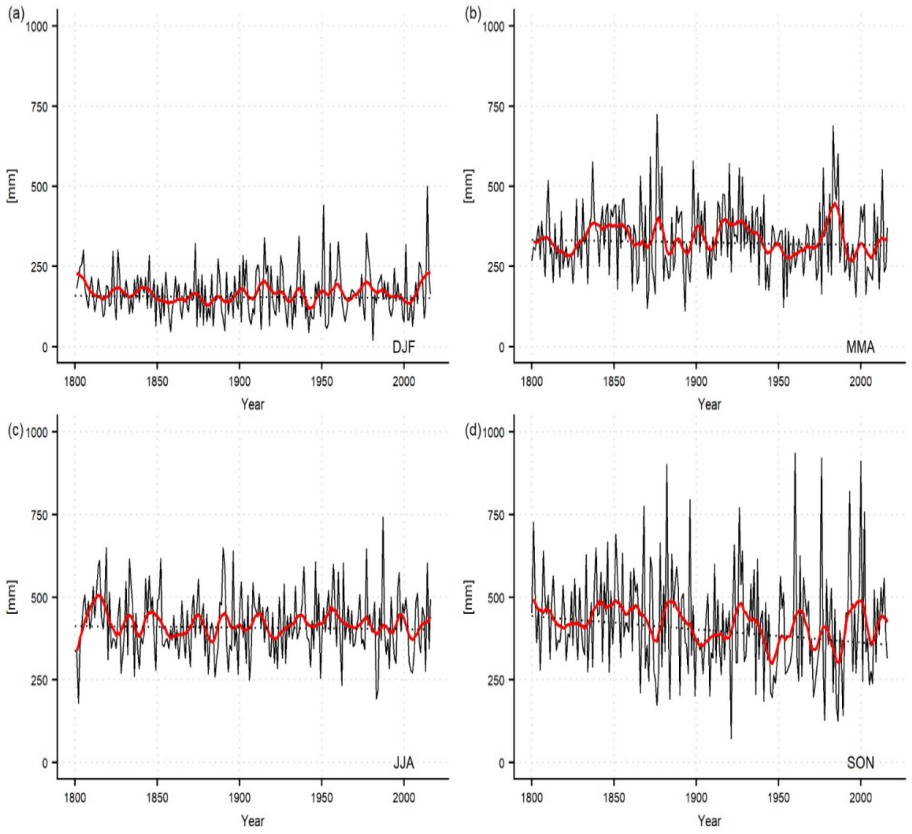

Figure 4.13: 1800–2016 seasonal series of total precipitation of upper Adda river basin (solid black lines) together with 11-year Gaussian filters with 3-year standard deviation (red solid lines) and the Theil-Sen linear fits (dotted lines).

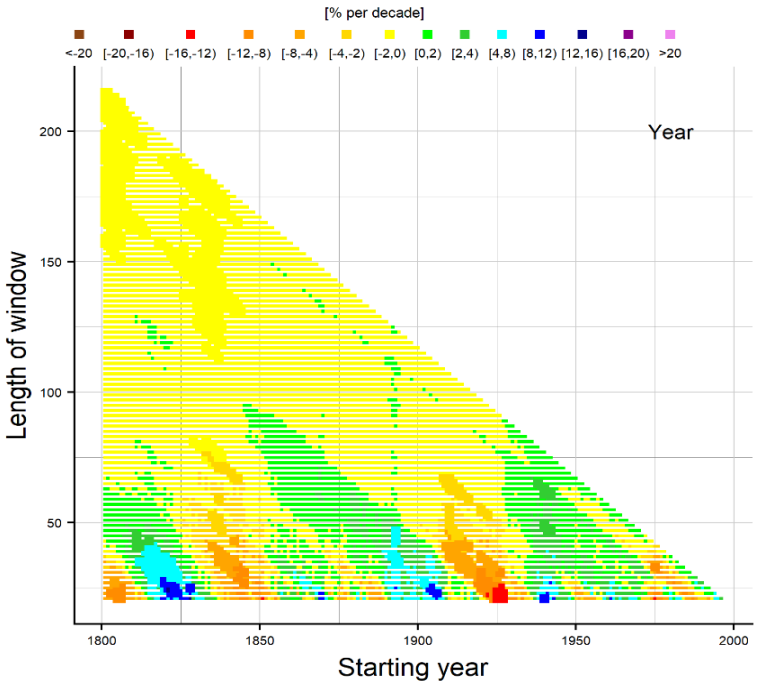

Figure 4.14: Running trend of annual precipitation anomalies. Trend values are expressed by colors while the significance of trend is represented by the size of the pixels.

running trend on annual values (Figure 4.14) confirms a negative long-term tendency which assumes statistically significant values only if almost the whole period is included in the evaluation. A more relevant variability in precipitation regime is evident at shorter time scales with a sequence of wetting and drying conditions, especially in the first half of the 20^{th} century. A high frequency variability was also pointed out for seasonal records (Figure 4.15), while only the autumn series shows a significant drying tendency at long-time scale, which persists even varying the beginning of the considered window. It is interesting to note the strong negative trend in spring precipitation series starting from the 1910s that is evident also for time windows longer than 50 years. This trend is due to the rather lower values from the 1940s which are interrupted by a relevant increase in the 1970s (Figure 4.13b). After this decade, however, spring precipitation returns to be rather low and it shows significant negative trends even over 100-year long periods starting around the 1910s.

Comparison of basin precipitation and runoff series

The computed study basin annual precipitation series was compared to the available annual runoff record covering the period 1845–2016. The comparison was performed by considering the hydrological year, i.e. from September to August, in order to take into account the fraction of precipitation falling as snow during autumn and winter which is released during the following spring and summer months. This is also the 12-month period providing the highest correlation between precipitation and runoff.

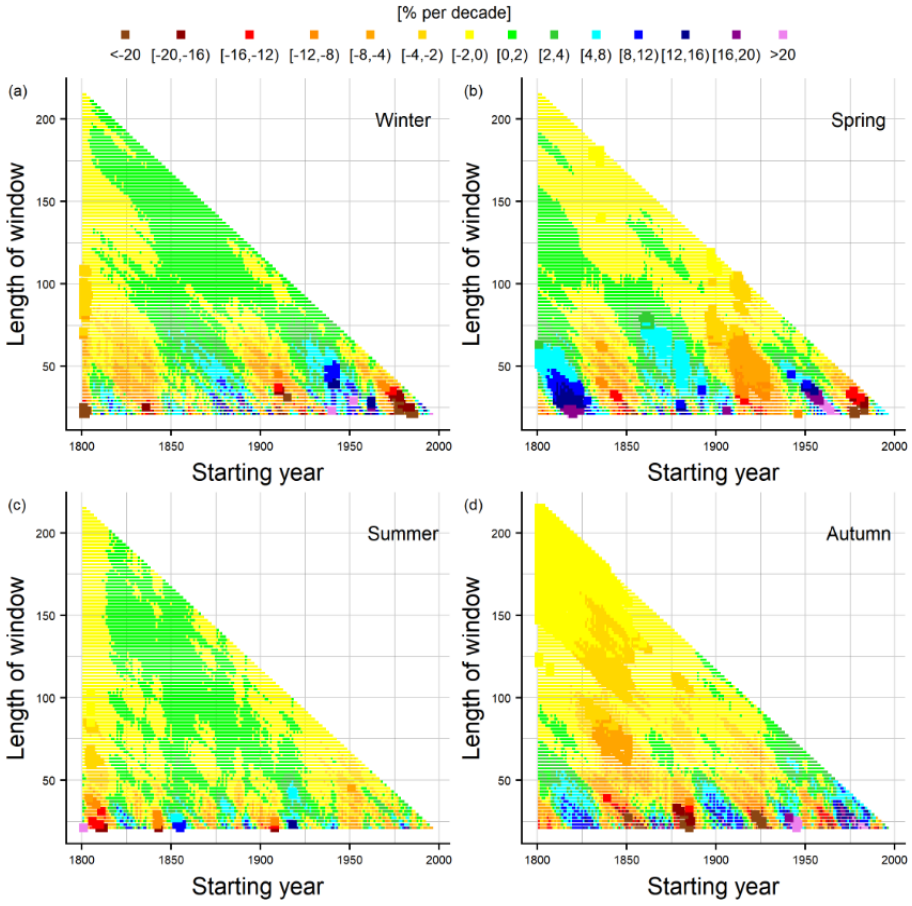

Figure 4.15: Running trend of seasonal precipitation anomalies. Trend values are expressed by colors while the significance of trend is represented by the size of the pixels.

The correlation coefficient between the annual series passes from 0.82 if solar year is considered (January-December) to 0.91 if the 12-month totals over the defined hydrological year are compared.
This comparison highlighted that the mean annual runoff coefficient, i.e. the the ratio between runoff and precipitation, on the whole period (September 1845 – August 2016, hereafter 1846–2016) is 0.867 corresponding to about 170 mm of precipitation losses per year by means of evapotranspiration which is the main mechanism acting on water loss from the basin. A previous study at specific sites located in Valtellina suggests that lower runoff coefficients should be expected for the area (Ranzi and Bacchi, 1998), moreover other literature works focusing on several basins of the Alpine region suggest that a more realistic estimate of the actual evapotranspiration in the study basin should range between 300 and 400 mm (Gurtz et al., 1999).
The rather high value we found for the mean annual runoff coefficient could be ascribed to an underestimation of the basin precipitation climatologies, which may reflect the difficulties of rain-gauges in correctly measuring precipitation, especially in presence of solid precipitation. This is a well-known problem, particularly affecting the Alpine region precipitation dataset. Isotta et al. (2014) warn about the contribution of rain-gauge undercatch in the uncertainty of gridded precipitation fields, especially during episodes of strong wind. The magnitude of the undercatch depends in fact on the wind speed, the precipitation type, rain or snow, and intensity, however it could account for up to several 10 percents (Sevruk, 1985; Richter, 1995; Frei and Schär, 1998). Ranzi et al. (1999) showed that the underestimation of snow precipitation for a rain-gauge network covering the area of the main northern tributaries to the Po river in Piedmont, Lombardy and Trentino-Alto Adige, thus including also the study basin, could range between 30% and 70% depending on the location.
It is however worth noting that the recent temperature increase could lead to a lower impact of gauge undercatch by reducing the fraction of precipitation falling in solid form in respect with the past. We could therefore expect that the actual long-term trend of basin precipitation could be more negative than the observed one. A detailed analysis aiming at assessing the influence of the undercatch variation over the years on the resulting long-term trend of precipitation could represent a very interesting issue to be investigated in the future.
Besides the undercatch of precipitation especially at high elevation, another source of errors in the precipitation climatologies may also rely on the fact that station coverage over the study area is not fully representative of the orography, with a relevant reduction of available in-situ observations towards the high-elevation areas. This uneven data distribution could influence the evaluation of precipitation-orography relationship performed by LWLR interpolation of climatologies where the higher number of low-elevation sites entering in the regression at grid cells over mountainous regions

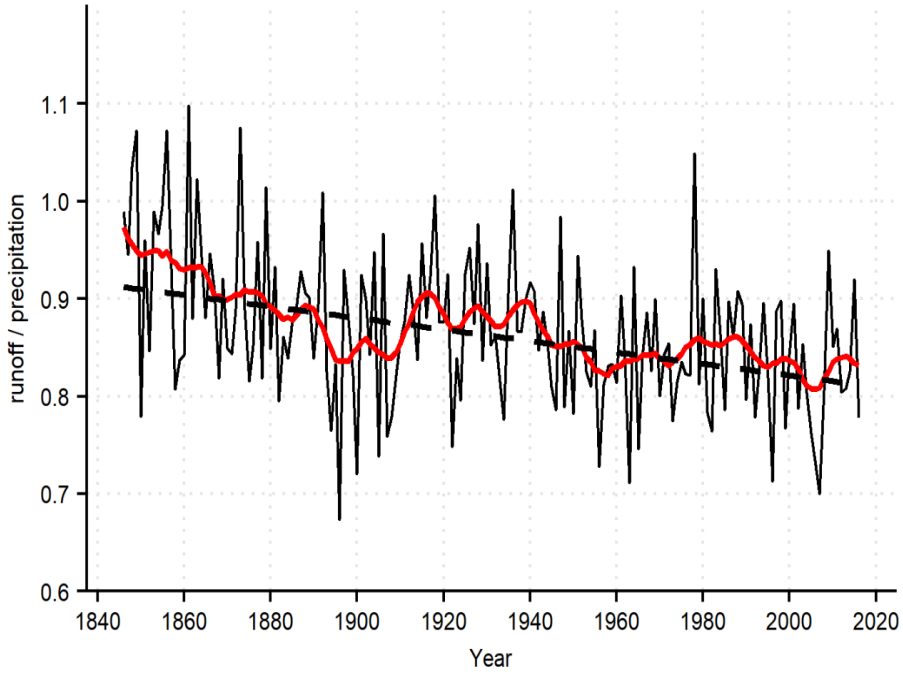

Figure 4.16: Evolution of the annual (hydrological year) runoff coefficient (black solid line) together with the 11-year Gaussian filter with 3-year standard deviation (red solid line) and the Theil-Sen linear fit (dashed line).

could smooth out the actual elevation-precipitation gradients. This problem is found to lead to large errors in the climatologies of high-level areas when the station network is strongly biased towards low-level areas as recently highlighted by Lussana et al. (2018) and Crespi at al. (2019) for Norway. Over the study basin, however, the station coverage is reasonable at least up to about 2000 m so that the problem of sparse network is expected to influence more the accuracy of the single grid cell estimates rather than to produce systematic errors. For this reason, the averaging over the entire study basin should reduce further this problem.

The comparison between basin precipitation and runoff pointed out that the yearly runoff coefficient is not constant over time (Figure 16). The trend analysis was performed over both the entire record and on shorter time intervals by applying the running trend evaluation (Figure 4.17). A significant negative trend of about -0.6% decade^{-1} was pointed out over the whole period. The running trend points out the occurrence of significant negative tendency for nearly all windows at least one century long. On shorter time scales significant decreases are also evident for windows starting around 1850 and in the 1910–1930 period. Strong decreases are finally evident for windows starting in the 1930s and in the 1980s and spanning 20-30 years.

Performing the trend evaluation on both annual runoff and precipitation over 1846–2016 period, runoff series experiences a clear negative trend of -10.6 mm decade^{-1} and a Mann Kendall p-value below 0.01, while precipitation values show

a decreasing tendency but not statistically significant.
One of the main causes of long-term runoff losses could be ascribed to the increase in evapotranspiration induced primarily by temperature increase. Another contribution to the variation of the evapotranspiration rate over the study basin could derive from changes in land-use occurred over the last century, with the expansion of forest coverage enhanced by the gradual decline of human communities and pastures over the mountainous areas (Guidi et al, 2014). Also the construction of artificial reservoirs could have an impact on the variation of measured runoff by modifying the natural hydrological processes. In particular, the large number of reservoirs built for electricity production from 1930 to 1970 reaching about $500 \cdot 10^6$ m^3 of total storage volume (Consorzio dell'Adda, 2003), which corresponds to slightly less than 10% of mean annual runoff, could have influenced the variability of runoff to precipitation ratio. In addition, the recent enhancement of glacier melting is likely to slightly mask the actual long-term tendency in the evolution of runoff coefficient which, without the contribution from glacier melting, could be even more negative than observed.
Looking at Figure 4.16, it is worth noting the presence of some years in which the runoff coefficient is greater than one, even if runoff excess is less than 10% in all cases. Most of these events occur before 1880 when the uncertainty on both runoff and precipitation reconstructions is higher, however, a sporadic excess of runoff could be produced by the many factors acting on the hydrological cycle on different time scales. In particular, the total runoff for a certain year could be significantly influenced by the precipitation that occurred in the previous hydrological year, especially in last months.

4.2.4 Conclusions

The 1800–2016 monthly areal precipitation series for the upper part of the Adda river basin was computed by means of the anomaly method and a quality-checked and homogeneous database of monthly station records spanning more than two centuries and covering a wide area centred on the study domain. More specifically, the monthly station anomalies in respect with the 1961–1990 normals were interpolated onto a 30-arc second resolution grid by a weighted averaging approach where station weights depend on distance and elevation difference from the target cell. The fields of 1800–2016 precipitation series in mm were finally obtained by multiplying the monthly anomaly fields times the gridded 1961–1990 climatologies which were computed by the LWLR scheme.
The errors of anomaly and climatology reconstructions were evaluated separately by means of the LOO validation on station data. In both cases, the greatest errors were observed for winter months which could be partly ascribed to a general difficulty in correctly measuring the solid precipitation, especially at the high-level sites. As re-

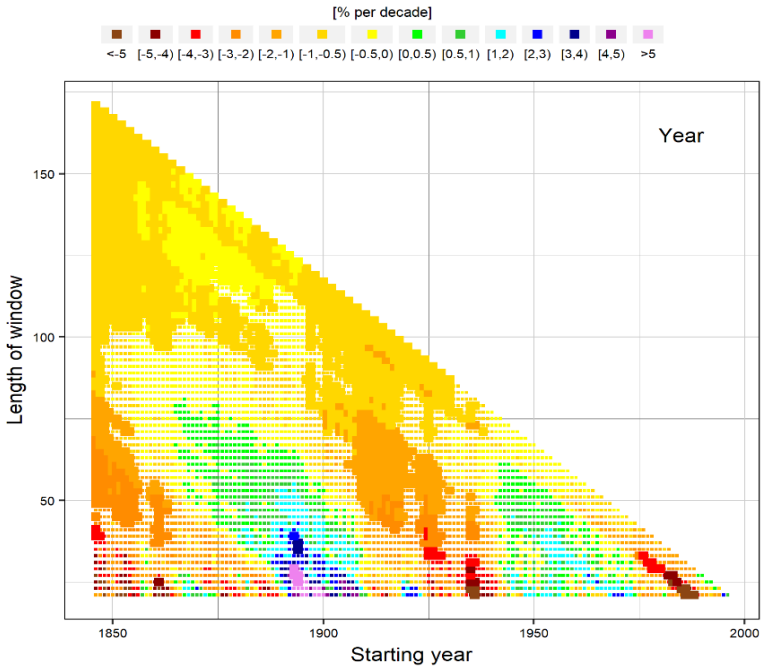

Figure 4.17: Running trend of annual runoff coefficient (hydrological year). Trend values are expressed by colors while the significance of trend is represented by the size of the pixels.

gards the temporal evolution of the anomaly reconstruction accuracy, MAE is about 0.2 already from the 1880s and BIAS is almost null over the whole analysed period. The evolution of the accuracy of areal precipitation estimates in respect with the variability in data coverage was evaluated too and MAPE on annual values turned out to rapidly decrease from about 10% over the first decades to 4% from the 1860s and to 2% from the 1880s. These rather low errors, even for the years of sparse data coverage, are obtained thanks to the anomaly-based method which has been proved to avoid significant underestimations in areal annual precipitation which occur by interpolating the station absolute values directly.

Significant negative trends were found for the 1800–2016 annual and autumn areal basin precipitation records only, even if the significance of the annual signal decreases far above 0.05 by excluding the first decades.

The comparison between the 1845–2016 basin precipitation and runoff records on hydrological year showed a high correlation. However, the resulting mean annual runoff index of 0.876, corresponding to about 170 mm of precipitation losses per year, was found to be higher than expected from literature for the area suggesting a possible underestimation of basin precipitation. The bias in the precipitation reconstruction could be partly due to the rain-gauge undercatch, especially in mountainous areas, and the uneven station coverage decreasing towards the high-level regions, even though further analyses are required in order to better assess the sources of uncertainty.

The Theil-Sen trend on runoff coefficient series showed a long-term significant decrease of -0.6% decade^{-1} corresponding to a relevant decrease in annual runoff of almost -11 mm decade^{-1}, as a consequence of the increase in evapotranspiration. Both warming and changes in land use and coverage could contribute to evapotranspiration variation. In addition, a small contribution to runoff could derive from the accelerated melting of basin Alpine glaciers during the recent decades which could slightly reduce the actual runoff losses. A more detailed discussion of natural and anthropogenic contributions to the variation of the hydrological cycle of the study basin will be provided in forthcoming papers.

4.2.5 References

- Auer, I., Böhm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schöner, W., Ungersboeck, M., Matulla, C., Briffa, K., Jones, P., Efthymiadis, D., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisseline, J. M., Begert, M., Müller-Westermeier, G., Kveton, V., Bochnicek, O., Stastny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic-Capka, M., Zaninovic, K., Majstorovic, Z., and Nieplova, E. (2007). HISTALP – Historical instrumental climatological surface time series of the Greater Alpine Region HISTALP. International Journal of Climatology, 27, 17–46. doi:10.1002/joc.1377

- Böhm, R, Auer, I., Brunetti, M., Maugeri, M., Nanni, T., and Schöner, W. (2001). Regional temperature variability in the European Alps 1760–1998 from homogenized instrumental time series. International Journal of Climatology, 21, 1779–1801. doi:10.1002/joc.689

- Brugnara, Y., Brunetti, M., Maugeri, M., Nanni, T., and Simolo, C. (2012). High-resolution analysis of daily precipitation trends in the central Alps over the last century. International Journal of Climatology, 32, 1406-1422. doi:10.1002/joc.2363.

- Brunetti, M., Maugeri, M., Nanni, T., Auer, I., Böhm, R., and Schöner, W. (2006a). Precipitation variability and changes in the greater Alpine region over the 1800–2003 period. Journal of Geophysical Research, 111, 111, D11107, doi:10.1029/2005JD006674

- Brunetti, M., Maugeri, M., Monti, F., and Nanni, T. (2006b). Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. International Journal of Climatology, 26, 345-381. doi:10.1002/joc.1251

- Brunetti, M., Lentini, G., Maugeri, M., Nanni, T., Auer, I., Böhm, R., and Schöner, W. (2009). Climate variability and change in the Greater Alpine Region over the last two centuries based on multi-variable analysis. International Journal of Climatology, 29, 2197-2225. doi:10.1002/joc.1857

- Casty, C., Wanner, H., Luterbacher, J., Esper, J., and Böhm, R. (2005). Temperature and precipitation variability in the European Alps since 1500. International Journal of Climatology, 25, 1855-1880. doi:10.1002/joc.1216

- Craddock, J. (1979). Methods of comparing annual rainfall records for climatic purposes. Weather, 34, 332-346. doi:10.1002/j.1477-8696.1979.tb03465.x

- Crespi, A., Brunetti, M., Lentini, G., and Maugeri, M. (2018). 1961-1990 high-resolution monthly precipitation climatologies for Italy. International Journal of Climatology, 38, 878–895. doi:10.1002/joc.5217

- Crespi, A., Lussana, C., Brunetti, M., Dobler, A., Maugeri, M., and Tveito, O. E. (2019). High-resolution monthly precipitation climatologies over Norway (1981–2010): joining numerical model datasets and in-situ observations. International Journal of Climatology (submitted).

- D'Agata, C., Bocchiola, D., Soncini, A., Maragno, D., Smiraglia, C., Diolaiuti, G. A. (2018). Recent area and volume loss of Alpine glaciers in the Adda River of Italy and their contribution to hydropower production. Cold Regions Science and Technology, 148, 172-184. doi:10.1016/j.coldregions.2017.12.010

- Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P. (2002). A knowledge-based approach to the statistical mapping of climate. Climate Research, 22, 99-113. doi:10.3354/cr022099

- Daly, C. (2006). Guidelines for assessing the suitability of spatial climate data sets. International Journal of Climatology, 26, 707-721. doi:10.1002/joc.1322

- Durand, Y., Laternser, M., Giraud, G., Etchevers, P., Lesaffre, B., and Mérindol, L. (2009). Reanalysis of 44 Yr of Climate in the French Alps (1958–2002): Methodology, Model Validation, Climatology, and Trends for Air Temperature and Precipitation. Journal of Applied Meteorology and Climatology, 48, 429–449. doi:10.1175/2008JAMC1808.1

- Consorzio dell'Adda (2003). Gli effetti della regolazione sulle portate dell'Adda e sulle piene del lago di Como. Vol. 12, Milano.

- EEA (European Environment Agency). 2009. Regional climate change and adaptation. The Alps facing the challenge of changing water resources. EEA Technical Report, 143 pp. doi:10.2800/12552.

- Efthymiadis, D., Jones, P. D., Briffa, K. R., Auer, I., Böhm, R., Schöner, W., Frei, C., and Schmidli, J. (2006). Construction of a 10-min-gridded precipitation data set for the Greater Alpine Region for 1800–2003. Journal of Geophysical Research, 111, D01105, doi:10.1029/2005JD006120

- Foresti, L., Sideris, I., Panziera, L., Nerini, D., and Germann, U. (2018). A 10-year radar-based analysis of orographic precipitation growth and decay patterns over the Swiss Alpine region. Quarterly Journal of the Royal Meteorological Society. doi:10.1002/qj.3364

- Frei, C., and Schär, C. (1998). A precipitation climatology of the Alps from high-resolution rain-gauge observations. International Journal of Climatology, 18, 873-900. doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9

- Garavaglia, R., Marzorati, A., Confortola, G., Bocchiola, D., Cola, G., Manzata, E., Senese, A., C. Smiraglia, C., and Diolaiuti, G.A. (2014). Evoluzione del Ghiacciaio dei Forni. Neve & Valanghe, 81, 60-67.

- Gyalistras, D. (2003). Development and validation of a high-resolution monthly gridded temperature and precipitation data set for Switzerland (1951-2000). Climate Research, 25, 55-83. doi:10.3354/cr025055

- Golzio, A., Crespi, A., Bollati, I. M., Senese, A., Diolaiuti, G. A., Pelfini, M., and Maugeri, M. (2018). High-Resolution Monthly Precipitation Fields (1913–2015) over a Complex Mountain Area Centred on the Forni Valley (Central Italian Alps). Advances in Meteorology, 2018, ID 9123814, pp. 17. doi:10.1155/2018/9123814

- Guidi, C., Vesterdal, L., Gianelle, D., and Rodeghiero, M. (2014). Changes in soil organic carbon and nitrogen following forest expansion on grassland in the Southern Alps. Forest Ecology and Management, 328, 103-116. doi:10.1016/j.foreco.2014.05.025

- Gurtz, J., Baltensweiler, A., and Lang, H. (1999). Spatially distributed hydrotope-based modelling of evapotranspiration and runoff in mountainous basins. Hydrological Processes, 13, 2751-2768. doi:10.1002/(SICI)1099-1085(19991215)13:17<2751::AID-HYP897>3.0.CO;2-O

- Isotta, F.A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertaničk, 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. International Journal of Climatology, 34, 1657-1675. doi:10.1002/joc.3794

- Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. Journal of Geophysical Research, 113, D20119. doi:10.1029/2008JD010201

- Kendall, M. G. (1975). Rank Correlation Methods. Griffin, London.

- Lussana, C., Saloranta, T., Skaugen, T., Magnusson, J., Tveito, O. E., and Andersen, J. (2018). seNorge2 daily precipitation, an observational gridded dataset over Norway from 1957 to the present day. Earth System Science Data, 10, 235–249. doi:10.5194/essd-10-235-2018

- Masson, D., and Frei, C. (2016). Long-term variations and trends of mesoscale precipitation in the Alps: recalculation and update for 1901-2008, International Journal of Climatology, 36, 492-500. doi:10.1002/joc.4343

- Mitchell, T. D., and Jones, P. D. (2005). An improved method of constructing a database of monthly climate observations and associated highresolution grids. International Journal of Climatology, 25, 693–712. doi:10.1002/joc.1181

- New, M., Todd, M., Hulme, M., and Jones, P. (2001). Precipitation measurements and trends in the twentieth century. International Journal of Climatology, 21, 1899–1922. doi:10.1002/joc.680

- Pavan, V., Antolini, G., Barbiero, R., Berni, N., Brunier, F., Cacciamani, C., Cagnati, A., et al. (2018). Climate Dynamics, 1-19. doi:10.1007/s00382-018-4337-6

- Ranzi, R., Caronna, P., and Tomirotti, M. (2017). Impact of climatic and land use changes on riverflows in the Southern Alps, in: "Sustainable Water Resources Planning and Management Under Climate Change", E. Kolokytha, S. Oishi, R.S.V. Teegavarapu (Eds.), Springer Science+Business Media, Singapore, 61-83. doi:10.1007/978-981-10-2051-3

- Ranzi, R., Grossi, G., and Bacchi, B. (1999). Ten years of monitoring areal snowpack in the Southern Alps using NOAA-AVHRR imagery, ground measurements and hydrological data. Hydrological Processes, 13: 2079-2095. doi:10.1002/(SICI)1099-1085(199909)13:12/13<2079::AID-HYP875>3.0.CO;2-U

- Ranzi, R., and Bacchi, B.(1998). Il bilancio idrologico nelle aree montane per la stima delle disponibilita idriche: alcuni problemi aperti. XXVI Convegno di Idraulica e Costruzioni Idrauliche, Catania, 3, 347-358.

- Richter, D. (1995). Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Messfehlers des Hellmannniederschlagsmessers, Vol. 194, Berichte des Deutschen Wetterdienstes.

- Scherrer, S. C., Fischer, E. M., Posselt, R., Liniger, M. A., Croci-Maspoli, M., and Knutti, R. (2016). Emerging trends in heavy precipitation and hot temperature extremes in Switzerland. Journal of Geophysical Research Atmosphere, 121, 2626–2637. doi:10.1002/2015JD024634

- Schmidli, J., Schmutz, C., Frei, C., Wanner, H., and Schär, C. (2002). Mesoscale precipitation variability in the region of the European Alps during the 20th century. International Journal of Climatology, 22, 1049– 1074. doi:10.1002/joc.769

- Servizio Idrografico (1920). Osservazioni pluviometriche raccolte a tutto l'anno 1915. Volume II: Bacino Imbrifero del Po. Fascicolo I. Ministero dei Lavori Pubblici – Consiglio Superiore delle Acque, Roma.

- Servizio Idrografico (1925). Osservazioni pluviometriche raccolte nel quinquennio 1916–1920. Pubblicazione numero 1 del Servizio Idrografico. Volume V: Bacino del Po, Roma.

- Servizio Idrografico (1959). Precipitazioni medie mensili ed annue e numero dei giorni piovosi per il trentennio 1921–1950. Pubblicazione numero 24 del Servizio Idrografico. Fascicolo XII b, Roma.

- Sevruk, B. (1985). Correction of precipitation measurements. Proceedings of Workshop on the Correction of Precipitation Measurements, Zurich, Switzerland, WMO/IAHS/ETH, 13–23.

- Sevruk, B., Ondrás, M., and Chvíla, B. (2009). The WMO precipitation measurement intercomparisons. Atmospheric Research, 92. 376-380. doi:10.1016/j.atmosres.2009.01.016

- Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. I. Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, A53, 386–392.

- Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., and Weingartner, R. (2007). Mountains of the world, water towers for humanity: Typology,

mapping, and global significance. Water Resources Research, 43, W07447. doi:10.1029/2006WR005653

Chapter 5

Variability and trends in monthly precipitation over Sardinia

The chapter presents the results obtained from the research activities performed over Sardinia in order to provide a high-resolution description of the spatial distribution of monthly precipitation normals over one of the main Mediterranean islands and to assess the precipitation evolution over the recent past. 1961-1990 monthly precipitation climatologies at 30-arc second resolution were estimated for the region and the 1922-2011 station records were then projected onto the same grid by the anomaly-based approach. At this aim, more than 300 monthly records were retrieved from hardcopy archives and subjected to quality assessment and homogenisation procedures. The 1922-2011 average monthly series for the region was computed and analysed for long-term trends, while the annual climatologies over consecutive 30-year periods from 1931 to 2010 were reconstructed and compared in order to assess the decadal variability in the spatial distribution of precipitation normals over the last century. The results were presented at the 6^{th} International Conference on Meteorology and Climatology of the Mediterranean and they will be discussed more in detail in a forthcoming paper.

5.1 Introduction

The availability of high-resolution datasets of precipitation records spanning long periods is a crucial requirement in order to perform detailed studies on the evolution and changes in water resources and climate variations over specific regions. Gridded datasets allow in fact to reduce possible biases due to the uneven distribution of data over the domain and to provide more robust areal precipitation estimates.
Global warming is expected to modify the water cycle (Huntington, 2006) and precipitation regime in the next future with relevant consequences on the economical activities, especially agriculture and energy production, the biodiversity and landscape,

as well as the development of the human communities. Recent studies have highlighted increased precipitation in northern Europe, North America and northern and central Asia while a drying tendency has been found in the Sahel, the Mediterranean basin, southern Africa, and parts of southern Asia (Hess et al., 1995; Sharma et al., 2000; Hamilton et al., 2001; Boyles and Raman, 2003; Liu et al., 2008; Lebel and Ali, 2009). Moreover, the climate projections depict in many mid-latitudinal and subtropical dry regions a possible decrease in mean precipitation with increasing frequency of long dry-spells, while in many mid-latitudinal wet regions, mean precipitation will be likely to increase. Extreme precipitation events over most of the mid-latitudinal and wet tropical regions are also likely to increase in both magnitude and frequency as a consequence of the increasing moisture in atmosphere (IPCC, 2013). However, the climate signal could be highly variable in relation to the spatial scale of analysis. In particular, the Mediterranean regions experience a high spatio-temporal variability in precipitation regimes due to the position between tropical and mid-latitude climate, their complex topography as well as the influence of the Atlantic Ocean (Lionello, 2006). Many climate studies focused on the reconstruction of long-term evolution of precipitation and its variability over regional and subregional Mediterranean domains. For most Mediterranean regions a general decrease in winter precipitation has been observed over the last decades, especially in Greece (Freidas et al., 2007), Turkey (Toros, 2012) and over the inland of Spain (González-Hidalgo et al., 2009; 2011). In Italy the analyses performed over secular and homogenised precipitation records spanning the last two centuries highlighted a decrease in total precipitation, especially on Central Italy, even though the signal shows a relevant spatial variability (Brunetti et al., 2006). Because of the spatial variability of the climate signal, studies at regional scale are needed, especially over the most vulnerable areas to changes in water cycle and precipitation regime. Reduction of fresh water disposal could have a relevant impact on those regions with scarce water resources and high demand, while the intensification of extreme meteorological conditions could increase the risk of natural hazards, such as landslides and floods, for complex orographically regions.

While many studies have been focused on the climate evolution in northern Italy and Alpine regions, fewer works concern the climate variability in central and southern Italy, where the decrease of data coverage and of long high-quality records, especially at daily resolution, for some specific areas prevents from performing accurate climate studies. Brunetti et al. (2012) set up a dense database of daily precipitation for Calabria region spanning the last century revealing a general decreasing tendency in annual precipitation and a negative trend in high intensities, especially in winter months, even though the trends and their significance vary over the region and show a relevent temporal variability.

In this framework, the presented work aims at investigating the precipitation variabil-

ity of Sardinia region, which is one of the main Mediterranean islands, by means of both spatial and temporal climate reconstruction. A dense database of monthly precipitation series covering the last century has been retrieved and subjected to careful quality-check and homogenisation activities. The set up database represents an improvement and update of the one used by Secci et al. (2010) to study the spatial distribution of the mean annual precipitation over the region for the period 1922-1991. In the presented work, the 1961-1990 monthly precipitation climatologies for Sardinia were computed at 30-arc second resolution, while the gridded dataset of 1922-2011 precipitation records was used to assess the long-term trend and variability of water disposal for the region over a 90-year period. The temporal extent of the reconstructed dataset was chosen according to the period of best data availability.

5.2 Data

The database used for the study is composed by more than 300 series which were partly retrieved from the Italian database already used for the construction of Italian climatologies (Crespi et al., 2018) and partly from the archive of the Hydrographic Service and the digitisation of its hardcopy yearbooks (Servizio Idrografico, 1918; 1925; 1958). More precisely, the monthly precipitation series of 397 Sardinian stations already included in the database for Italian climatologies were merged to the 271 monthly records recovered from the daily database of Hydrographic Service which has been recently released on the website of the Sardinian regional service (source: http://www.regione.sardegna.it/j/v/25?s=131338&v=2&c=5650&t=1) and spans the period 1922-2012. Most of the recovered stations were already included in the Italian archive, however the new records allowed to fill monthly gaps and to extend the previous temporal coverage both in the past and in the most recent years, while in some cases new sites were added. In addition, the temporal coverage of the ten longest series was extended in the past by the digitisation of data from the historical yearbooks of the Hydrographic Service.

The merging activities were managed by a careful analysis of available metadata and led to 408 resulting records. All the series were subjected to quality-control procedures as described in section 1.6 in order to detect and remove outliers and spurious entries as well as to highlight erroneous station locations (both coordinates and elevation). The records showing the lowest agreement with their simulation by means of neighbouring observations were discarded from the database together with the series containing less than 10 years of available data. After these activities, the database included 367 monthly precipitation series covering the whole Sardinia region (Figure 5.1). The temporal evolution of station coverage is shown in Figure 5.2. 9 series have more than 100 years of data, 6 out of which starts before 1900, however most records begin in 1922

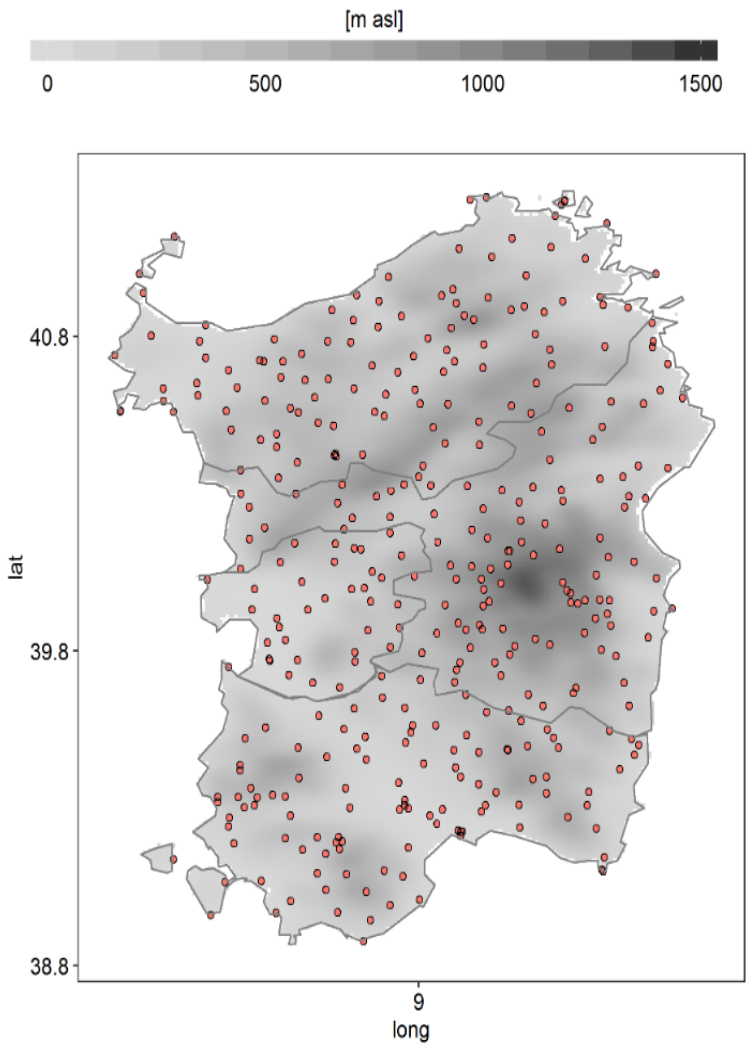

Figure 5.1: Spatial distribution of the 367 available stations over Sardinia. The orography of the region is extracted from a smoothed version of a 30-arc second resolution DEM (GTOPO30).

and the station availability remains almost constant up to present. Temporary reduction of data coverage occurred in the first half of the 20^{th} century during the II World War and in the last decades of the century due to the transition from mechanical to automatic monitoring networks.

In order to compute the high-resolution climatologies for the region, the monthly normals over the period 1961-1990, which corresponds to the interval of best data availability, were computed for each station. In order to avoid biases in climatological values due to the uneven data availability over the reference interval, before computing the monthly average, the missing data in the period 1961-1990 were reconstructed for each station by means of the procedure described in Crespi et al. (2018) and based on the relative anomalies of neighbouring stations. After this procedure, for each month 367 station normals for Sardinia were available corresponding to a spatial density of about one station per 65 km^2.

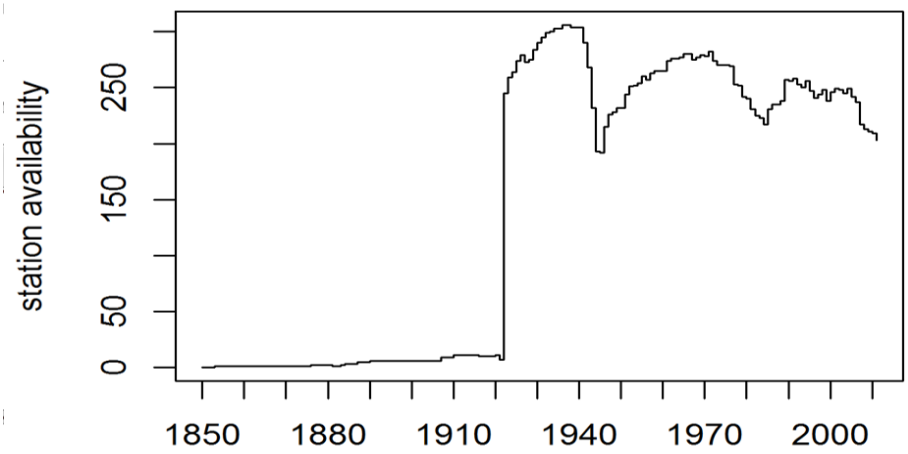

Figure 5.2: Temporal evolution of the station availability over the whole period spanned by the database.

In addition, in order to evaluate the long-term temporal evolution of precipitation over the region, all the series were checked for homogeneity by applying the Craddock test (section 1.6) and 27 monthly records showing relevant breaks were homogenised.

5.3 Methods

The spatio-temporal distribution of precipitation over Sardinia was reconstructed over a 30-arc second resolution Digital Elevation Model (USGS GTOPO30) by applying the anomaly-based approach to the observation database. More specifically, the 1961-1990 high-resolution monthly climatologies for the region were computed by means of the LWLR of precipitation normals *versus* elevation. For each month and at each point of the grid, the 15 surrounding stations with the highest weights were selected to enter in the weighted linear fit. The station weights are expressed as the product of Gaussian functions taking into account the nearness and geographical similarity of the stations to the target grid cell in terms of elevation, sea distance, slope orientation and steepness. The weight decays were defined locally by means of a minimisation procedure of the model errors. From regression coefficients estimated at each grid cell the precipitation normal for the month m was obtained:

$$p_m(x,y) = a_m(x,y) + b_m(x,y) \cdot h(x,y) \tag{5.1}$$

where $h(x,y)$ is the elevation of cell (x,y) and $a_m(x,y)$ and $b_m(x,y)$ are the regression coefficients at target point.
In order to take into account the actual spatial scale at which atmosphere is expected to interact with the surface features, orographic details contained in the original DEM version were smoothed by assigning to each cell the weighted mean elevation of sur-

rounding cells. The cell weights were defined by a Gaussian function of their distance from the target with halving coefficient at 3 km from the cell to estimate. The smoothing degree was set by minimising the model errors.
The gridded dataset of 1922-2011 monthly precipitation records was obtained by converting into anomalies the station records as the ratio with corresponding 1961-1990 normals and by interpolating them onto the 30-arc second resolution grid by a weighted averaging approach. The interpolated anomaly at the target cell $a(x, y)$ for the time step t was obtained as follows:

$$a_t(x, y) = \frac{\sum_{j=1}^{N} w_{j,t}(x, y) \cdot a_{j,t}}{\sum_{j=1}^{N} w_{j,t}(x, y)} \tag{5.2}$$

where N is the total number of stations and w are the station weights at the cell (x, y) which are a Gaussian function of their distance from the point. In order to take into account the variation of station availability over the whole study period, the halving distance of the station weight was computed for each year as the mean distance from the grid cell including at least three stations with available data. The dataset of gridded monthly anomalies was then converted into millimeters by means of the product for the corresponding field of monthly climatologies.
The accuracy of the results was assessed by reconstructing in LOO approach the monthly normals and the 1922-2011 anomaly records of all the stations contained in the database and by comparing modelled and observed values in terms of error estimators (BIAS, MAE and RMSE).

5.4 Results and Discussion

5.4.1 High-resolution monthly climatologies for Sardinia

The robustness of 1961-1990 LWLR monthly climatologies over Sardinia was assessed by the LOO reconstruction of station monthly normals. The bias is almost null in all months suggesting the absence of systematic biases in the model estimates. The MAE and RMSE values range from 10.8 and 14.1 mm in December to 2.0 and 2.7 mm in July, respectively. The higher errors of winter months are mainly due to the greater precipitation normals in respect with the summer season which features the driest conditions. In Figures 5.3 and 5.4 the seasonal and annual 1961-1990 precipitation climatologies for Sardinia are reported. On seasonal scale, the wettest conditions occur during winter with mean precipitation over the whole region of about 270 mm, while summer is characterised by relevant drier conditions with 44 mm as regional average. Spring and autumn feature intermediate precipitation regimes with 177 and 224 mm as regional

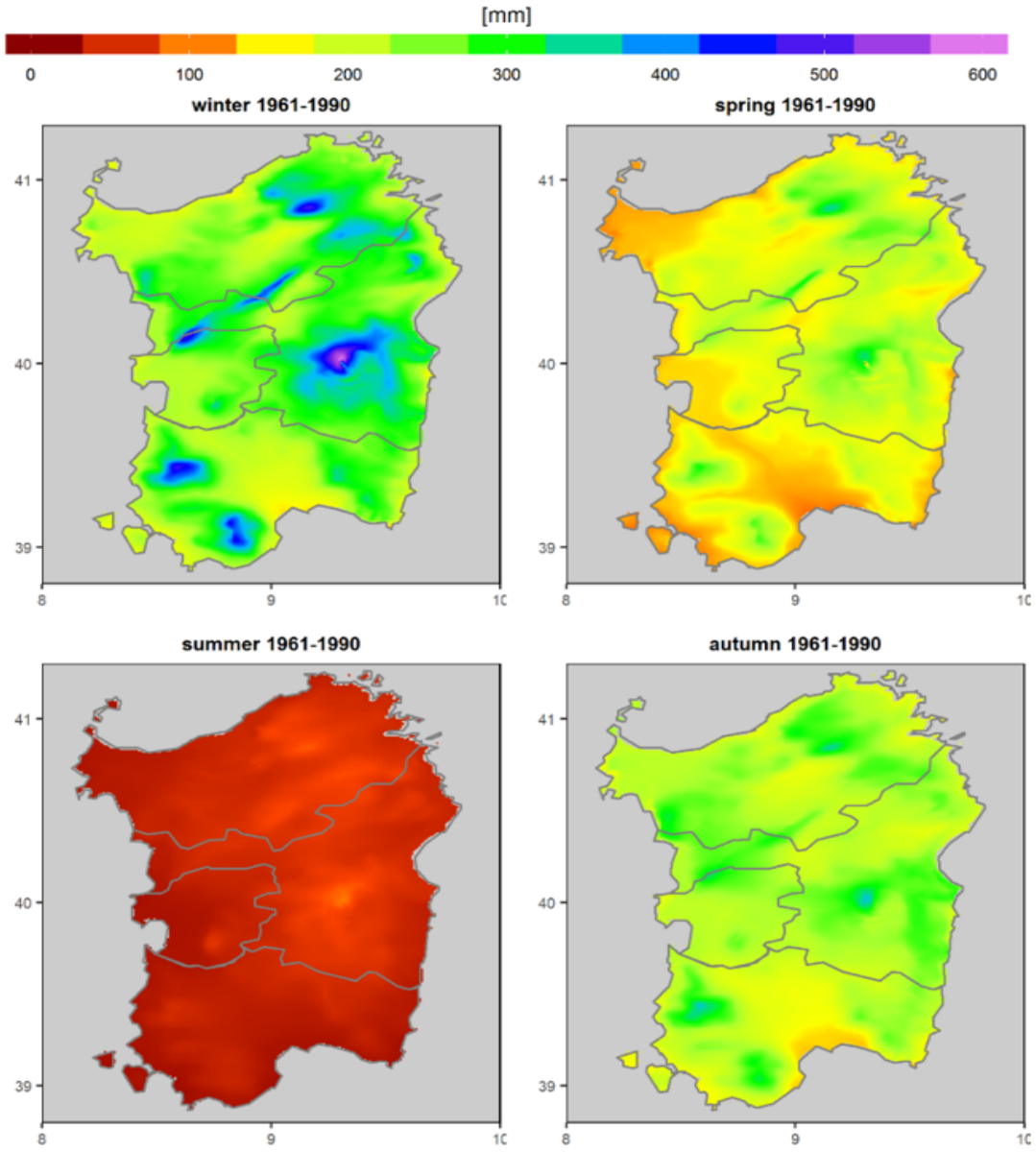

Figure 5.3: 30-arc second resolution 1961-1990 seasonal precipitation climatologies for Sardinia.

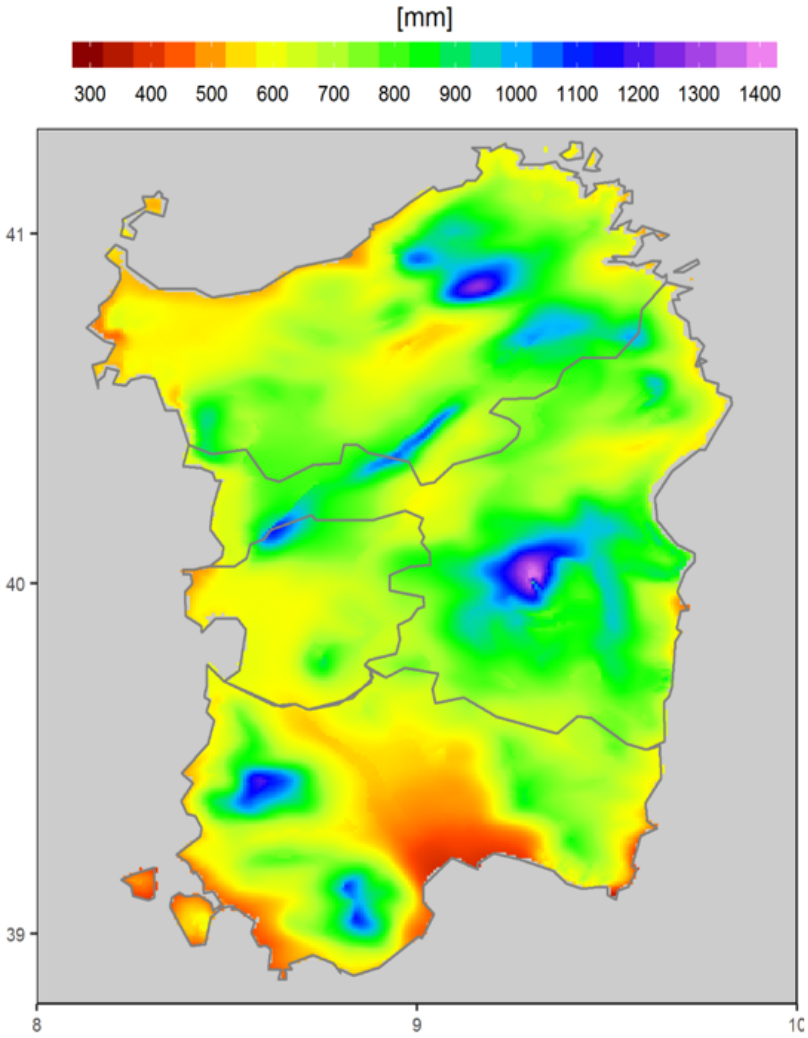

Figure 5.4: 30-arc second resolution 1961-1990 annual precipitation climatologies for Sardinia.

averages, respectively. The highest precipitation values are located in the eastern part of the region, especially around Mount Gennargentu (1834 m a.s.l.) which is the highest elevation of Sardinia with more than 1300 mm as annual maxima. Other wet areas are Gallura in the north-eastern portion of the region and the mountainous areas around Macomer and Iglesias located in western Sardinia. These areas are mainly characterised by orographic precipitation regimes due to the interaction of the moist air masses from the sea with the rough terrains close to the coasts. The driest zones are located in southern Sardinia, especially around Cagliari with mean annual values lower than 500 mm.

5.4.2 1922-2011 monthly precipitation records over Sardinia

The robustness of the gridded dataset of long-term precipitation records over Sardinia was evaluated by computing in LOO approach the available station anomaly records over the 90-year reconstructed period. For each station the MAE value between simulated and observed station anomaly series was evaluated for all months. As average over all the stations, the lowest accuracy is depicted for summer estimates with average MAE of about 45% while the best performances are found for winter reconstructions for which the average error is around 20%. The mean common variance between modelled and measured anomaly records over all months is about 77% with the highest agreement for winter.
The average monthly precipitation record for Sardinia over the period 1922-2011 was finally computed as the average over all the cells of the high-resolution grid. The 90-year record was analysed for long-term trend on both annual and seasonal scales (Figures 5.5 and 5.6). The trend slope was estimated by means of the Theil-Sen test (Theil, 1950; Sen, 1968), while the significance of trends was assessed by the Mann-Kendall estimator (Kendall, 1970).

Significant negative trends with Mann-Kendall p-value below 0.01 were found for the annual and winter precipation records with Theil-Sen slopes of -2.3% and -4.1% decade^{-1}, respectively. As regards the other seasons, a decreasing tendency was pointed out for spring and autumn records while a positive trend emerged for summer precipitation, even though in all cases the trend is not statistically significant with significance levels far above 0.05. These outcomes are in agreement with those reported in Brunetti et al. (2012) for the 1916-2006 total annual and seasonal precipitation records for Calabria (southern Italy).

5.4.3 Variability of annual precipitation climatologies

The anomaly-based approach described in section 5.3 was also applied to reconstruct the gridded precipitation climatologies for Sardinia over consecutive 30-year periods

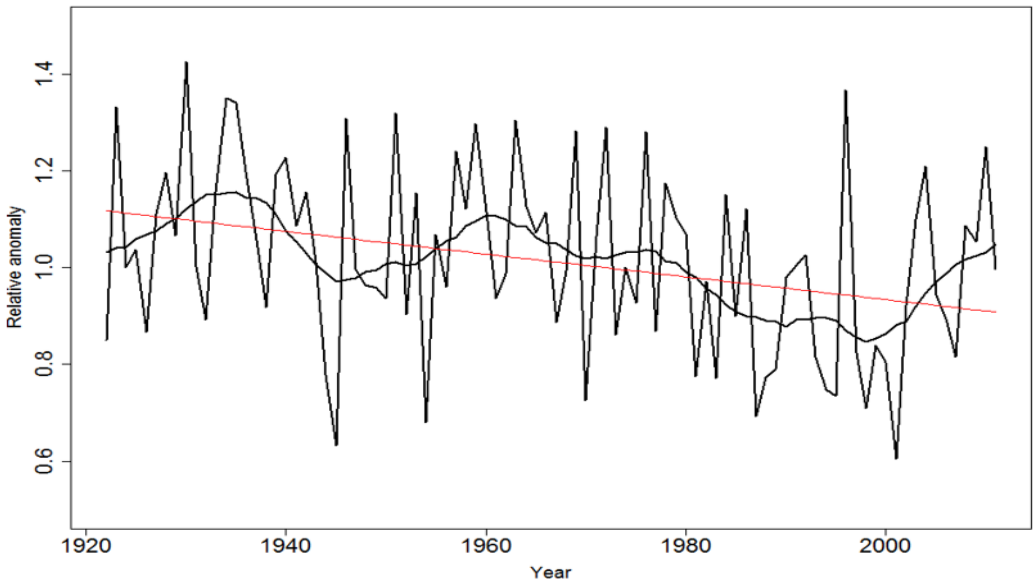

Figure 5.5: 1922-2011 annual average precipitation record over Sardinia together with the 11-year Gaussian filter with 3-year standard deviation. The red solid line represents the Theil-Sen linear fit.

from 1931 to 2010. The estimated fields allowed to evaluate the evolution and variability of precipitation distribution and magnitude over the region during the last century. More precisely, for each interval under consideration (1931-1960, 1941-1970, 1951-1980, 1971-2000, 1981-2010) the missing monthly data over the 30-year period in each station record were reconstructed by interpolating the anomalies from 1961-1990 normals of available surrounding stations (equation 5.2) and rescaling the modelled value for the 1961-1990 monthly normal of the station under reconstruction. After completing the station series, their mean monthly precipitation values over the analysed 30-year period were computed and interpolated onto the 30-arc second grid by means of LWLR (equation 5.1).

The evolution of annual precipitation climatologies from 1931 to 2010 (Figure 5.7) confirms the significant negative trend in regional annual precipitation depicted by Theil-Sen analysis. The annual precipitation normal, as average over all the grid points, decreases from 761 mm in 1931-1960 climatology to 658 mm in the last considered interval. In particular, 1961-1990 climatology points out a remarkable decrease in mean annual precipitation in respect with the previous 30-year mean conditions, with an average decrease of about 40 mm corresponding to a loss of about 5% in respect with the 1951-1980 mean annual precipitation over the region. The most relevant drying tendency regarded both mountainous areas, especially Mount Gennargentu and Gallura, and the coastal zones.

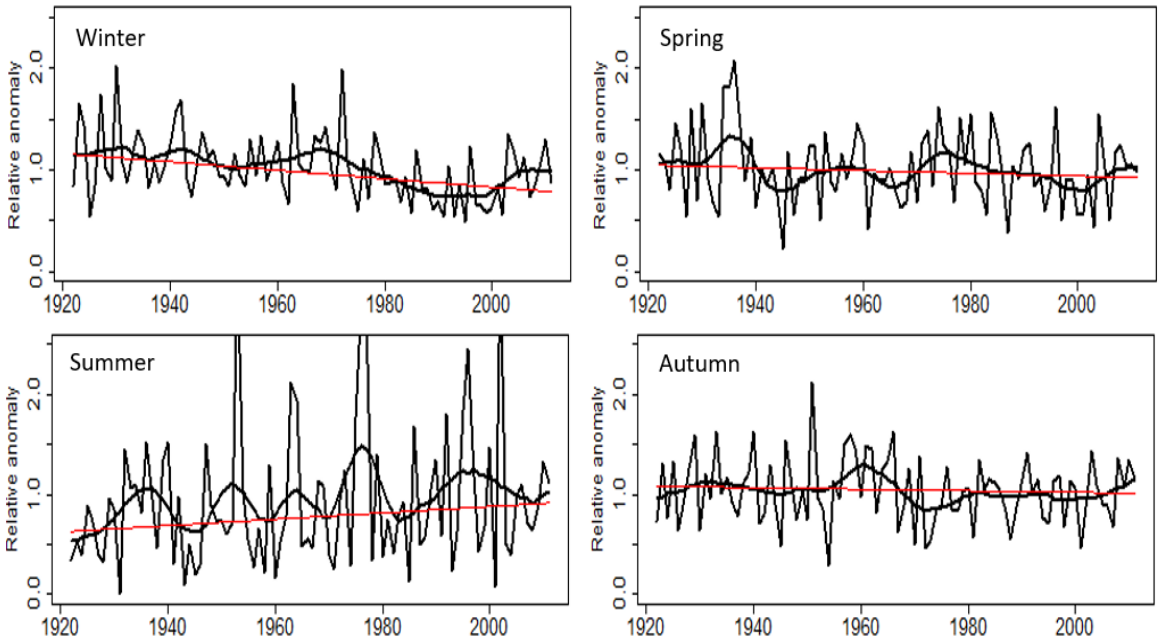

Figure 5.6: 1922-2011 seasonal average precipitation records over Sardinia together with the 11-year Gaussian filters with 3-year standard deviation. The red solid lines represent the Theil-Sen linear fits.

5.5 Conclusions

The long-term spatio-temporal evolution of monthly precipitation over Sardinia, one of the main Italian Mediterranean islands, was analysed by computing the 1961-1990 monthly climatologies and the gridded dataset of 1922-2011 monthly precipitation records at 30-arc second resolution. At this aim, a dense database of monthly precipitation records for the region spanning the last century was set up by recovering both the past non-digitised data and the recent available automatic records. All the retrieved series were controlled for data quality and homogeneity and the homogenisation procedure was applied on the records showing the most relevant breaks.

The temporal reconstruction was performed by means of the anomaly-based method and the accuracy of the results was evaluated by simulating in LOO approach the station data. The highest errors in station normal reconstruction was pointed out for winter, which corresponds to the wettest season in Sardinia, with the maximum MAE in December (14.1 mm). As regards the temporal reconstruction the mean common variance between the simulated and observed station anomaly series is about 77% while the greatest errors in station anomaly simulation are obtained for summer values (45%), due to the very low summer precipitation which are more likely to be affected by relevant fluctuations in modelled values in relative terms.

The long-term trends and variability of the 1922-2011 regional precipitation record in both annual and seasonal scales were evaluated by means of Theil-Sen test and

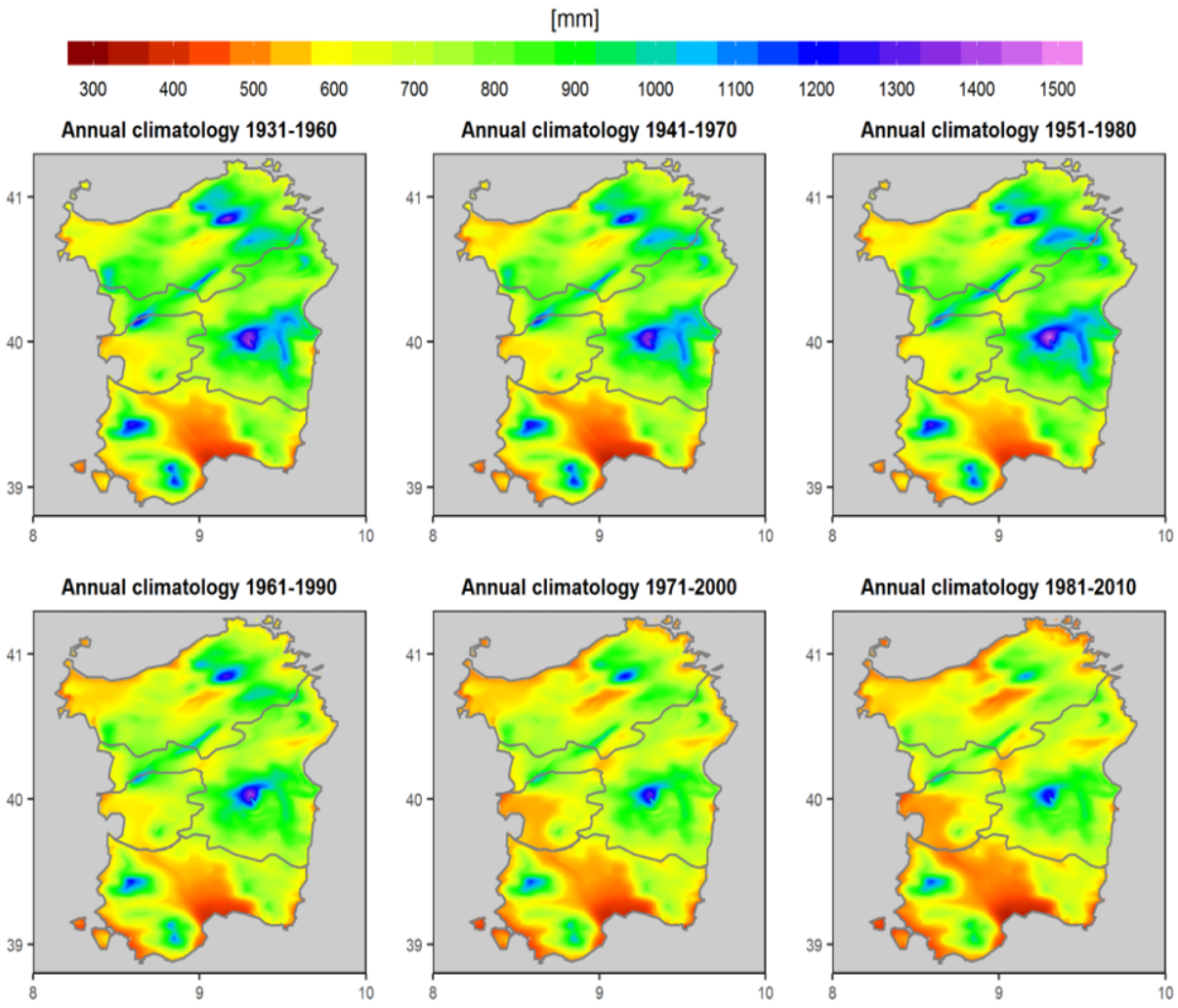

Figure 5.7: Evolution of annual precipitation climatologies for Sardinia over consecutive 30-year periods.

Mann-Kendall estimator. A statistically significant negative trend (p-value<0.01) was depicted for annual and winter precipitation while an increasing but not significant tendency was depicted for the summer series. The same drying tendency in annual precipitation was observed by reconstructing and comparing the annual precipitation climatologies for Sardinia over consecutive 30-year periods from 1931 to 2010.
The presented outcomes confirm the long-term trends in precipitation evolution which was found for other Mediterranean areas and southern Italian regions. However, in order to analyse the long-term trends on precipitation regime on a finer time scale and to assess possible acting changes in the frequency of extremes, in the next future we will focus our analyses on the daily resolution too. In addition, running-trend analyses are required in order to study the short-term variability in precipitation evolution over Sardinia and to evaluate the trends and their significance over different subperiods. These activities are in course and the results will be discussed in detail in forthcoming

papers.

5.6 References

- Boyles, R. P., and Raman, S. (2003). Analysis of climate trends in North Carolina (1949–1998). Environment International, 29, 263–275. doi:10.1016/S0160-4120(02)00185-X

- Brunetti, M., Buffoni, L., Maugeri, M., and Nanni, T. (2006). Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. International Journal of Climatology, 26, 345–381. doi:10.1002/joc.1251

- Brunetti, M., Caloiero, T., Coscarelli, R., Gullá, G., Nanni, T., and Simolo, C. (2012). Precipitation variability and change in the Calabria region (Italy) from a high resolution daily dataset. International Journal of Climatology, 32, 57-73. doi:10.1002/joc.2233

- Crespi, A., Brunetti, M., Lentini, G., and Maugeri, M. (2018). 1961-1990 high-resolution monthly precipitation climatologies for Italy. International Journal of Climatology, 38, 878–895. doi:10.1002/joc.5217

- Feidas, H., Noulopoulou, C., Makrogiannis, T., and Bora-Senta, E. (2007). Trend analysis of precipitation time series in Greece and their relationship with circulation using surface and satellite data: 1955–2001. Theoretical and Applied Climatology, 87, 155–177. doi:10.1007/s00704-006-0200-5

- González-Hidalgo, J. C., Lopez-Bustins, J., Štepánek, P., Martin-Vide, J., and de Luis, M. (2009). Monthly precipitation trends on the Mediterranean fringe of the Iberian Peninsula during the second-half of the twentieth century (1951–2000). International Journal of Climatology, 29, 1415–1429. doi:10.1002/joc.1780

- González-Hidalgo, J. C., Brunetti, M., and de Luis, M. (2011). A new tool for monthly precipitation analysis in Spain: MOPREDAS database (Monthly precipitation trends December 1945–November 2005). International Journal of Climatology, 31, 715–731. doi:10.1002/joc.2115

- Hamilton, J. P., Whitelaw, G. S., and Fenech, A. (2001). Mean annual temperature and total annual precipitation trends at Canadian biosphere reserves. Environmental Monitoring and Assessment, 67, 239–275. doi:10.1023/A:1006490707949

- Hess, T. M., Stephens, W., and Maryah, U. M. (1995). Rainfall trends in the North East Arid Zone of Nigeria 1961–1990. Agricultural and Forest Meteorology, 74, 87–97. doi:10.1016/0168-1923(94)02179-N

- Huntington, T. G. (2006). Evidence for intensification of the global water cycle: review and synthesis. Journal of Hydrology, 319, 83–95. doi:10.1016/j.jhydrol.2005.07.003

- Kendall, M. G. (1970). Rank Correlation Methods, Griffin, London, UK.

- Lebel, T., and Ali, A. (2009). Recent trends in the Central and Western Sahel rainfall regime (1990–2007). Journal of Hydrology, 375, 52–64. doi:10.1016/j.jhydrol.2008.11.030

- Lionello, P., Malanotte-Rizzoli, P., and Boscolo, R. (2006). The Mediterranean Climate: An Overview of the Main Characteristics and Issues. Elsevier: Amsterdam.

- Liu, Q., Yang, Z., and Cui, B. (2008). Spatial and temporal variability of annual precipitation during 1961–2006 in Yellow River Basin, China. Journal of Hydrology, 361, 330–338. doi:10.1016/j.jhydrol.2008.08.002

- Secci, D., Patriche, C. V., Ursu, A.,and Sfîcă, L. (2010). Spatial interpolation of mean annual precipitations in Sardinia. A comparative analysis of several methods. Geographia Technica, 1, 67–75.

- Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. Journal of the American Statistical Association, 63, 1379–1389.

- Servizio Idrografico (1918). Osservazioni pluviometriche raccolte a tutto l'anno 1915. Volume I, Sardegna. Fascicolo I. Ministero dei Lavori Pubblici – Consiglio Superiore delle Acque, Roma (in Italian).

- Servizio Idrografico (1925). Osservazioni pluviometriche raccolte nel quinquennio 1916–1920. Pubblicazione numero 1 del Servizio Idrografico. Volume IV: Italia peninsulare e isole. Ministero dei Lavori Pubblici – Consiglio Superiore, Roma (in Italian).

- Servizio Idrografico (1958). Precipitazioni medie mensili ed annue e numero di giorni piovosi per il trentennio 1921–1950, Sezione idrografica di Cagliari, Ministero dei Lavori Pubblici – Consiglio Superiore, Roma (in Italian).

- Sharma, K. P., Moore, B. III, and Vorosmarty, C. J. (2000). Anthropogenic, climatic, and hydrologic trends in the Kosi basin, Himalaya. Climatic Change, 47, 141–165. doi:10.1023/A:1005696808953

- Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. Proceedings of Koninklijke Nederlandse Akademie von Wetenschappen A, 53, 1397–1412.
- Toros, H. (2012). Spatio-temporal precipitation change assessments over Turkey. International Journal of Climatology, 32, 1310-1325. doi:10.1002/joc.2353
- USGS - United States Geological Survey, (1996). website: http://eros.usgs.gov/

Chapter 6

High-resolution monthly precipitation climatologies over Norway (1981–2010): joining numerical model datasets and in-situ observations

The chapter presents the results of the research activity managed during the 6-month visiting period spent by the Norwegian Meteorological Institute of Oslo. More specifically, the 1981–2010 monthly precipitation climatologies over Norway at 1 km resolution are presented. They are computed by an interpolation procedure combining the output from a numerical model with the in-situ observations in order to overcome the reconstruction limits of conventional interpolation methods due to the uneven distribution of the meteorological network over the country, especially in the northermost Norwegian regions and in mountainous areas. At this aim, the regional climate model dataset HCLIM-AROME, based on the dynamical downscaling of the global ERA-Interim reanalysis onto 2.5 km resolution, is considered together with quality-checked records of the 2009 rain-gauges located within the model domain. The precipitation climatologies are defined by superimposing the grid of 1981–2010 monthly normals from the numerical model and the kriging interpolation of station residuals. The integration of rain-gauge data proved to remove the original numerical model biases and the errors obtained from the LOO station validation turned out to be lower than those provided by two conventional interpolation schemes based on observations only. In addition the combined approach was found to be less sensitive to the spatial variability of station distribution over the domain and to prevent from the reconstruction of discontinuities and outliers, especially over those areas not covered by the rain-gauge observations. The obtained climatologies clearly depict the main West-to-East gradient occurring from the orographic precipitation regime of the coast to the more continental

climate of the inland and it allows to point out the features of the climatic subzones of Norway.
The chapter is based on the manuscript submitted to the *International Journal of Climatology* and the released version of the gridded dataset is freely available at https://zenodo.org/record/1304042#.Wzti8Bx9i90.
The research activities were presented at an international conference and described in a technical report freely available on the Norwegian Meteorological Institute website:

- Crespi, A., Lussana, C., Bunetti, M., Dobler, A., Maugeri, M., and Tveito, O. E. (2018). A new combined interpolation approach for 1981–2010 monthly precipitation climatologies over Norway: joining numerical model output with in-situ observations. EMS Annual Meeting: European Conference for Applied Meteorology and Climatology, Budapest (Hungary), Vol. 15, EMS2018-236.

- Crespi, A., Lussana, C., Bunetti, M., Dobler, A., Maugeri, M., and Tveito, O. E. (2018). High-resolution monthly precipitation climatologies over Norway: assessment of spatial interpolation methods. MET Report, The Norwegian Meteorological Institute, Oslo, Norway, No. 03/2018, ISSN 2387-4201.

6.1 Introduction

Gridded climatological datasets of surface precipitation represent valuable information sources for both researchers and decision makers in a wide range of fields, such as energy production, management and conservation of natural resources and agriculture (Prein and Gobiet, 2017). The availability of accurate high-resolution descriptions of spatial distribution of the precipitation normals over the territory is particularly relevant for those countries, such as Norway, whose electricity sector relies primarily on hydropower (Dyrrdal et al., 2015).
The reconstruction of conventional climatological datasets, as defined by Simmons et al. (2017), requires both dense rain-gauge networks and statistical interpolation approaches able to capture the interactions between atmospheric circulation and the surface (Henn et al., 2018). In fact, precipitation distribution is found to be strongly influenced by the main geographical features, such as elevation, sea nearness and slope conditions, and these relationships could highly vary at local level (Daly et al., 2008). A wide selection of statistical schemes has been developed and applied so far to project monthly precipitation normals from station sites onto the unsampled points of high-resolution grids, such as geostatistical approaches based on kriging and all its variants (Goovaerts, 2000; Hengl, 2009), regression-based models (Crespi et al., 2018; Daly, 2002), spline and inverse distance weighting interpolation (Boer et al., 2001). The uneven spatial coverage of observational data could significantly affect the interpolation

accuracy especially in highly variable terrains where the model predictions could be derived from rain-gauges located at distant and very different environments.
The Norwegian Meteorological Institute (MET) provided two versions of the seNorge gridded dataset for daily precipitation over Norway at 1 km resolution: one is based on a triangulation procedure with altitude corrections (Tveito et al., 2005; Mohr, 2008; 2009) and the other is based on a multi-scale optimal interpolation approach (Lussana et al., 2018). The distribution of Norwegian station network is very sparse over the Northern inland and the scarcely sampled high-elevated areas significantly limit the ability of observation-based models to capture the actual conditions occurring over these regions. In particular, remarkable underestimations in gridded precipitation values are found to occur over the mountainous inland where the complex environment prevents the management of manual stations and limits the availability of dense observations (Lussana et al., 2018). Only in the recent years, some new automatic rain-gauges have been established in these areas and their observations have improved the monitoring applications even though the length of the new records is not yet suitable for climatological studies.
In this study, we aim to combine two data sources, such as in-situ observations and numerical model outputs in order to obtain monthly precipitation climatologies of reasonable quality even in data-sparse areas over complex terrains. The advancements in the accuracy and spatial resolution of numerical models in fact provided in the last years new data and information, which started to be used in the inter-comparisons with observation-based gridded datasets (Haylock et al., 2008; Dyrrdal et al., 2018) as well as in climatological studies (Karger et al., 2017; Berthou et al., 2018). In particular, regional reanalysis models are primarily relevant for climatological applications because of the high spatial resolution (almost comparable to operational numerical weather prediction systems, NWP) of their long-term gridded datasets (Isotta et al., 2015; Jermey and Renshaw, 2016). They do not assimilate in-situ precipitation observations so the two data sources can be considered independent from each other. The high-resolution of the regional reanalysis datasets is found to improve the description of precipitation fields in comparison with the global ones, especially for extreme events, even though reanalyses are affected by significant biases, shifts in regional anomalies and inaccuracies in mountain-valley contrasts.
MET has recently produced a climate model dataset of precipitation covering the Norwegian mainland at 2.5 km resolution over the period July 2003 to December 2016. The dataset is based on the dynamical downscaling of the global ERA-Interim reanalysis (Dee et al., 2011) using HCLIM-AROME, the climate model version of the HARMONIE (HIRLAM ALADIN Research towards Mesoscale Operational NWP In Europe) NWP model framework (Lind et al., 2016). The model includes a set of different physics packages adapted for different horizontal resolutions. The high spatial resolu-

tion of the HCLIM-AROME model is suitable to provide fine-scale realistic precipitation patterns over the whole Norwegian domain, where the circulation is dominated by Westerlies, and it supplies useful information over the remote regions not covered by in-situ observations. The overall model setup is similar to the operational AROME-MetCoOp model (Müller et al., 2017), which provides realistic precipitation patterns especially where orographically forced precipitation is the dominant mechanism. In this framework, we developed and applied an interpolation approach in which rain-gauge observations and the numerical model output are used together to provide the 1981–2010 monthly precipitation climatologies over Norway at 1 km grid spacing. In this scheme, hereafter referred to as HCLIM+RK, the grids of monthly precipitation normals derived from HCLIM-AROME dataset are considered as background and a kriging interpolation procedure is applied to adjust the original fields by means of station residuals, i.e. the differences between the station normals and the numerical model estimates at the closest grid cells. This procedure allows both to exploit the spatial precipitation gradients resolved by HCLIM-AROME, especially over the regions not covered by stations, and to improve the accuracy of resulting climatologies by integrating the available in-situ observations to correct the numerical model biases. In order to compute the 30-year climatologies, the HCLIM-AROME dataset was extended back to 1981 using available station records and a reconstruction procedure based on multiplicative anomalies.
In the present work the features of HCLIM+RK method are extensively investigated and discussed and the computed 1981–2010 monthly precipitation climatologies over Norway are presented. Moreover, in order to assess the improvements provided to Norwegian climatologies by the combination of numerical model datasets and rain-gauge data, HCLIM+RK performances are compared to those obtained by two statistical methods based on observations only: a Multi-linear Local Regression Kriging (MLRK) and a Local Weighted Linear Regression (LWLR) of precipitation versus elevation, based on PRISM framework (Daly, 2002).

6.2 Data

6.2.1 Precipitation data from the rain-gauge network

The observational database considered to reconstruct the 1981–2010 monthly precipitation climatologies over Norway is composed by more than 2000 station records of monthly precipitation covering the HCLIM-AROME domain (Figure 6.1). The Norwegian station series were retrieved from the MET Norway Climate daily database (KDVH) and they were integrated with the daily records contained in the European Climate Assessment and Dataset (ECA&D) for the surrounding countries.

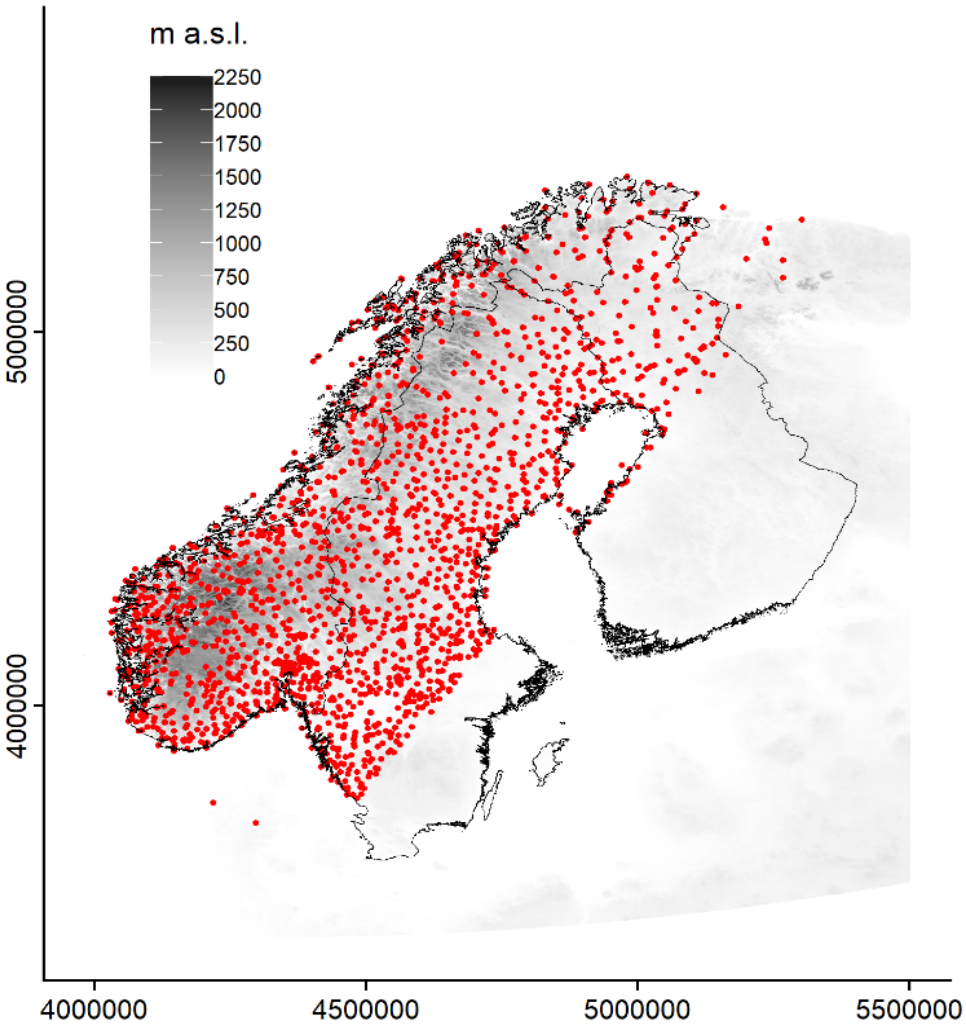

Figure 6.1: Distribution of the 2009 stations considered for the Norwegian climatology reconstruction.

The monthly database was checked for quality in order to detect and correct spurious entries, duplicates and erroneous locations. In particular, following the procedure described in Crespi et al. (2018), each measured series was compared to the simulated one by means of neighbouring station data and reconstruction errors were computed in terms of Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). High errors allowed to point out and remove single problematic periods in a record or to discard stations whose observed values completely mismatched the simulated ones. After removing gross data errors in both monthly and daily series, the gap-filling procedure described in Golzio et al. (2018) was applied to daily records in order to maximise the number and length of monthly data series available for climatological purposes. Monthly precipitation series were computed again for each station from filled daily records and whenever daily data were still missing, the corresponding monthly total was not computed. On average, the filling procedure allowed to re-

construct 3% of daily gaps which led to increase the available monthly precipitation totals for each station of about 12%. Stations with less than 10 years of available data, also after the gap-filling, were definitely discarded from the database. After these activities, the 1981–2010 monthly precipitation normals were computed for each series and, whenever this period was partially or completely unavailable, missing months were reconstructed by a procedure based on multiplicative anomalies of neighbouring stations (Crespi et al., 2018). Since a remarkable fraction of stations (52%) had more than 30% of missing data in the reference period, this procedure allowed to prevent monthly normals being biased by a lower fraction of monthly data entering in the average computation.
Finally, the 1981–2010 precipitation normals were available for 2009 sites, 1043 out of them located in Norway. However, Norwegian sites are unevenly distributed over the country: data coverage is generally higher in the South (below 63°18'N) with about one station per 250 km^2 and it decreases significantly towards the North (above 63°18'N) with about one station per 500 km^2 (Figure 6.1). Moreover, most rain-gauges are located at low-elevation and only 15% and 1% of stations are above 500 and 1000 m a.s.l., respectively, so that the grid cells at higher elevation (47% and 15% of the total are above 500 and 1000 m a.s.l., respectively) are mostly or completely uncovered by in-situ observations (Figure 6.2).

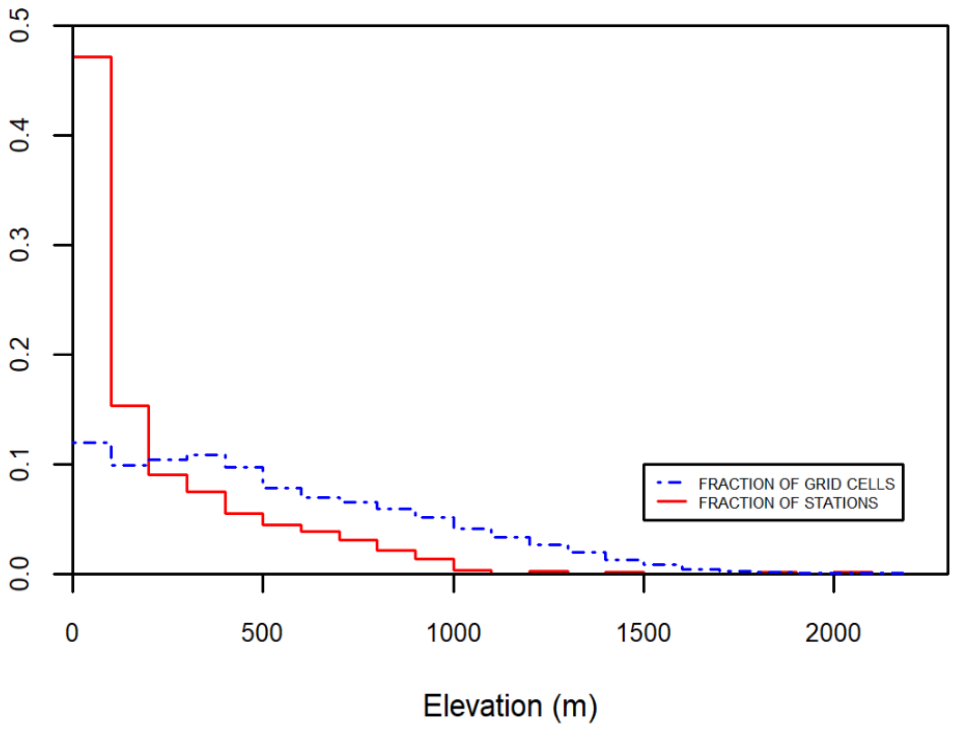

Figure 6.2: Vertical distribution of the 1043 Norwegian stations (red solid line) compared to the grid-cell elevation distribution (blue dashed line) over Norway.

6.2.2 HCLIM-AROME numerical model dataset of monthly precipitation

The HCLIM-AROME dataset of monthly precipitation over Norway was retrieved from the hourly precipitation fields at 2.5 km grid spacing computed by the climate model version of HARMONIE (version cy38h1.2). To perform the high-resolution convection permitting simulations, the model was set-up with the AROME physics (Seity et al., 2011) and the SURFEX surface scheme (Masson et al., 2013). The model was run from July 2003 to December 2016 and it covers the whole Norwegian mainland and part of Sweden, Finland and Russia.
The series of monthly totals were computed for each grid cell by summing the hourly precipitation and the grid was converted into ETRS89 reference system with the Lambert Azimuthal Equal Area (LAEA) projection.
In order to obtain the 1981–2010 mean monthly HCLIM-AROME fields, the 2003–2016 monthly precipitation series of all the cells in the 2.5 km grid were extended back to January 1981 by means of available in-situ observations. For this purpose, the same reconstruction method applied to fill monthly station gaps and based on multiplicative anomalies was considered. More precisely, each missing monthly value of all the HCLIM-AROME grid points ($p_m^{(HCLIM-AROME)}$) was reconstructed by selecting the 5 closest stations with at least 9 years of common data in the period 2003–2016 and a valid entry for the month under reconstruction. From the selected station data, 5 simulated values for month m at the target cell were computed by rescaling each rain-gauge entry $p_{m,i}$ by the ratio between the monthly means of HCLIM-AROME cell values ($\overline{p_{m,i}^{HCLIM-AROME}}$) and station observations ($\overline{p_{m,i}}$) over the available years in the period 2003–2016:

$$p_{m,i}^{HCLIM-AROME} = p_{m,i} \cdot \frac{\overline{p_{m,i}^{HCLIM-AROME}}}{\overline{p_{m,i}}} \qquad i = (1, \ldots, 5) \tag{6.1}$$

The final estimate is defined as the weighted mean of the 5 simulations where the station weight is expressed as a Gaussian function of the radial distance from the considered cell (Crespi et al., 2018).
After the reconstruction, the 1981–2010 HCLIM-AROME monthly normals were computed and downscaled to the target 1 km resolution grid by means of a bilinear interpolation procedure.

6.3 Methods

6.3.1 The combined interpolation scheme: HCLIM+RK

In HCLIM+RK scheme the residuals between each station normal and the corresponding HCLIM-AROME value at the closest grid cell are computed for each month and their spatial distribution is modelled by the kriging procedure. The sample variogram is reconstructed from all the station pairs within 300 km and by setting the bin width to 15 km, while the fitted variogram is obtained by considering the exponential model. The final value of the 1981–2010 precipitation normal for the month m at cell (x, y) is obtained by adding the interpolated station residuals to the corresponding monthly normal from the downscaled HCLIM-AROME background:

$$p_m(x,y) = p_m^{HCLIM-AROME}(x,y) + \boldsymbol{k}^{\boldsymbol{T}}(x,y) \cdot \boldsymbol{\epsilon} \tag{6.2}$$

where $\boldsymbol{k}(x,y)$ is the vector of kriging weights at the cell (x,y) and $\boldsymbol{\epsilon}$ are the station residuals.

6.3.2 Observation-based interpolation methods

In MLRK the precipitation normal for month m at each grid cell (x,y) is computed by applying the residual regression kriging with the regression based on a local multilinear relationship between precipitation and several geographical predictors:

$$p_m(x,y) = \sum_{j=1}^{N} \alpha_{j,m}(x,y) \cdot q_{j,m}(x,y) + \boldsymbol{k}^{\boldsymbol{T}}(x,y) \cdot \boldsymbol{\epsilon} \tag{6.3}$$

where N is the numbers of geographical predictors, $\boldsymbol{\alpha}$ and $\boldsymbol{q}$ are the vectors of regression coefficients and predictor values at target site (x,y), respectively, $\boldsymbol{k}(x,y)$ is the vector of kriging weights and $\boldsymbol{\epsilon}$ are the residuals between observed normals and regression estimates at station locations. Latitude, longitude, elevation and sea distance are chosen as regression geographical predictors for all months while $\boldsymbol{\alpha}$ are monthly estimated for each grid cell by least-square method considering precipitation normals and geographical features of all sample sites within 100 km from the point. If less than 100 stations are available for regression within this distance, searching radius is incremented by 10 km until the minimum threshold is reached. Both the searching radius and the number of stations entering in the regression are defined by the minimisation of model errors. Also in this case, the experimental variogram is defined by considering a bin width of 15 km and all the station pairs within 300 km and it is fitted by the exponential model.

In LWLR the monthly precipitation normals at each grid cell (x,y) are modelled by

using elevation as the main geographical predictor:

$$p_m(x,y) = a_m(x,y) + b_m(x,y) \cdot h(x,y) \tag{6.4}$$

where $h(x,y)$ is the elevation of cell (x,y) and $a_m(x,y)$ and $b_m(x,y)$ are the regression coefficients at target point. The precipitation-elevation relationship is estimated for each month and at each grid point from a weighted linear regression involving neighbouring stations. In order to take into account the influence of local surface features on precipitation distribution, the stations enter in the regression with weights depending on their nearness and orographic similarities (elevation, slope steepness, slope orientation and sea distance) to the target cell. The weighting function decay rate is locally optimised and evaluated for each month. In particular, the radial weighting function presents an optimal halving distance ranging from about 11 km in the South during winter months to 70 km over the northernmost regions in summer.
In both MLRK and LWLR the orographic information is extracted from a smoothed version of a 1 km resolution digital elevation model (DEM). The smoothing allows to remove too fine terrain details and to consider a spatial scale more similar to that at which the interaction between atmospheric circulation and surface is expected to occur (Foresti and Pozdnoukhov, 2012). The smoothed DEM is obtained by substituting the elevation of each cell by the weighted average of the elevations of surrounding cells, with weighting functions (of gaussian shape) decreasing to 0.5 at a distance d (halving distance) from the considered cell, such that the 1 km resolution is preserved. Different halving distances were considered and $d = 3$ km turned out to be the smoothing degree that minimises the model errors.

6.3.3 Validation procedures of interpolation schemes

The accuracy of each considered method was evaluated by reconstructing the 1981–2010 monthly precipitation normals of all the 1043 stations located in Norway. The reconstruction was performed by means of the leave-one-out (LOO) approach, i.e. by excluding the station data under estimation, in order to avoid self-influence. Actually, in order to contain the computational time demand, in the LOO procedure for all the kriging-based approaches, the covariance matrix was computed from the full dataset, while the kriging weight of the station to be reconstructed was set to zero and the remaining station weights renormalised. The modelled values for all the N stations considered in the validation subset were compared month-by-month to the observations by means of BIAS, MAE, MAPE and RMSE (see section 1.5).

6.4 Results

6.4.1 HCLIM+RK validation and comparison with observation-based methods

The LOO errors obtained for HCLIM+RK, MLRK and LWLR 1981–2010 climatologies are shown in Table 6.1 while the errors of the corresponding climatologies from the original HCLIM-AROME dataset are listed in Table 6.2. Thanks to the integration of in-situ observations in HCLIM+RK scheme, the significant BIAS of the original HCLIM-AROME fields is almost completely corrected in all months, especially in winter, and MAE and RMSE are reduced accordingly. The bias of the original HCLIM-AROME fields is highlighted by the distribution of the relative differences in the 1981–2010 normals between the station records and model estimates: they are shown in Figure 6.3 for January and July. The highest differences occur along the south-western coast and over the Lofoten islands, especially in January, where HCLIM-AROME climatologies underestimate the precipitation normals. In addition to the coastal bias, an overall tendency to overestimate precipitation normals is depicted in July over the inland in central-southern Norway. These findings are consistent with the results obtained by Müller et al. (2017) for AROME-MetCoOp: detailed studies focusing on identifying the possible sources of AROME biases will be undertaken in the next future.

On the contrary, the BIAS is almost zero for HCLIM+RK, MLRK and LWLR, suggesting that no method is affected by significant systematic errors. However, the combined HCLIM+RK approach provides the best performance in all months. The mean MAPE

	HCLIM+RK				MLRK				LWLR			
	BIAS	MAE	MAPE	RMSE	BIAS	MAE	MAPE	RMSE	BIAS	MAE	MAPE	RMSE
1	0.3	13.9	12.5	21.2	0.6	15.9	14.4	25.1	-0.1	16.5	14.3	26.4
2	0.2	11.3	13.4	17.5	0.5	12.9	15.3	20.3	-0.3	13.4	15.2	21.2
3	0.2	10.4	13.5	15.9	0.4	12.1	15.5	18.9	-0.2	12.5	15.4	19.8
4	0.1	7.3	12.9	10.8	0.2	8.4	14.7	12.9	-0.3	8.6	14.4	13.1
5	0.1	6.3	10.4	9.2	0.0	7.0	11.5	10.8	-0.4	7.0	11.3	10.4
6	0.0	6.9	9.9	9.6	0.0	7.1	10.1	10.7	-0.3	7.2	10.0	10.5
7	0.1	7.9	9.1	11.2	0.0	7.8	8.9	11.7	-0.1	7.8	8.8	11.8
8	0.1	9.0	8.9	12.9	0.0	9.5	9.2	14.6	-0.3	9.7	9.2	14.7
9	0.2	10.3	8.7	16.2	0.1	12.4	10.7	20.4	-0.3	12.5	10.3	20.0
10	0.2	12.2	9.4	18.8	0.3	14.3	11.3	22.9	-0.3	14.6	11.1	22.8
11	0.3	11.8	11.0	18.4	0.5	13.9	12.9	21.7	-0.1	14.1	12.6	22.0
12	0.2	13.1	12.2	20.3	0.5	15.1	14.0	24.2	-0.3	15.5	13.9	24.8

Table 6.1: Monthly leave-one-out reconstruction errors of HCLIM+RK, MLRK and LWLR for the 1043 Norwegian stations. Except for MAPE, all the values are expressed in mm and BIAS is defined as the difference between simulation and observation.

	HCLIM-AROME		
	BIAS	MAE	RMSE
1	-18.3	26.0	36.9
2	-14.0	20.3	29.0
3	-13.5	20.6	29.4
4	-5.1	14.8	19.7
5	-0.1	12.0	15.9
6	-3.4	16.6	21.2
7	-0.9	20.1	26.2
8	-7.1	22.9	29.6
9	-21.3	27.4	38.8
10	-26.4	31.1	42.0
11	-20.9	28.0	37.9
12	-20.6	26.5	36.8

Table 6.2: Errors of the 1981–2010 monthly normals from HCLIM-AROME original fields at the 1043 Norwegian station sites. All the values are expressed in mm and BIAS is defined as the difference between HCLIM-AROME estimate and observation.

over all months turns out to be about 11% for HCLIM+RK and about 12% for both MLRK and LWLR. The average correlation with observations is 0.95 for HCLIM+RK and 0.94 for the other two approaches with the best agreement in winter for which HCLIM+RK reconstructions explain up to 94% of the variance of the observed station normals.

HCLIM+RK, MLRK and LWLR reconstruction errors were also evaluated for northern (above 63°18'N) and southern (below 63°18'N) Norwegian stations, separately, in order to assess the benefits provided by the integration of numerical model information where the station density is lower. In Figure 6.4 the resulting monthly distribution of MAE values for the three methods is reported. In both subdomains, the median is generally lower for HCLIM+RK, especially in winter, as well as the range of outliers. It is worth noting that MAPE passes on average from 10% in the South to 13% in the North for HCLIM+RK, while the increase is slightly greater for MLRK and LWLR (from 11% to 15% for both). The lower error differences between southern and northern station validations could suggest that the accuracy of the HCLIM+RK reconstruction is less influenced by the variability in data distribution over the domain.

In order to further evaluate the sensitivity of the models to data distribution, a reconstruction test was performed for southern Norway, which is the subdomain with the highest and most homogenous station coverage. The LOO normals were computed both by considering the full data availability over the subregion (728 stations) and by reducing the data density to 70% and 50%. For both data reductions, the reconstruction was iteratively performed with ten random resamples of the original station availability and the resulting errors were computed as the average of the ten simulations. Even though the errors turned out to increase for all the methods with reducing data cover-

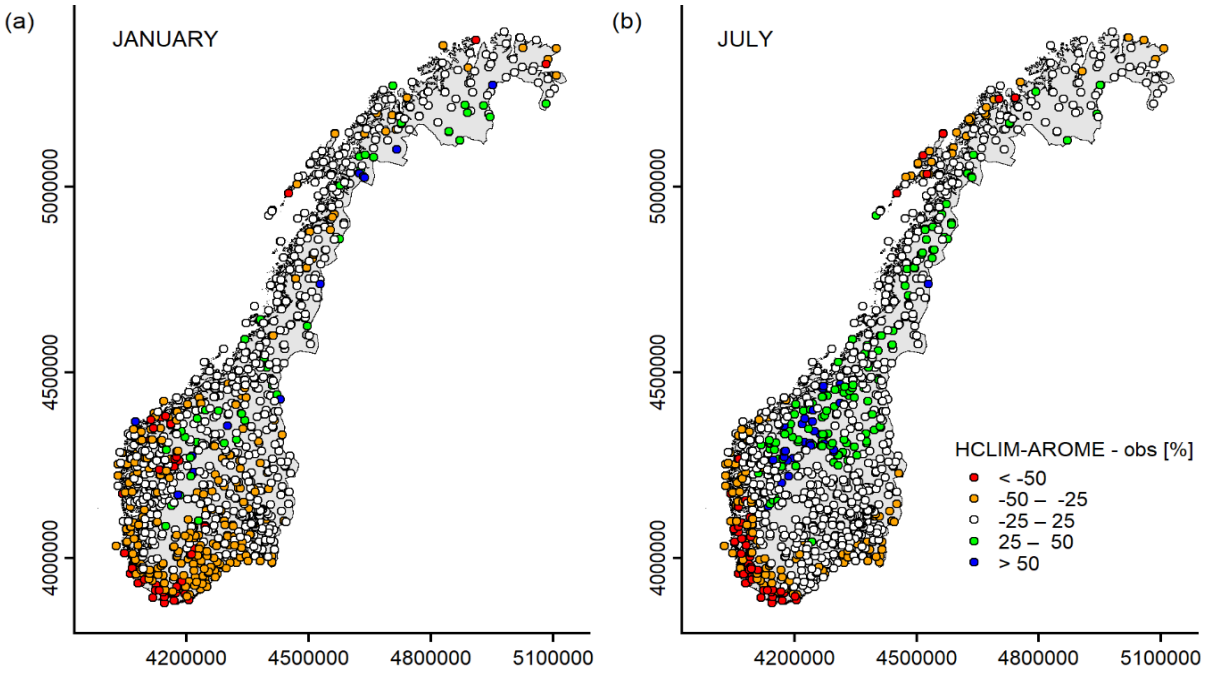

Figure 6.3: Distribution over Norway of the relative BIAS [%] of HCLIM-AROME dataset for a) January and b) July. The values are obtained by comparing the 1981–2010 monthly climatologies of HCLIM-AROME with the station normals and by normalising the differences for the average of numerical model and rain-gauge values.

age, the accuracy decrease obtained by using 70% and 50% of stations in comparison with the results provided by the original availability is lower for HCLIM+RK. In fact, MAPE for HCLIM+RK remains almost invariant (about 10%) with 70% of stations and it increases to 11% with halved density, while MLRK and LWLR turn out to be more affected by the sparse station coverage with average MAPE increasing for both methods from 11% to 12% with 70% of data availability and to 14% with halved coverage.
The above observations suggest that integrating in-situ information with numerical model output could be a valuable approach to improve the robustness of reconstructed climatological fields over Norway, especially in the North where the rain-gauge network gets sparser. In these areas in fact station-based methods are forced to extrapolate precipitation gradients by means of very few and far observations and they could produce unrealistic artefacts. On the contrary, the problem of areas scarcely covered by stations is less relevant for HCLIM+RK as observations are only used to correct model biases whose high spatial coherence allows to get reliable results even with rather low station density.
It is worth noting that HCLIM+RK climatologies have lower errors than MLRK and LWLR ones even if the HCLIM-AROME simulations on which they are based cover only a small fraction of the 1981–2010 period. The errors of HCLIM+RK estimates depend therefore both on HCLIM+RK method and on the approach we used to extend the numerical model dataset and make it representative of the 30-year reference period.

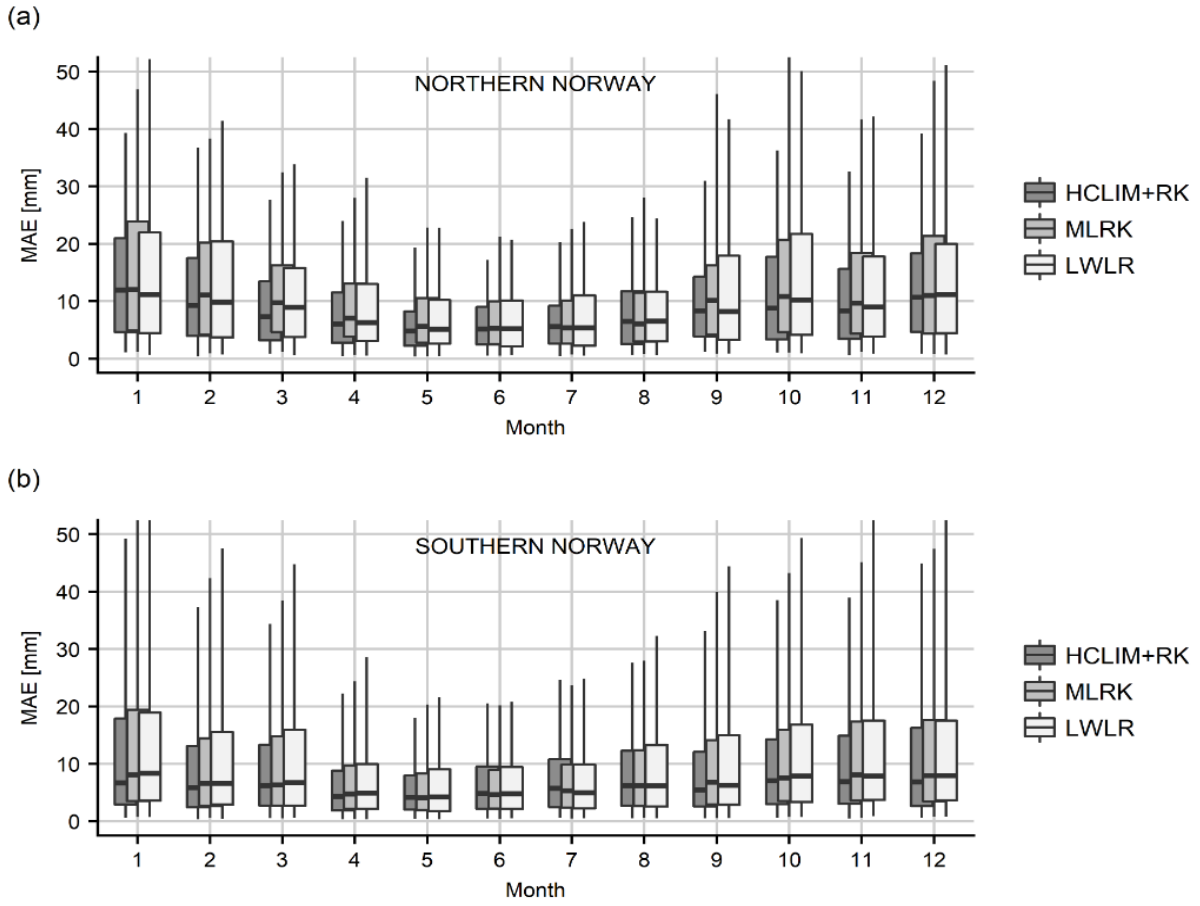

Figure 6.4: Monthly MAE distribution of the reconstructed normals by the three methods for stations located in a) northern Norway (above 63°18′N) and b) southern Norway (below 63°18′N). The boxes represent the inter-quartile range of the distribution where the median is reported by the bold line; the whiskers represent the 5–95% quantile range.

The contribution of the latter factor was evaluated by means of the station records. Specifically, the monthly series of all Norwegian station sites were reconstructed over the entire period 1981–2010 by considering only the station data corresponding to the years covered by HCLIM-AROME simulations and by applying the same method used for HCLIM-AROME. The 1981–2010 station normals resulting from the reconstructed records were then compared with the corresponding observed climatologies. The results are listed in Table 6.3 The BIAS is almost zero in all months proving that the reconstruction procedure did not lead to systematic under or overestimations, while correlation between estimated and observed monthly normals is always greater than 0.99. RMSE and MAPE, as averages over the months, are 4.3 mm and 3.3%, respectively, suggesting that the errors due to the reconstruction of missing data can be considered rather low in comparison with those ascribing to HCLIM+RK method. The small influence of the missing data period simulation on the HCLIM+RK accuracy can be explained by the fact that the reconstruction procedure is equivalent to multiplying

	1981–2010 station simulation			
	BIAS	MAE	MAPE	RMSE
1	0.1	3.7	3.5	5.6
2	0.2	3.4	4.1	5.3
3	0.1	3.0	4.1	4.6
4	0.1	2.2	3.8	3.2
5	0.0	1.9	3.1	2.7
6	0.1	2.2	3.1	3.0
7	0.0	2.7	3.0	3.7
8	0.1	3.2	3.2	4.5
9	0.1	3.0	2.8	4.3
10	0.0	3.4	2.8	5.0
11	0.0	3.2	2.9	4.7
12	0.0	3.6	3.5	5.4
MEAN	0.1	3.0	3.3	4.3

Table 6.3: Errors of the 1981–2010 monthly precipitation normals computed from the simulated records of the 1043 Norwegian stations. Except for MAPE, all the values are expressed in mm and BIAS is defined as the difference between simulation and observation.

the 2003–2016 climatologies by the ratio between the 1981–2010 and 2003–2016 normals of surrounding stations and this ratio shows a very high spatial coherence over the domain in all months. The distribution of this ratio for January and July at all the station sites is reported in Figure 6.5. The sites showing values strongly deviating from those of the neighbouring sites are more likely to be affected by inhomogeneities in their records than to represent actual outliers in climatological ratio.

The low errors provided by the extension of the 2003–2016 model records to the 1980–2010 window highlight that the same method we used to get the 1981–2010 climatologies can be applied also over other previous 30-year periods (e.g. 1971–2000, 1961–1990), even though the errors are expected to increase due to the lack of overlapping years of data and to the increasing probability of break occurrences in station records.

Furthermore, it is worth noting that the above findings prove the relative low contribution of missing period reconstruction on the errors of 1981–2010 HCLIM+RK climatologies, however they also suggest that the accuracy of the interpolation method is not expected to be drastically improved even with the availability of numerical model simulations over a longer period.

In addition to the model comparison by station validation, HCLIM+RK results were compared with the output of MLRK and LWLR also at grid point level in order to detect the specific features of the continuous precipitation field provided by the different approaches. In Figures 6.6 and 6.7 the differences of MLRK and LWLR gridded monthly precipitation climatologies for January (a) and July (b) in respect with

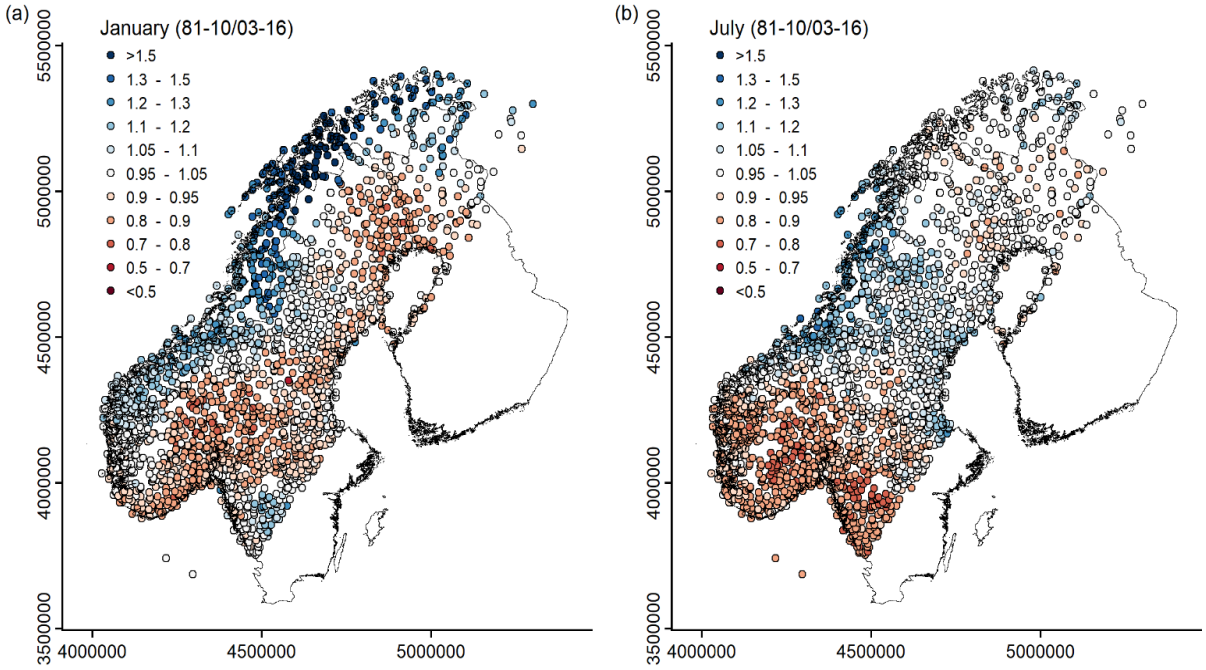

Figure 6.5: Distribution of the ratios between the 1981–2010 and 2003–2016 normals of all the stations in the database for a) January and b) July.

HCLIM+RK fields are reported. Both models show an overall tendency to underestimate winter precipitation especially along the coastal reliefs throughout the country, with the lowest estimates at the highest elevated grid points.

On average, MLRK and LWLR precipitation normals in January are lower than HCLIM+RK values of about 20 mm in northern Norway and of about 7 and 11 mm, respectively, in the South. However, if only points above 800 m a.s.l. are considered, the mean underestimation increases for both methods with the most negative discrepancies for LWLR in both subregions. As regards July normals, the discrepancies with HCLIM+RK are more spatially heterogeneous for both methods and relevant discontinuities are evident, especially in the LWLR climatology over central Norway where negative biases turn into relevant overestimations within very short distances. The drying tendency and wet outliers mostly occur over the lowest sampled areas where the local available information entering into MLRK and LWLR interpolation are not enough to model reliable precipitation gradients. This is particularly evident if the discrepancies with HCLIM+RK output are clustered for elevation ranges (Figure 6.8). In January they turn out to gradually increase with elevation for both MLRK and LWLR which could be a consequence of the decrease in data availability over the mountainous regions. The distribution of discrepancies for July does not show a significant trend over elevation ranges below 1200 m a.s.l. and the medians for MLRK and LWLR are almost comparable. At the grid cells above 1200 m a.s.l. (about 8% of the total) MLRK and LWLR provide on average higher July normals than HCLIM+RK and this discrep-

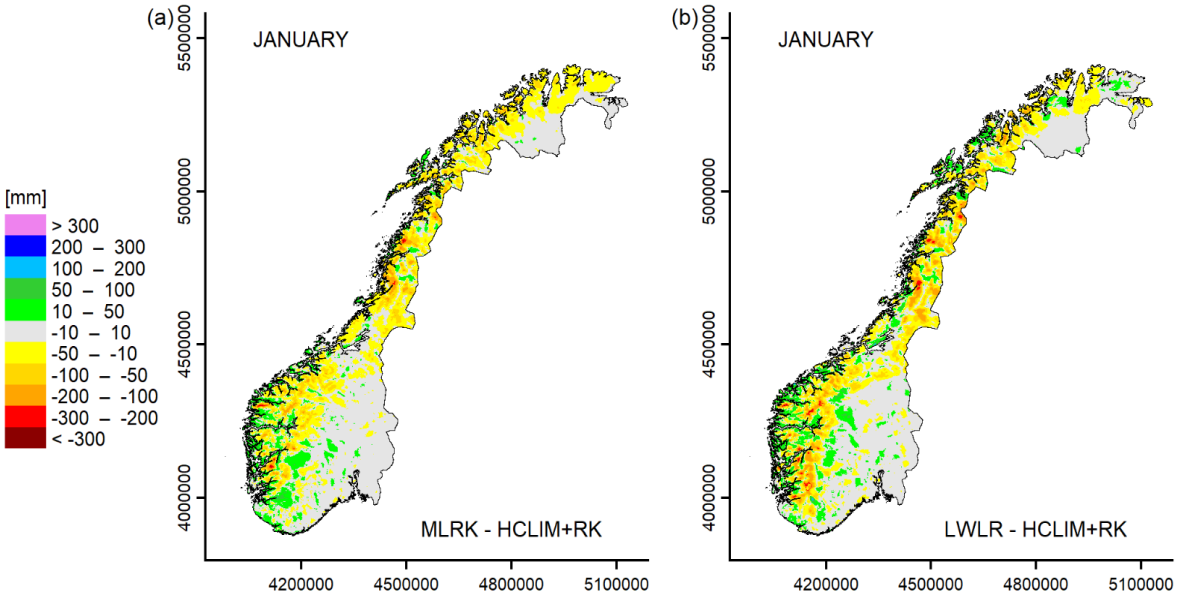

Figure 6.6: Discrepancies of a) MLRK and b) LWLR with HCLIM+RK in 1981–2010 gridded precipitation climatologies for January.

ancy could be partly due to the original negative bias affecting the HCLIM-AROME summer precipitation fields (Figure 6.3b) over the mountainous coastal regions not completely corrected by the RK on station data. It is worth noting that in both months the interquartile ranges and extremes of outliers are generally greater for LWLR suggesting a major instability of its modelled fields. Moreover, the LWLR discrepancies are expected to be even greater. In fact over the least sampled areas this method reconstructed some negative precipitation, which were automatically corrected within the algorithm by substituting the regression with a simple weighted average based on station distance. This correction could be suitable for adjusting single points, whereas it could give rise to evident discontinuities when the extent of negative precipitation areas gets wider. LWLR negative estimates are mostly likely to occur if no significant elevation gradients exist at the target cell and/or if the uneven station distribution leads to misjudge the local precipitation-elevation relationship. This problem is partly reduced in MLRK thanks to the inclusion of more geographical predictors and to the larger spatial scales considered in the interpolation, however some isolated and slightly negative normals still occur for some months.

The above findings suggest that the interpolated climatologies provided by the observation-based methods are more likely to be affected by discontinuities and outliers where the station coverage gets sparser and pure extrapolation is performed, whereas the integration of numerical model fields into the HCLIM+RK procedure could help to reconstruct more stable precipitation gradients and to reduce the underestimation of winter normals over the highest elevated areas of Norway.

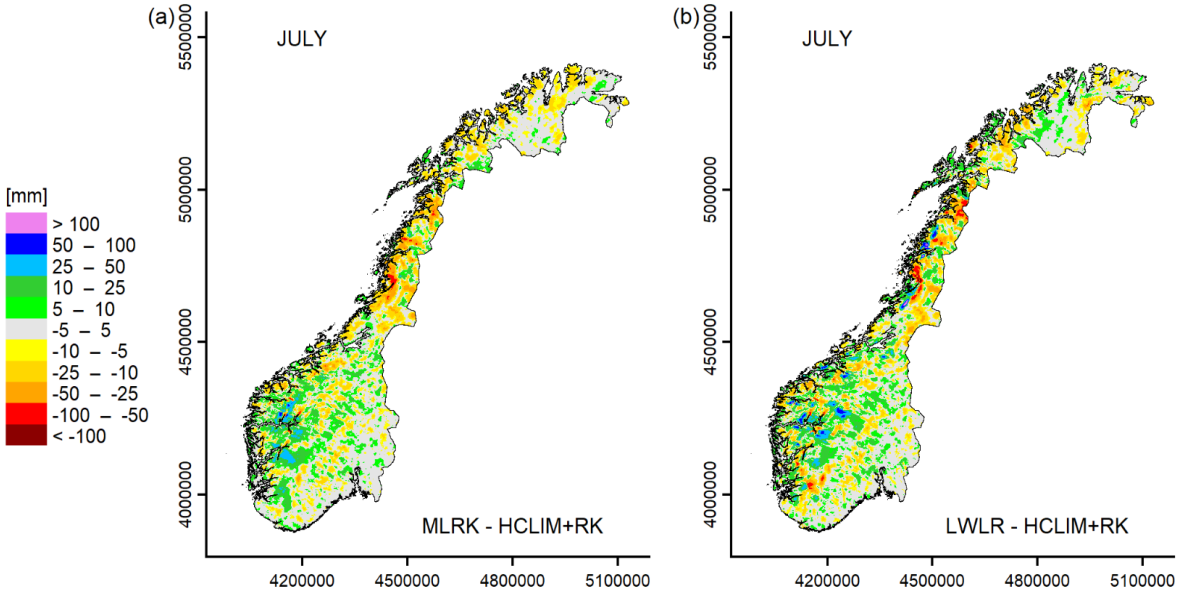

Figure 6.7: Discrepancies of a) MLRK and b) LWLR with HCLIM+RK in 1981–2010 gridded precipitation climatologies for July.

6.4.2 HCLIM+RK 1981–2010 monthly precipitation climatologies over Norway

According to the results discussed in the above section we selected the HCLIM+RK as the reference climatology for Norway and we analysed it to assess the spatial distribution of precipitation over the country. The HCLIM+RK seasonal and annual precipitation climatologies for the period 1981–2010 are presented in Figures 6.9 and 6.10, respectively. The average seasonal precipitation normals (winter, spring, summer and autumn) over the whole domain are 339, 220, 263 and 364 mm, respectively, while the average annual precipitation is 1186 mm. The mean precipitation totals reconstructed by MLRK on both seasonal and annual scales are in agreement with HCLIM+RK results (correlation values always above 0.9) even if they are underestimated in all cases of about 10% in respect to the HCLIM+RK reference values. The spatial distribution of HCLIM+RK precipitation climatologies is mainly dominated by a strong and well-defined West-to-East gradient along the whole country in all seasons, with the highest normals in correspondence to the coastal reliefs, acting as orographic barrier to the wet air masses from the sea, and a quick transition to rather dry conditions over the leeward side and inland. The wettest area in Norway is located around the Ålfotbreen glacier (1385 m), near the Nordfjorden. The mean annual precipitation over this area exceeds 5700 mm with the greatest contributions from winter precipitation accounting for more than 30% of annual totals. Another very wet region is depicted around the Svartisen glacier in northern Norway with annual precipitation values around 3000 mm. On the contrary the driest conditions occur between Otta and Gudbrand Valleys,

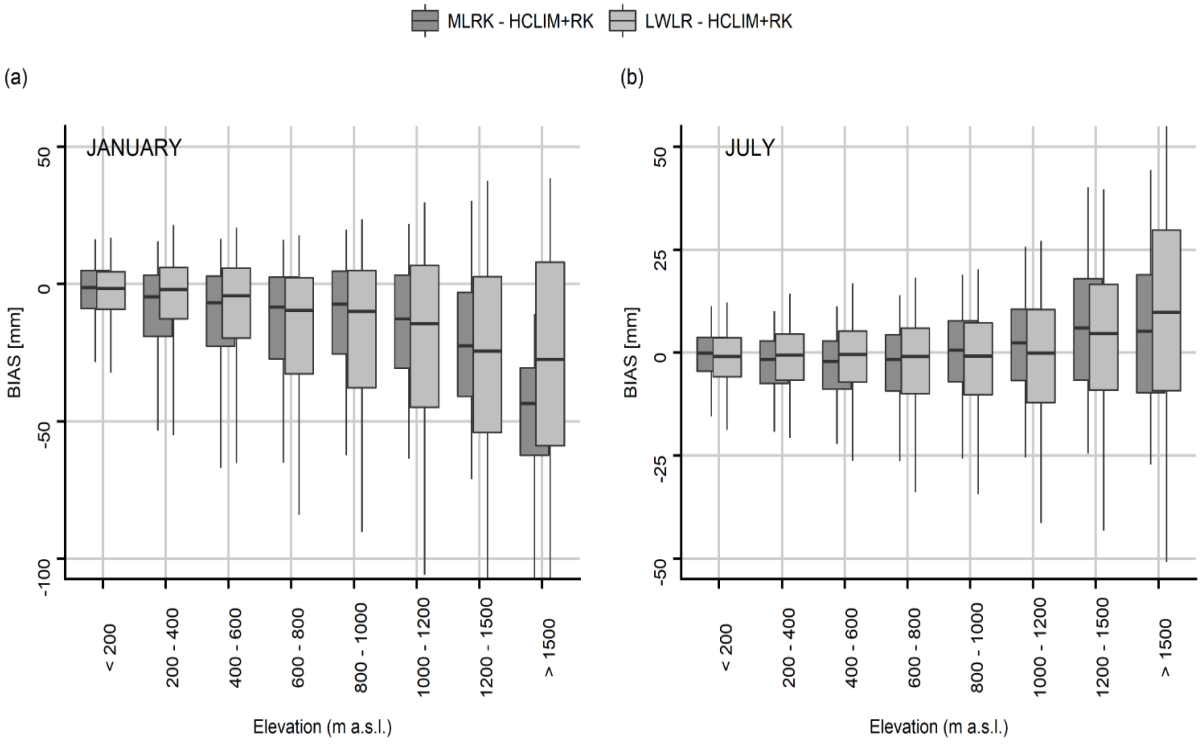

Figure 6.8: Distribution of MLRK and LWLR discrepancies with HCLIM+RK gridded a) January and b) July climatologies on elevation ranges. The boxes represent the interquartile range of the distribution where the median is reported by the bold line; the whiskers represent the 5–95% quantile range.

located on the bottom of the leeward side of the west mountain chain, where the absolute minimum annual normal is reconstructed (243 mm). Furthermore, precipitation is found to reduce significantly also towards the northernmost part of the country, where the annual normals decrease below 300 mm, especially over Finnmark. It is worth noting that over northern Norway a remarkable positive precipitation gradient occurs along the whole year from the flat regions of inland to the more mountainous coastal areas. This gradient is reconstructed thanks to the background information supplied by HCLIM-AROME numerical model fields, while it is not captured or turns negative if only in-situ observations enter into the climatological computation, as in MLRK and LWLR, whose extrapolation over under-sampled areas leads to underestimations in most cases. By taking into account the grid cells above 69°N and 500 m a.s.l. only, the average annual precipitation in HCLIM+RK climatology is in fact about 300 mm higher than in MLRK and, except for summer, the HCLIM+RK seasonal precipitation normals are almost twice greater than MLRK estimates.

Further interesting information about Norwegian climate can be retrieved from the distribution of annual precipitation cycles over the country (Figure 6.11). The average yearly patterns were computed over consecutive 10000 km^2 areas covering the whole domain, after filtering the normal cycle of each grid point by means of a trigonometric

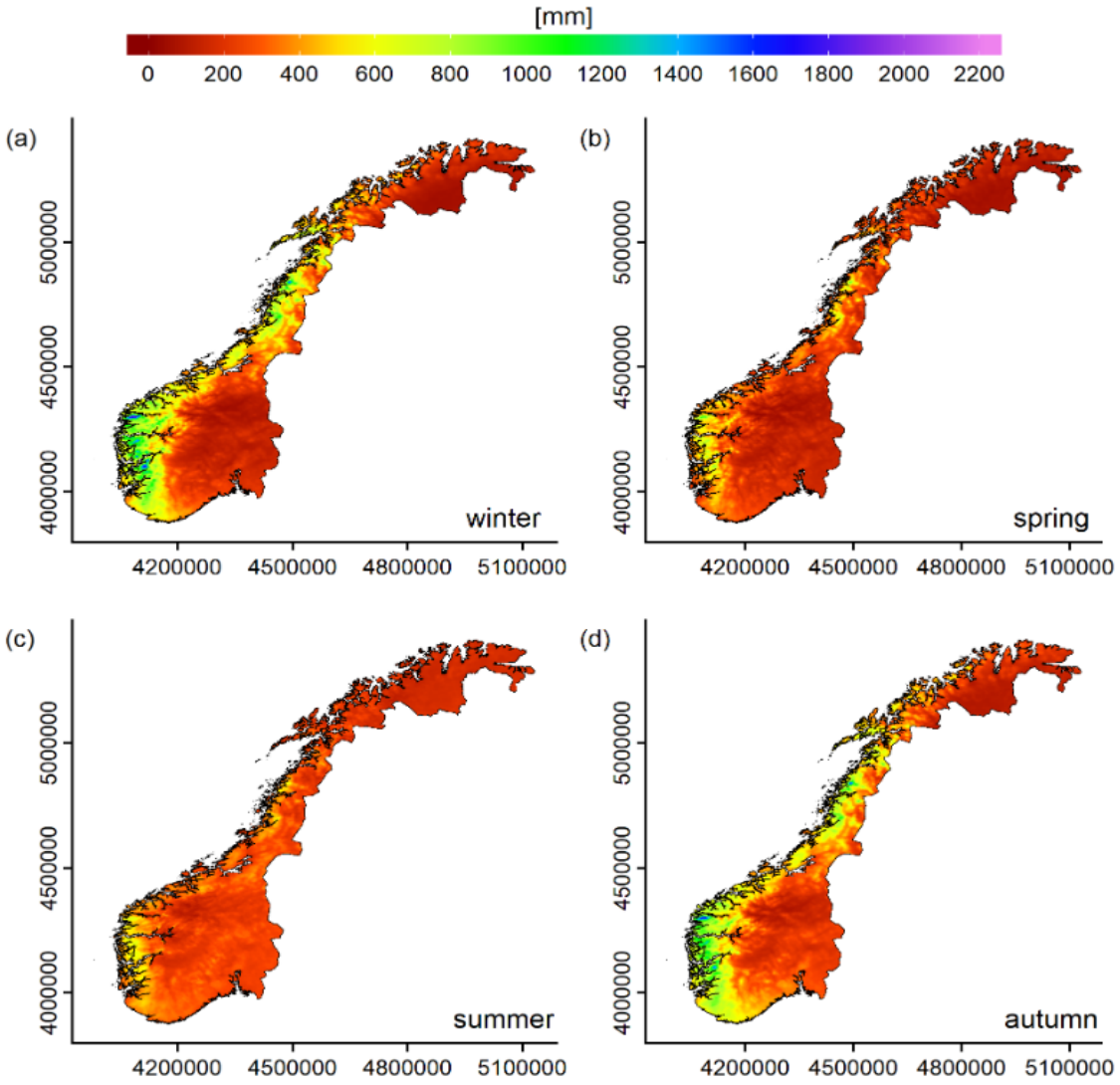

Figure 6.9: Seasonal HCLIM+RK precipitation climatologies.

function in order to reduce the discontinuities from one month to another. All the Norwegian subregions which are close to the sea experience the highest relative monthly contributions to annual precipitation in winter and the lowest ones in late spring or summer and this annual pattern turns out to be almost constant with latitude. Moving from the coast towards the inland, the climate turns to be characterised by maximum precipitation in summer and minima in winter or early spring. Moreover, the relative contributions of summer precipitation to the annual cycles over the inner areas increase towards the North of about 3%, except for the northernmost subregions where the very dry conditions lead to more homogeneous precipitation regimes along the year and slightly higher contributions in late summer and autumn. Also the area around Oslo Fjord on the southern coast experiences a more smoothed yearly cycle with the main contributions in late summer when the convective phenomena prevail. In Figure 11, the annual cycles over the considered subregions are superimposed to the spatial dis-

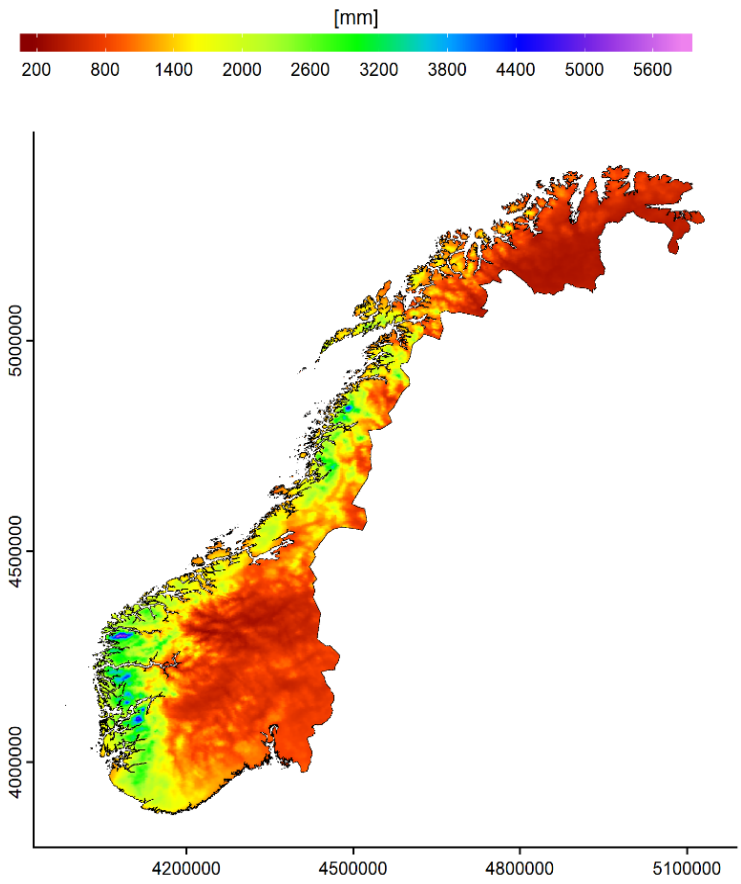

Figure 6.10: Annual HCLIM+RK precipitation climatologies.

tribution of the ratio between winter and summer precipitation normals at each point of the 1 km grid. It confirms the existence of two main distinctive climatic regimes with the prevalence of orographic enhancement mechanisms for precipitation over the mountainous coast and of more continental conditions over the inland. More specifically, the winter to summer precipitation ratio ranges on average between 1.5 and 3 over the coastal Norway, with the greatest values over the highest areas, whereas the ratio is below the unity over the inland where summer precipitation normals become twice the winter ones for a large fraction of grid cells. In addition, it is worth noting the occurrence of a "transition zone" crossing the whole country from North to South characterised by quite comparable contributions of winter and summer precipitation amounts.

6.5 Conclusions

The 1981–2010 monthly precipitation climatologies over Norway were computed at 1 km resolution by applying a new combined interpolation approach, named HCLIM+RK. The method joins a database of quality-checked rain-gauge observations

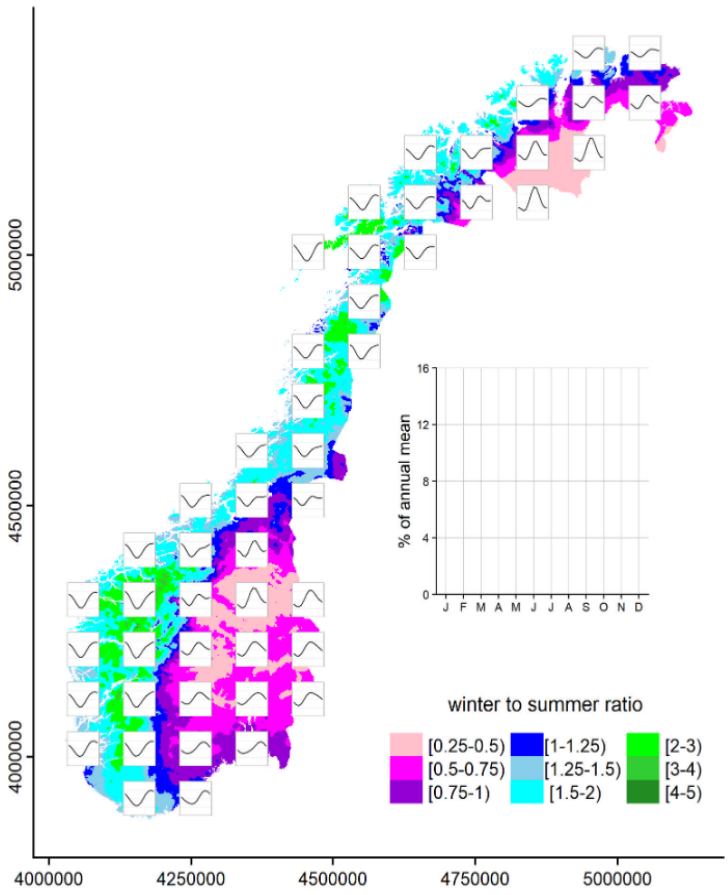

Figure 6.11: Distribution of winter (DJF) to summer (JJA) precipitation ratio over the domain and of the average yearly precipitation cycles over different 10000 km^2 subdomains covering Norway. The inset box defines the range of axes that is the same for all the plots. The values are expressed as percentage to the total annual precipitation.

with the 1981–2010 monthly precipitation fields derived from the 2003–2016 reanalysis driven regional climate HCLIM-AROME model at 2.5 km resolution.
The ability of this approach to deal with the uneven data coverage of Norway and to provide reliable precipitation patterns also where station density gets lower was evaluated by the comparison with two interpolation schemes using only observations: MLRK and LWLR. HCLIM+RK turned out to provide the lowest errors in reconstructing the station normals in all months with MAPE ranging from 8.7% (September) to 13.5% (March). Moreover, the significant biases affecting HCLIM-AROME numerical model fields were almost removed thanks to the integration with station observations. The better performance of HCLIM+RK are also evident if the reconstruction is evaluated on northern and southern stations, separately. MLRK and LWLR provide higher errors and larger ranges of outliers, especially in the North where the network is sparser, suggesting the greater influence of the variation in data availability on their accuracy. A sensitivity test was also performed on the three methods and it confirmed the greater stability of HCLIM+RK results with varying station distribution and den-

sity. The reconstruction errors of HCLIM+RK remained almost unaltered by reducing the data availability, while MLRK and LWLR errors increased more significantly.
By comparing the gridded precipitation climatologies for January and July, MLRK and LWLR showed systematic underestimations over coastal regions in respect with HCLIM+RK reconstruction, especially in winter with increasing grid cell elevation. In addition, the observation-based models are affected by relevant spatial discontinuities in precipitation distribution and some negative estimates also occur, especially for LWLR, where the rain-gauge coverage is sparser.
The contribution to the global HCLIM+RK errors provided by extrapolating the 1981–2010 normals from the 2003–2016 HCLIM-AROME fields was also evaluated by simulating the same reconstruction approach on station data. The series reconstruction provided much lower errors than those of the whole HCLIM+RK procedure (average MAPE was 3% and 11%, respectively) and the overall coherence in climatological ratios among stations proved the robustness of HCLIM-AROME run extension.
The HCLIM+RK precipitation climatologies for the period 1981–2010 are characterised by a sharp West-to-East transition from the wet mountainous coast, which is interested by orographic precipitation regime, to the drier inland experiencing more continental climate. Even if the precipitation normals decrease significantly in the northernmost part of Norway, a positive gradient with elevation is preserved thanks to the information provided by HCLIM-AROME numerical fields. The total lack of observations at the high-elevated areas of northern Norway leads in fact the observation-based approaches to mainly extrapolate decreasing precipitation with altitude.
The distributions of annual cycles as well as of the winter-to-summer ratios obtained from HCLIM+RK climatologies confirm the main precipitation gradients over the country and depict the more specific climatic features of the different subregions. The presented findings prove that the interpolation approach combining numerical model information with in-situ observations allows to better handle with the uneven station network over Norway and to significantly increase the accuracy of resulting climatologies, especially over the most remote regions. Future analyses aiming at identifying the sources of numerical model biases as well as the integration of the most recent automatic station observations at remote sites could help to further increase the accuracy of the available monthly precipitation climatologies over the country.

6.6 References

- Berthou, S., Kendon, E. J., Chan, S. C., Ban, N., Leutwyler, D., Schär, C., and Fosser, G. (2018). Pan-European climate at convection-permitting scale: a model intercomparison study. Climate Dynamics. doi:10.1007/s00382-018-4114-6

- Boer, E. P. J., De Beurs, K. M., and Hartkamp, A. D. (2001). Kriging and thin plate splines for mapping climate variables. International Journal of Applied Earth Observation and Geoinformation, 3, 146–154. doi:10.1016/S0303-2434(01)85006-6

- Crespi, A., Brunetti, M., Lentini, G., and Maugeri, M. (2018). 1961-1990 high-resolution monthly precipitation climatologies for Italy. International Journal of Climatology, 38, 878–895. doi:10.1002/joc.5217

- Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P. (2002). A knowledge based approach to the statistical mapping of climate. Climate Research, 22, 99–113. doi:10.3354/cr022099

- Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., and Pasteris, P. A. (2008). Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. International Journal of Climatology, 28, 2031–2064. doi:10.1002/joc.1688

- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, D. P., and Bechtold, P. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553–597. doi:10.1002/qj.828

- Dyrrdal, A. V., Stordal, F. and Lussana, C. (2018). Evaluation of summer precipitation from EURO-CORDEX fine-scale RCM simulations over Norway. International Journal of Climatology, 38, 1661–1677. doi:10.1002/joc.5287

- Dyrrdal A. V., Lenkoski A., Thorarinsdottir T. L., and Stordal F. (2015). Bayesian hierarchical modeling of extreme hourly precipitation in Norway. Environmetrics, 26, 89–106. doi:10.1002/env.2301

- Foresti, L., and Pozdnoukhov, A. (2012). Exploration of alpine orographic precipitation patterns with radar image processing and clustering techniques. Meteorological Applications, 19, 407–419. doi:10.1002/met.272

- Golzio, A., Crespi, A., Bollati, I. M., Senese, A., Diolaiuti, G. A., Pelfini, M., and Maugeri, M. (2018). High-Resolution Monthly Precipitation Fields (1913–2015) over a Complex Mountain Area Centred on the Forni Valley (Central Italian Alps), Advances in Meteorology, 2018, ID 9123814, pp. 17. doi:10.1155/2018/9123814

- Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology, 228, 113–129. doi:10.1016/S0022-1694(00)00144-X

- Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, Journal of Geophysical Research, 113, D20119. doi:10.1029/2008JD010201

- 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.

- Henn, B., Newman A. J., Livneh, B., Daly, C., and Lundquist, J. D. (2018). An assessment of differences in gridded precipitation datasets in complex terrain. Journal of Hydrology, 556, 1205–1219. doi:10.1016/j.jhydrol.2017.03.008

- Isotta, F., Vogel, R., and Frei, C. (2015). Evaluation of European regional reanalyses and downscalings for precipitation in the Alpine region. Meteorologische Zeitschrift, 24, 15–37. doi:10.1127/metz/2014/0584

- Jermey, P. M., and Renshaw, R. J. (2016). Precipitation representation over a two-year period in regional reanalysis. Quarterly Journal of the Royal Meteorological Society, 142, 1300–1310. doi:10.1002/qj.2733

- Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., and Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. Scientific Data, 4: 170122. doi:10.1038/sdata.2017.122

- Lind, P., Lindstedt, D., Kjellström, E., and Jones, C. (2016). Spatial and temporal characteristics of summer precipitation over central europe in a suite of high-resolution climate models. Journal of Climate, 29(10), 3501–3518. doi:10.1175/JCLI-D-15-0463.1

- Lussana, C., Saloranta, T., Skaugen, T., Magnusson, J., Tveito, O. E., and Andersen, J. (2018). seNorge2 daily precipitation, an observational gridded dataset over Norway from 1957 to the present day. Earth System Science Data, 10, 235–249. doi:10.5194/essd-10-235-2018

- Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., ... and Voldoire, A. (2013). The surfexv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. Geoscientific Model Development, 6, 929–960. doi:10.5194/gmd-6-929-2013

- Mohr, M. (2008). New routines for gridding of temperature and precipitation observations for "seNorge.no", Met. no Report, 8, the Norwegian Meteorological Institute, Oslo, Norway.

- Mohr, M. (2009). Comparison of versions 1.1 and 1.0 of gridded temperature and precipitation data for Norway, Norwegian Meteorological Institute, met no note, 19, the Norwegian Meteorological Institute, Oslo, Norway.

- Müller, M., Homleid, M., Ivarsson, K., Køltzow, M. A., Lindskog, M., Midtbø, K. H., Andrae, U., Aspelien, T., Berggren, L., Bjørge, D., Dahlgren, P., Kristiansen, J., Randriamampianina, R., Ridal, M., and Vignes, O. (2017). AROME-MetCoOp: A Nordic Convective-Scale Operational Weather Prediction Model. Weather and Forecasting, 32, 609–627. doi:10.1175/WAF-D-16-0099.1

- Prein, A. F., and Gobiet, A. (2017). Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. International Journal of Climatology, 37, 305–327. doi:10.1002/joc.4706

- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V. (2011). The AROME-France Convective-Scale Operational Model. Monthly Weather Review, 139, 976–991. doi:10.1175/2010MWR3425.1

- Simmons, A. J., Berrisford, P., Dee, D. P., Hersbach, H., Hirahara, S., and Thépaut, J. (2017). A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets. Quarterly Journal of the Royal Meteorological Society, 143, 101–119. doi:10.1002/qj.2949

- Tveito, O. E., Bjørdal, I., Skjelvåg, A. O., and Aune, B. (2005). A GIS-based agro-ecoglogical decision system based on gridded climatology. Meteorological Applications, 12, 57–68. doi:10.1017/S1350482705001490

Chapter 7

Conclusions and future developments

The recent increase in the global mean temperature is found to influence all the components of the complex climate system including the hydrological cycle. The future climate projections highlight changes in the precipitation regime at global level both in intensity and frequency, in particular an increasing occurrence of long dry-spells and short events of extreme precipitation is expected (IPCC, 2013). The increasing temperature is supposed to lead to a decrease in solid precipitation contributions together with the enhancement of glacier melting, permafrost reduction and evapotranspiration rates. These phenomena are likely to influence the economical and social development as well as the safety of the most vulnerable areas, such as mountains or marine coasts, by increasing the soil instability, the probability of landslides, floods or avalanches and by modifying the water disposal. However the magnitude and effects of climate changes and impacts are highly variable on regional and local scales.
In this framework, the availability of high-resolution information about precipitation distribution over specific areas of interest and their temporal evolution over the recent past is crucial to analyse the spatial behaviour and long-term variability of the climatic signal. At this aim, long and accurate observation records are needed together with suitable interpolation methods able to capture the climatic signal at local level and to project it onto regular grids (Jones and Hulme, 1996; Daly et al., 2002). However, the reduced availability of meteorological data over the remote regions, especially the high-elevation areas, and for the most ancient periods could highly limit the reconstruction ability and the robustness of gridded results. For these reasons, the main aims of my PhD activity were:

- to set up a dense and quality-checked archive of monthly precipitation records for the whole Italian territory in order to produce the 1961-1990 monthly precipitation climatologies for Italy on a 30-arc second resolution grid by implementing an interpolation approach able to model the local interaction between precipitation and the complex Italian orography;

- to apply the anomaly method on databases of long and homogenised station records to reconstruct and analyse the long-term precipitation evolution over vulnerable Italian regions: Sardinia and the upper part of the Adda river basin in Central Alps with an additional focus over the Forni glacier;
- to evaluate possible applications of the gridded datasets of secular monthly precipitation, such as the reconstruction of specific past episodes of intense precipitation;
- to compute the 1981-2010 monthly precipitation climatologies over Norway by implementing an interpolation approach able to deal with the uneven distribution of the meteorological stations over the country, especially in the northernmost regions and over the mountainous areas.

The PhD activities allowed also to demonstrate the suitability of the developed methods to be adapted to the climate reconstruction on different domains characterised by specific geographical and climatic features. The detailed results achieved are point-by-point summarised in the following subsections.

7.1 The observation database and the Italian precipitation climatologies

The first step to compute the high-resolution Italian precipitation climatologies concerned the construction of a dense and quality-checked observation database covering the whole national territory. It was set up starting from a great number of daily and monthly records (together with the corresponding metadata) which were retrieved by the research group over more than 10 years of data collection and rescue activities. Most series were derived from the archives of the former Italian Hydrographic Service and from the Italian regional and subregional agencies which took in charge the management of the meteorological network after its closure at the end of the 1980s. Other sources of data were the Italian Air Force and previous projects of data rescue and digitisation. Moreover, some data from Switzerland and Slovenia were included in order to improve the station coverage along the Italian boundaries.
The original database included more than 7000 monthly series. All the series were controlled in order to detect duplicates and possible merging for records available from more than one source. In addition all the station metadata, location and elevation, were checked and, whenever possible, the errors were corrected. An automatic procedure of quality-check was also developed in which each measured series is compared to a simulated one by means of the neighbouring station data. The comparison and the evaluation of reconstruction errors allowed to detect and invalidate the gross errors, such as

outliers, spurious sequences of null precipitation as well as evident inhomogeneities; the series showing low agreement with the simulated ones even after the corrections and those records with less than 10 years of available data were definitely discarded from the climatology computation. After the quality control, the same procedure was also used to complete the monthly gaps of the station series in the climatological period of reference (1961-1990) in order to prevent the normals being biased because of the uneven data coverage of the 30-year interval.
The final version of the database counted more than 4500 series which was integrated by the 1961-1990 normals of station sites available at the meteorological services of Austria, Switzerland and France.
The climatologies were produced on a 30-arc second resolution Digital Elevation Model (DEM) covering Italy and the trans-national portion of Po basin by applying a Local Weigthed Linear Regression (LWLR) of precipitation *versus* elevation, which is based on the PRISM modelling framework (Daly et al., 1994). The local approach was justified by the complex orographic features of the Italian territory which could influence the precipitation gradients on a very small scale. This procedure allows in fact to define the regression coefficients at each grid point by giving more importance to the closest stations with the greatest orographic similarity (elevation, sea distance, slope steepness and orientation) with the target location. The station weights are Gaussian functions of the geographical parameters whose decays are set month-by-month for each cell by an iterative optimisation procedure. In addition, in order to remove the too fine orographic details in respect with the spatial scales at which the atmospheric processes are expected to interact with the surface, a smoothed version of the DEM was considered in which the elevation of each cell is computed by a distance weigthed mean of the surrounding cell elevations.
The accuracy of LWLR was computed by the leave-one-out validation on station normals and it turned out to be greater than that provided by interpolation methods based on global regression or not modelling the elevation-precipitation relationship at all. However, LWLR could produce artefacts over orographically complex areas covered by a very low number of rain-gauges.
As an alternative interpolation method to LWLR, the Local Regression Krging (LRK) was tested showing very similar performances. LRK considers a greater spatial scale which reduces the occurrence of artefacts over the lowest sampled regions in respect with LWLR even though it provides a lower spatial detail in the reconstructed gradients.
Both versions of the 1961-1990 high-resolution precipitation climatologies allowed to analyse the mean annual cycles and precipitation regimes over Italy and to assess the location of climatological extremes. The gridded dataset of LWLR monthly climatologies has been released on the ISAC-CNR website for research and operative purposes.

It represents the most recent dataset providing the precipitation normals for the whole national territory at a very high-resolution scale. However, the updates of the climatologies to the most recent 30-year period (1981-2010 and, in few years, 1991-2020) will be required so that a relevant activity of rescue and integration of the most recent data from the currently operating meteorological networks of all the Italian regions is going to be planned in the next future.

7.2 Variability and long-term trends of the monthly precipitation (1800-2016) and runoff (1845-2016) over the Adda river basin

A 30-arc second resolution grid of 1800-2016 monthly precipitation series covering the upper part of Adda river basin was obtained by applying the anomaly method to an observation database of long and homogenised rain-gauge records. In order to maximise the data availability over the most remote period, digitisation activities of the hardcopy yearbooks were performed and the available series over a wide area around the study domain were considered. In addition, the data from the automatic weather stations were integrated and merged with the records from the previous mechanical rain-gauges of some regional networks in northern Italy allowing to extend the data coverage to present.
All the series underwent the quality-check procedure (Crespi et al., 2018) and, in order to avoid the influence of non-climatic signals on the trend evaluation, their homogeneity was checked and the most relevant breaks were corrected by means of the Craddock test.
The 1800-2016 gridded dataset was obtained by superimposing the monthly climatologies and the fields of monthly anomalies and it was used to compute the 1800-2016 areal total monthly precipitation series for the basin. The accuracy of the results and its variability in relation to the changes in station coverage over the whole analysed period was evaluated by the leave-one-out validation on station observations and by performing a sensitivity test on the areal basin precipitation series reconstruction. In all cases, the anomaly method allows to obtain a reliable signal also for the very beginning period and the accuracy rapidly increases after the second half of the 19^{th} century thanks to the increasing number of operating stations over the study domain. Theil-Sen and Mann-Kendall tests on annual and seasonal areal precipitation series pointed out significant negative trends for the annual and autumn series only. However if the very first decades were excluded from evaluation the annual trend is still negative but turns out to be not significant. The comparison with runoff (Ranzi et al., 2017) was performed on annual scale from 1845 when the available runoff records starts and it

highlighted a significant negative trend in runoff coefficients of -0.6% decade^{-1}, as a consequence of the remarkable decrease in annual runoff from the basin of about -11 mm decade^{-1}. This behaviour could be mainly ascribed to the increase in evapotranspitation driven by both the increasing temperature and the changes in land-use and coverage occurred over the last century on the basin (Birkinshaw et al., 2014). However, further analyses are needed i) to assess the role of natural and anthropogenic forcings, ii) to quantify the contribution of increasing glacier melting on the basin runoff and iii) to investigate the possible underestimation of winter precipitation due to the lower station coverage of mountainous areas and the difficulty of high-elevation rain-gauges to measure the solid precipitation (Sevruk et al., 2009; Eccel et al., 2012). As an example of application of the dataset, specific events of intense precipitation in the past were reconstructed at monthly scale, in particular the event of November 2002 was analysed and the robustness of the resulting spatial field was discussed. Further developments of the presented study will concern the improvement of daily data availability over the area, especially before 1951, by the digitisation of precipitation records still in hardcopy yearbooks at the aim of extending the analyses to the daily resolution. The same methodology was also applied to reconstruct a gridded dataset of 1913-2015 monthly precipitation series over a smaller area centred over the Forni glacier where some additional station data were retrieved. The resulting dataset allowed to assess the local trends and variability of the areal precipitation series and, in collaboration with another PhD student, it was used to compare the estimated precipitation values for the period 2005-2015 with the measures provided by the Automatic Weather Station located on the glacier tongue.

7.3 1922-2011 monthly precipitation records over Sardinia: long-term trends and variability

In order to analyse the variability and trends of monthly precipitation over one of the main Italian Mediterranean islands, the anomaly method was applied to provide a 30-arc second resolution dataset of 1922-2011 monthly precipitation series for Sardinia. A dense database of long rain-gauge observations was retrieved from the archive of the Hydrographic Service, integrated with the records which were already available in the database used for the construction of Italian climatologies and extended in the past by newly digitised data. The resulting 367 monthly series were analysed for quality and homogeneity before performing the interpolation procedures. In particular, spurious monthly values were invalidated, metadata checked and corrected, while 27 series containing relevant breaks due to the station relocations or the merging of close rain-gauge records were homogenised.

The trend evaluation was performed on the 1922-2011 average regional precipitation series retrieved from the final gridded dataset obtained from the combination of 1961-1990 climatologies and the interpolated anomalies. The trends are in agreement with previous studies concerning other Mediterranean regions: annual and winter precipitation records show negative long-term trends of -2.3% and -4.1% decade^{-1}, respectively, with significance levels above 99%, while the tendency for the summer season is positive even though not statistically significant (e.g. Freidas et al., 2007; Gonzàlez-Hidalgo et al., 2011; Brunetti et al., 2012). The same outcome was observed by the evolution of annual precipitation normals over consecutive 30-year periods spanning the whole reconstructed interval.
The gridded dataset is also available for the analysis of specific episodes of intense precipitation or dry spells occurred over the region in the past. The recovery of new digitised data for the past will allow to extend back the reconstruction and the application of the anomaly-based procedure on a daily database could be useful to analyse the variability of extremes and the long-term precipitation evolution on a finer time scale.

7.4 The integration of numerical model datasets with in-situ observations for the reconstruction of high-resolution climatologies

The 1981-2010 monthly precipitation climatologies over Norway were computed on a regular grid of 1 km resolution by a new spatialisation scheme called HCLIM+RK combining the monthly precipitation fields retrieved from HCLIM-AROME numerical model dataset (Lind et al., 2016) with the rain-gauge observations by means of the residual kriging interpolation. The available daily records over Norway and surrounding countries were retrieved from the MET Norway Climate Database and the ECA&D archive, analysed for quality, subjected to a daily gap-filling procedure and converted into monthly resolution. These activities allowed to get a large observation database covering the whole Scandinavian peninsula with more than 1000 stations located in Norway. However, due to the severe climate conditions together with the orographic complexity, the availability of rain-gauges is significantly reduced over mountains and in the remote northern regions. The choice of the combined method helps to overcome the limits in climatology reconstruction due to the very sparse meteorological network of Norway and, at the same time, the use of available observations allows to correct the precipitation biases affecting the numerical model outputs.
The leave-one-out validation on station normals and the intercomparison with conventional interpolation methods using in-situ observations only proved the higher accuracy of HCLIM+RK climatologies and the lower sensitivity to the variations in data

coverage. In particular, thanks to the information provided by the numerical model fields where no other observations are available, HCLIM+RK allows to reduce the underestimation of precipitation at the high-elevation regions generally affecting the station-based datasets (Lussana et al., 2018).
The gridded dataset was used to discuss the climatic zones of Norway and their features in terms of precipitation seasonality and magnitude. In particular, they point out the strong west-to-east gradient occurring all along the country between the wet west coast, where mountains act as orographic barrier for the moist oceanic air masses, and the drier inland featuring more continental conditions. The grid of 1981-2010 monthly precipitation climatologies over Norway has been freely released at https://zenodo.org/record/1304042#.Wzti8Bx9i90 for both applicative and research purposes.
The work allowed to suggest a promising interpolation approach which could be applied over any domain unevenly covered by meteorological station networks and having available high-resolution numerical model datasets. At the same time, the settlement of new automatic stations at high-elevation sites is expected to improve the data availability for the remote regions in the next future and to contribute in increasing the accuracy of the climatological estimates.

7.5 References

- Birkinshaw, S.J., Bathurst, J.C., and Robinson, M. (2014). 45 years of non-stationary hydrology over a forest plantation growth cycle, Coalburn catchment, Northern England, Journal of Hydrology, 519, 559-573. doi:10.1016/j.jhydrol.2014.07.0502014

- Brunetti, M., Caloiero, T., Coscarelli, R., Gullá, G., Nanni, T., and Simolo, C. (2012). Precipitation variability and change in the Calabria region (Italy) from a high resolution daily dataset. International Journal of Climatology, 32, 57-73. doi:10.1002/joc.2233

- Crespi, A., Brunetti, M., Lentini, G., and Maugeri, M. (2018). 1961-1990 high-resolution monthly precipitation climatologies for Italy. International Journal of Climatology, 38, 878–895. doi:10.1002/joc.5217

- Crespi, A., Brunetti, M., Maugeri, M., Ranzi, R., and Tomirotti, M. (2018). 1845–2016 gridded dataset of monthly precipitation over the upper Adda river basin: a comparison with runoff series, Advances in Science and Research, 15, 173-181. doi:10.5194/asr-15-173-2018.

- Daly, C., Neilson, R. P., and Phillips, D. L. (1994). A statistical-topographic model for mapping climatological precipitation over Mountainous Terrain. Journal of Applied Meteorology, 33, 140-158. doi:10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2

- Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P. (2002). A knowledge based approach to the statistical mapping of climate. Climate Research, 22, 99–113. doi:10.3354/cr02209

- Eccel, E., Cau, P., and Ranzi, R. (2012). Data reconstruction and homogenization for reducing uncertainties in high-resolution climate analysis in Alpine regions. Theoretical and Applied Climatology, 110, 345-358. doi:10.1007/s00704-012-0624-z

- Feidas, H., Noulopoulou, C., Makrogiannis, T., and Bora-Senta, E. (2007). Trend analysis of precipitation time series in Greece and their relationship with circulation using surface and satellite data: 1955–2001. Theoretical and Applied Climatology, 87, 155–177. doi:10.1007/s00704-006-0200-5

- González-Hidalgo, J. C., Lopez-Bustins, J., Štepánek, P., Martin-Vide, J., and de Luis, M. (2009). Monthly precipitation trends on the Mediterranean fringe of the Iberian Peninsula during the second-half of the twentieth century (1951–2000). International Journal of Climatology, 29, 1415–1429. doi:10.1002/joc.1780

- IPCC (2013): Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press.

- Jones, P., D., and Hulme, M. (1996). Calculating regional climatic time series for temperature and precipitation: methods and illustrations. International Journal of Climatology, 16, 361-377. doi:10.1002/(SICI)1097-0088(199604)16:4<361::AID-JOC53>3.0.CO;2-F

- Lind, P., Lindstedt, D., Kjellström, E., and Jones, C. (2016). Spatial and temporal characteristics of summer precipitation over central europe in a suite of high-resolution climate models, Journal of Climate, 29(10), 3501–3518. doi:10.1175/JCLI-D-15-0463.1

- Lussana, C., Saloranta, T., Skaugen, T., Magnusson, J., Tveito, O. E., and Andersen, J. (2018). seNorge2 daily precipitation, an observational gridded dataset over

Norway from 1957 to the present day. Earth System Science Data, 10, 235–249. doi:10.5194/essd-10-235-2018

- Ranzi, R., Caronna P., and Tomirotti, M. (2017). Impact of climatic and land use changes on riverflows in the Southern Alps, in: "Sustainable Water Resources Planning and Management Under Climate Change", E. Kolokytha, S. Oishi, R.S.V. Teegavarapu (Eds.), Springer Science+Business Media, Singapore, 61-83. doi:10.1007/978-981-10-2051-3

- Sevruk, B., Ondrás, M., and Chvíla, B. (2009). The WMO precipitation measurement intercomparisons. Atmospheric Research, 92. 376-380. doi:10.1016/j.atmosres.2009.01.016