**Advances in
Science and
Research**

1961–1990 high-resolution Northern and Central Italy monthly precipitation climatologies

M. Brunetti[1], G. Lentini[2], M. Maugeri[2], T. Nanni[1], C. Simolo[1], and J. Spinoni[2]

[1]Istituto per le Scienze dell'Atmosfera e del Clima, via Gobetti 101, 40129 Bologna, Italy
[2]Dipartimento di Fisica, via Celoria 16, 20133 Milano, Italy

Abstract. Monthly Northern and Central Italy precipitation climatologies, obtained by means of a procedure aiming at capturing the local dependence of precipitation on elevation, are presented. The procedure consists in considering each cell of a high-resolution grid and in making use of the closest 15 stations in order to perform a weighted linear regression of precipitation versus elevation: the weights are determined by the distances of the stations from the grid point and by the level of similarity between the stations' cells and the considered grid cell in terms of elevation, slope steepness, slope orientation and distance from the sea. The study is based on a dataset of about 3200 1961–1990 monthly precipitation normals. The performance of the procedure is evaluated by reconstructing the monthly precipitation normals of each station and by comparing them with the corresponding observed values. Such comparisons are performed by means of five error estimators: the mean absolute error and the mean error, both evaluated in absolute and relative form, and the root mean squared error. The mean absolute error is slightly more than 10 mm for the monthly precipitation, which turns out to be between 10% and 14% of the monthly precipitation.

1 Introduction

High-resolution datasets of monthly precipitation climatological normals (i.e. high-resolution precipitation climatologies) have proved increasingly important in the recent past, and they are likely to become even more important in the future: they are requested by a variety of models and decision support tools, such as, just to cite a few, those used in agriculture, engineering, hydrology, ecology and natural resource conservation (Daly et al., 2002, 2006). These datasets have to be set up with the aim of providing a realistic representation of the major forcing factors that affect the spatial patterns of precipitation, in order to give reasonable estimates also for areas with poor station coverage such as high elevation sites in mountain areas (Daly, 2008).

In spite of their importance, in the last decades only little effort was put forward in order to obtain precipitation climatologies for the whole of Italy and, up to now, the best climatologies for the entire national territory are still the ones constructed in the 1950s by the Italian Hydrographical Service. They were obtained by means of the 1921–1950

normals of about 3500 stations (Servizio Idrografico, 1957) and they consist of maps with hand-drawn isohyets (Frosini, 1961) which cover the whole Italian territory: neither the maps nor the data which were used to construct them are available in digital form. It is also worth noticing that even the areas of Northern and Central Italy encompassed by the 1.25-arc-minute-resolution Alpine precipitation climatology presented by Schwarb (2000) have not been studied satisfactorily. In fact, the Schwarb climatology is mainly based on the well known database set up by Frei and Schär (1998): Italy took part in the construction of this database with rather dis-homogenous regional contributions, which resulted in a very poor station density over a significant fraction of the considered Italian area (see Fig. 3 in Frei and Schär, 1998). Such a problem concerns in particular the Po river basin and the areas located at the South of the 44th parallel.

We have recently set up a research programme with the aim of developing new 1961–1990 Italian high-resolution precipitation climatologies. The first assumption we adopt is that precipitation climatologies are linked to the physiographical features of the Earth's surface. Such an assumption allows the integration of the information contained in the meteorological records with the one arising from digital elevation models (DEMs): this permits the estimation of the

precipitation normals for a number of points which can be of several orders of magnitude larger than the number of the available records. The second assumption we adopt is that the link between precipitation normals and physiographical variables has to be studied at a spatial scale which allows to obtain a good compromise between the need of considering a small area and the need of working with a reasonable number of stations. The third assumption we adopt is that the leading physiographical variable is elevation: more in detail, we assume that, after considering a local scale and fixing all the other physiographical variables, the relation between precipitation and elevation can be approximated as linear. In other terms, we use the PRISM (Parameter-elevation Regressions on Independent Slopes Models) interpolation scheme: for a more detailed discussion of the PRISM methodology see Daly et al. (1994, 2002, 2006, 2008). The PRISM approach has also been applied to the construction of the Scwarb (2000) 1.25-arc-minute-resolution Alpine precipitation climatology. Schwarb performed also a comparison of PRISM with other interpolation methods: he found that PRISM proved effective especially in regions which are difficult to model because the influence of topography on precipitation was represented in the station data only partially. His results confirm that the PRISM approach proves to be particularly suitable for the construction of high-resolution precipitation climatologies of areas with complex orography such as Italy.

The first step of our research programme concerned Northern and Central Italy. The aim of this paper is to present the climatologies obtained so far: they are constructed on a grid based on a digital elevation model (DEM) with a 30-arc-second longitude and latitude resolution. These climatologies represent the first step of a wider project aiming at i) obtaining 1961–1990 precipitation climatologies for the whole of Italy and ii) allowing the estimation of local secular absolute value precipitation records for any point of the national territory. The methodology we adopt in order to estimate the local absolute value precipitation records is conceptually analogous to the one we present in Brunetti et al. (2009) for the temperature records. It makes use of both high-resolution climatologies and of a low resolution network of secular series (for a detailed discussion on the availability and quality of the Italian secular precipitation records, see Brunetti et al. (2006)).

One of the most interesting applications of the availability of high-resolution climatologies lies in their role in completing the information which can be extracted from a sparse secular precipitation network (i.e. the spatio-temporal behaviour of the precipitation multiplicative anomalies) and in permitting the estimation of local secular absolute value records for any point of the national territory. The ability of estimating such records turns out to be particularly useful in the frame of climate change impact-related projects. The estimated local records not only permit the best assessment of local past precipitation variability and change, but they also allow the

General Circulation Models (GCM) information to be tailored to the need of assessing local impacts of future climate change, via downscaling procedures.

2 Data

The 1961–1990 Northern and Central Italy monthly precipitation climatologies presented in this paper are based on a dataset that was constructed with the main goal to get the largest possible number of Italian records covering at least a 20-year period.

The principal data source was the former Italian Hydrographical Service (this national service is no longer operational and its competences are now attributed to the Italian administrative regions). Most of the records were provided by ISPRA (Institute for Environmental Protection and Research, formerly known as APAT, or Agenzia per la Protezione dell'Ambiente e per i Servizi Tecnici), even though in some cases the data were provided directly by local offices of the former Hydrographical Service, such as the Tuscan one based in Pisa. The records provided by ISPRA generally start in 1951 and are updated to the 1980s or 1990s, depending on the region. Beside the considerably large amount of data coming from the former Hydrographical Service, a wide range of other sources was exploited. Among these, the most relevant ones were the datasets of UCEA (Ufficio Centrale di Ecologia Agraria), of the Italian Air Force, of some ARPA (Agenzia Regionale per la Protezione dell'Ambiente) agencies and the one collected on Alpine catchments by the former Italian Electric National Board (ENEL). Moreover, some regional or sub-regional datasets set up by local authorities were considered as well, such as the one of Meteotrentino for the Trento autonomous province and the one of the Autorità di Bacino del Reno for the area of the Reno river basin, in Emilia-Romagna.

It is worth noticing that the availability of Italian digital records has rapidly expanded over the last few years, thanks to several digitization programmes recently completed or still in progress. These projects were often triggered by the stable scenario in terms of data holders that was finally achieved roughly ten years after the closure of the Hydrographical Service. The new data availability made it possible to expand the 1951–1986 time span of the former Hydrographical Service records. The 1951–1986 period was also the one considered in the Frei and Schär (1998) database for the Italian data.

The dataset was completed with some Slovenian series, kindly provided by the Environmental Agency of the Republic of Slovenia (ARSO), and with the 1961–1990 monthly precipitation normals of a number of Austrian and Swiss stations available, respectively, at ZAMG (Zentral Anstalt für Meteorologie und Geodynamik) and Meteoswiss.

All records were subjected to a quality check procedure which consisted in checking all sites for their position (the consistency among position and elevation was the

main constraint) and in correcting the coordinates, when possible, or deleting the series any time the correct position could not be identified with a reasonable confidence. This checking procedure was performed by: i) comparing the elevation of each site with the elevation of the closest grid point of a 30-arc-second-resolution digital elevation model (USGS GTOPO30, http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.htm), ii) studying all the largest discrepancies by means of metadata and/or supporting mapping tools like Google-Earth. Moreover the monthly normal values of all stations were compared with the ones of the neighbouring sites, in order to check the data interconsistency. A further check concerned the stations which were available from more than one source: in this case only the most reliable version was included in the final version of the dataset. It was usually selected the one with the smaller amount of missing values.

This detailed check of the data, though extremely time-consuming, turned out to be very important and a number of errors were corrected as a result. In particular it was absolutely necessary to enhance the reliability of the information concerning the station elevations, as non-correct elevation values may induce significant errors in the estimation of the precipitation-elevation dependence, which is the chief relation on which our climatologies are constructed. Actually, the precipitation versus elevation regressions on which our climatologies are constructed are not performed on the basis of the station elevations themselves, but on the basis of the elevations of the corresponding DEM cells. It is therefore clear how a correct identification of the station location is of key importance for the correct identification of the elevation to be used in the regression model. Among possible further errors, it is necessary to mention the ones related to the DEM itself: for elevation, GTOPO30 has a root mean squared error (RMSE) of about 18 m (USGS, 1996). This 18-m error is fairly small, especially if compared to the errors due to the uncertainty in stations' locations which, for some areas, can be remarkably bigger, even after the quality checks.

Once the dataset was set up and checked, the 1961–1990 monthly precipitation normals were calculated for all the available stations. Whenever the 1961–1990 period was not completely available, we considered the closest station without missing data in 1961–1990 and with an overlapping period of at least 10 years with the incomplete series, in order to re-scale the normals of the incomplete series to the 1961–1990 period. The re-scaling procedure was based, as in Brunetti et al. (2006), on the assumption that the ratios among precipitation amounts in neighbouring stations are constant in time. It was not always possible to re-scale to the 1961–1990 normals and some stations were rejected. The final dataset used for the Northern and Central Italy precipitation climatologies consists of about 3200 stations.

3 Methods

The spatial distribution of Italian precipitation is strongly linked to the complex orography of the Italian territory. Italy includes both the Alps and the Apennines and presents the highest peaks in Europe, west of the Caucasus: the existing precipitation climatologies generally show an increase of precipitation with elevation, even though such a relation is extremely site-dependent (Frosini, 1961).

In this context we set up a procedure aiming at the estimation of a "local" precipitation-elevation relationship for each of the Italian grid cells of the USGS GTOPO30 digital elevation model (DEM). For a more detailed discussion on models, as the PRISMs – Parameter-elevation Regressions on Independent Slopes Models – which aim at parametrizing the dependence of a meteorological variable on elevation, see the PRISM Group website http://www.prism.oregonstate.edu/; in particular, for precipitation, see Daly et al. (1994, 2002, 2006).

Actually instead of using directly the USGS GTOPO30 DEM we preferred to smooth it by means of the following relation:

$$h(x,y) = \frac{\sum\limits_{i=-k}^{k} \sum\limits_{j=-k}^{k} \frac{h'(x+i,y+j)}{(1+\max(|i|,|j|))}}{\sum\limits_{i=-k}^{k} \sum\limits_{j=-k}^{k} \frac{1}{(1+\max(|i|,|j|))}}$$

Where $h(x,y)$ is the elevation we attribute to the grid cell (x,y), $h'(x,y)$ is the corresponding elevation according to the USGS GTOPO30 DEM and k, which in our case is set to 2, is the range of the smoothing, expressed in number of contiguous grid cells from the central one. In this smoothing procedure the central grid cell has the maximum weight, while the weights decrease with increasing distance. The smoothing was introduced in order to avoid an unnecessarily high spatial resolution, while considering a spatial resolution closer to the actual scale at which the interactions of atmospheric circulation with orography occur.

The procedure we adopt for the estimation of the precipitation of any grid cell (x,y) of this smoothed version of the USGS GTOPO30 DEM (SmGTOPO30) is organized in three steps:

1. Slope steepness, slope orientation, and distance from the sea are calculated for each grid point of the SmGTOPO30.

2. Each station is associated to the closest grid point of the SmGTOPO30.

3. A local precipitation-elevation regression is performed for each grid point of the SmGTOPO30 and the estimated precipitation is assigned to the corresponding grid cell.

The precipitation-elevation regression is performed considering the 15 closest stations to the grid point and performing a weighted linear regression of precipitation versus elevation. So, the precipitation at the grid cell location (x, y) is estimated by substituting the grid cell elevation $h(x, y)$ in the equation:

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

where a and b are the coefficients of the weighted linear regression of precipitation versus elevation.

In order to avoid the overestimation of the precipitation normals at very high elevations, where almost no station is located, $h(x, y)$ is allowed to range only up to 2500 m, whereas it is set to 2500 m for all the grid cells with higher elevation. The behaviour of high-elevation precipitation in Italy will be furtherly analysed in the future, mainly thanks to the collection and digitization of data from nivo-pluviometric totalizers, available on the Hydrographical Service yearbooks.

In a procedure like this, the definition of the regression weights plays a key role. In our procedure, the weight of the i-th station involved in the linear regression yielding to the estmation of the precipitation at the grid cell (x, y) is the product of the following weighting factors:

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

All the weighting factors (position (rad), height (h), slope steepness (st), slope orientation (facet), distance from the sea (sea)) range from 0 to 1 and are based on Gaussian functions of the form:

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

where par is the specific geographical parameter which is being considered, c^{par} is a coefficient which regulates the decrease of the weighting factor and $\Delta_i^{\mathrm{par}}(x, y)$ is the difference between the values of the considered parameter at the grid cells corresponding to station i and grid point (x, y) (for rad and h such differences are rescaled in order to set to 1 the weighting factor of the closest station to the grid point, where the term "close" has to be intended, respectively, in terms of distance and elevation).

In our procedure the c^{par} values have been selected in order to reduce the weighting factors to 0.5 for: 25 km of radial distance, 250 m/km of difference in slope steepness, π of difference in slope orientation and 10 km of difference of distance from the sea. For the elevation c^{par} has been set equal to $(h(x, y) + 100\,\mathrm{m})^2 / \ln(0.5)$ in order to have a stronger decrease close to the foothills and a weaker decrease at higher elevations.

A minimum of 15 stations is required in order to evaluate the precipitation-elevation regression for each grid cell. The procedure includes an algorithm which searches for stations within a distance of 10 km from the grid point and, if

less than 15 stations are found, it increases such distance by steps of 5 km until the required minimum number of stations is reached up to a maximum distance of 50 km. If 15 stations are not found within 50 km the grid cell precipitation is not evaluated: this has never been the case for Northern and Central Italy.

To improve the stability of the precipitation estimation, the regression parameters corresponding to each grid point (x, y) are evaluated also by excluding one of the considered stations at a time. Then the station whose elimination from the regression causes the greatest variation in $b(x, y)$ is excluded from the regression if such a variation (in absolute value) is greater than 0.1 mm/m.

4 Results and conclusions

The January and July Northern and Central Italy precipitation climatologies are shown in Fig. 1a and b whereas Fig. 2 shows the annual precipitation climatology.

The performance of the model was evaluated by reconstructing the monthly 1961–1990 precipitation normals of each station and by comparing the results with the observed values. This was done by using a jackknife-like procedure, that involves the removal of each station before its reconstruction is carried out. This avoids "self-influence" of the station data that are being estimated.

To quantify the accuracy of the model five error-estimators were used: the mean absolute error (MAE) and the mean error (BIAS), both evaluated as absolute and relative errors (MAER and BIASR respectively), and the root mean squared error (RMSE).

The results are shown in Table 1 for all months. The table shows that the procedure is affected by a mean monthly precipitation absolute error of slightly more than 10 mm, which turns out to be between 10% and 14% of the monthly precipitation. It is also interesting to notice that the BIAS is very close to zero, indicating that no systematic overestimation or underestimation of precipitation is observed. The best performance of the model concerns the summer months, which turn also out to be the ones that have the strongest precipitation-elevation relations.

Figure 3 shows the scatter-plots of the reconstructed versus observed 1961–1990 precipitation normals for January, July and the year. This figure confirms the good performance of the procedure, which is also highlighted by the fact that the model explains between 89% (in February and March) and 96% (in July) of the variances of the stations precipitation normals (see Table 1).

The ability in estimating local secular anomaly records for any point of Northern and Central Italy (Brunetti et al., 2006) and the availability of high resolution climatologies enable us to estimate secular absolute value precipitation records for any point of this area. Such estimated local secular records can simply be obtained by a method conceptually analogous
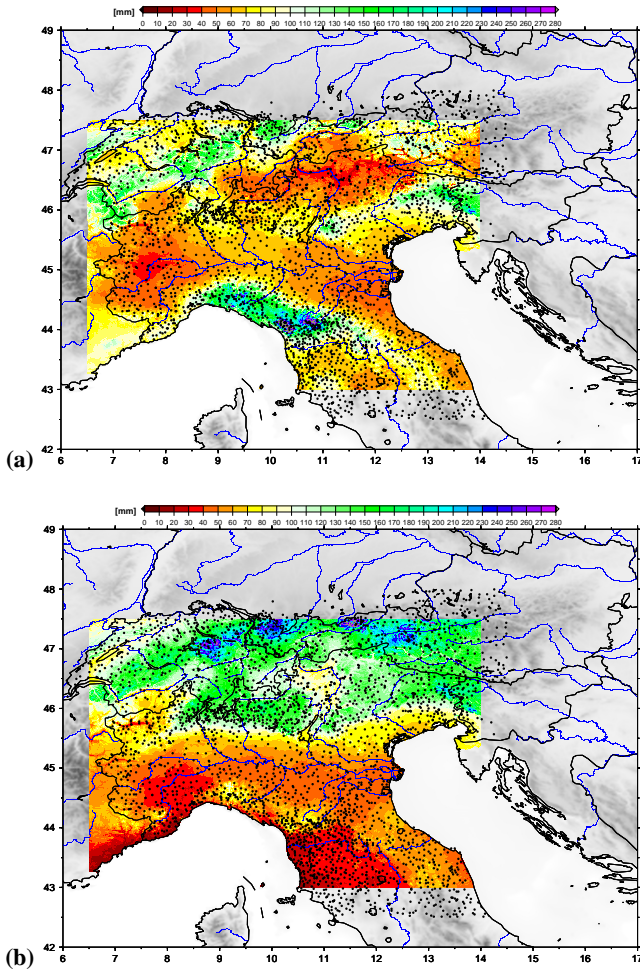
Figure 1. (a) January and **(b)** July 1961–1990 Northern and Central Italy precipitation climatologies. The figure also shows the stations included in the dataset of the 1961–1990 precipitation normals. Maps have been produced thanks to Honolulu University Generic Mapping Tools (GMT).
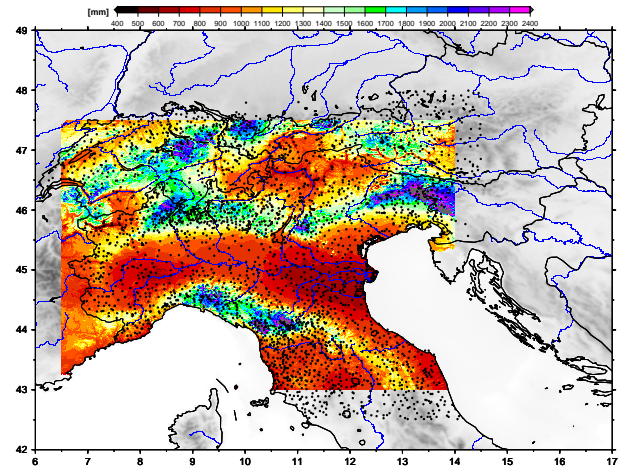

Figure 2. Yearly 1961–1990 Northern and Central Italy precipitation climatologies. The figure also shows the stations included in the dataset of the 1961–1990 precipitation normals. Maps have been produced thanks to Honolulu University Generic Mapping Tools (GMT).

Table 1. Accuracy of the climatologies quantified by means of: mean absolute error (MAE), mean error (BIAS), mean absolute relative error (MAER), mean relative error (BIASR) and root mean squared error (RMSE). The table also shows the fraction of the monthly variance (R^2) of the stations precipitation normals explained by the model.

	MAE (mm)	BIAS (mm)	MAER [%]	BIASR [%]	RMSE (mm)	R^2
1	11.1	0.2	14.3	3.9	16.8	0.91
2	10.2	0.1	13.9	3.6	15.0	0.89
3	11.2	0.0	12.9	3.1	16.4	0.89
4	12.6	0.0	12.5	2.9	18.4	0.91
5	12.0	−0.2	10.6	2.0	18.0	0.92
6	10.8	−0.2	10.4	2.0	15.9	0.95
7	9.6	−0.2	11.9	2.4	14.5	0.96
8	10.9	−0.2	10.6	1.9	15.7	0.94
9	10.2	0.1	10.5	2.1	15.1	0.91
10	12.2	0.3	10.9	2.4	17.6	0.93
11	13.6	0.1	11.6	2.6	20.1	0.92
12	11.0	0.1	13.2	3.2	16.6	0.91
AVG	11.3	0.0	11.9	2.7	16.7	0.92

to the one we present in Brunetti et al. (2009): for precipitation, it consists in multiplying the local estimated monthly secular multiplicative anomaly records for the corresponding 1961–1990 monthly precipitation normals.

The methodology for estimating local secular absolute value precipitation records has already proved useful for Northern and Central Italy in the frame of several climate change impact-related projects, such as the CARIPANDA (Lombardy Regional Project coordinated by Parco dell'Adamello, http://www.parcoadamello.it/progetti/default.htm) and the Kyoto Lombardia Projects (Lombardy Regional Project coordinated by Fondazione Lombardia per l'Ambiente, http://www.flanet.org/ricerca/kyoto.asp). The estimated records were also used in order to downscale to the local level future records produced by the General Circulation Models (GCMs); thus, the estimated local records not only permit the best assessment of local past precipitation variability and change, but they also allow the GCM information to be tailored for the need of assessing local impacts of future climate change.
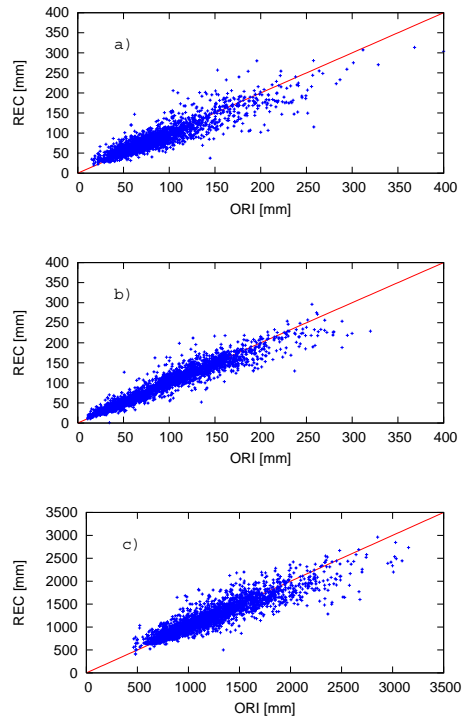
Figure 3. Scatter-plots of the reconstructed (REC) versus observed (ORI) 1961–1990 precipitation normals for January **(a)**, July **(b)** and the year **(c)**.

Naturally, the estimated local records must be considered with a clear awareness of the actual spatial resolution achievable with the available data. Such a spatial resolution ought still to be improved both for the secular anomaly records and for the 1961–1990 climatologies.

References

Brunetti, M., Maugeri, M., Monti, 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, 2006.

Brunetti, M., Lentini, G., Maugeri, M., Nanni, T., Simolo, C., and Spinoni, J.: Estimating local records for Northern and Central Italy from a sparse secular temperature network and from 1961–1990 climatologies, Adv. Sci. Res., 3, 63–71, 2009.

Daly, C., Neilson, R. P., and Phillips, D. L.: A statistical–topographic model for mapping climatological precipitation over mountainous terrain, J. Appl. Meteorol., 33, 140–158, 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.

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

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

Frei, C. and Schär, C.: A precipitation Climatology of the Alps from high-resolution rain-gauge observations, Int. J. Climatol., 18, 873–900 (also available at the website http://www.map.meteoswiss.ch/sop-doc/rr_sop/FreiSchaer1998.pdf), 1998 .

Frosini, P.: La carta della precipitazione media annua in Italia per il trentennio 1921–1950, in: Precipitazioni medie mensili ed annue e numero di giorni piovosi per il trentennio 1921–1950, Pubblicazione n. 24, Istituto Poligrafico dello Stato, Servizio Idrografico – Ministero Lavori Pubblici, Fascicolo 13, 1961 (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).

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