DELIVERABLE 1.1E. Temperature and precipitation high-resolution (30 arc-seconds) 1961-1990 climatologies for the italian alpine area (h>1500m a.s.l.)

The 1961-1990 30-arc-second resolution temperature and precipitation climatologies for the Italian Alpine region are presented together with the 1951-2012 monthly temperature and precipitation records (at the same spatial resolution) for three Italian National Parks located in the Alpine region (Gran Paradiso, Stelvio and Paneveggio – Pale di San Martino).

Monthly temperature and precipitation high-resolution climatologies

Stations' database description

The dataset used to produce the 1961-1990 temperature and precipitation climatologies over the Italian Alpine area are recovered from many different providers at international, national, regional and local level for the Italian territory and the surrounding areas. The dataset is mainly based on the archive of the former Italian Hydrographic Service (Servizio Idrografico, SI), updated with the new regional networks of automatic stations provided by Regional Environmental Agencies, integrated with data provided by SMI, single Observatories, and by foreign weather services of MeteoSwiss, MeteoFrance, ZAMG, and DWD. For more details see Brunetti et al. (2014) and Crespi et al. (2018) for temperature and precipitation, respectively.

All the records are subjected to quality controls by checking all sites for their position and correcting the coordinates when possible, or discarding the series if the correct location could not be identified. This step has been necessary, as incorrect elevations may induce significant errors in the estimation of the temperature/precipitation-elevation dependence. A further control on the station coordinates is also performed checking for each pair of stations if the highest correlation is associated to the closest station.

The precipitation series are further checked by comparison with simulated series estimated from neighbouring stations. Specifically, each monthly datum in each station (test station) is estimated by means of the ten closest stations (reference stations) with an available value in correspondence with the entry under consideration (p_{ref_i}) and with a sufficient number of data (15 monthly values) for that month in common with the test series. Then, the test series monthly datum under consideration (p_{test}) is calculated as median of the single values ($\tilde{p}_{test,i}$) estimated from the ten reference series using the anomaly method:

$$\tilde{p}_{test,i} = p_{ref,i} \cdot \frac{\bar{p}_{test,i}}{\bar{p}_{ref,i}} \qquad (i = 1, \ldots, 10) \qquad (1)$$

where $\bar{p}_{test,i}$and $\bar{p}_{ref,i}$ are the mean of test and reference series for the considered month over their common period.

After the quality-check, the 1961-1990 monthly temperature and precipitation normals are computed at each station site. To overcome the problem of missing data in the 1961-1990 period, the temperature normals are firstly calculated with the available data, and then re-adjusted to the 1961-1990 period by means of the Italian temperature anomaly records presented by Brunetti et al. (2006). Specifically, for each series (test series) with the 1961-1990 period partly (or fully) missing, a complete monthly local temperature anomaly (relative to the 1961-1990 period) record is calculated by a weighted average of neighbouring stations using the data and the interpolation method discussed in Brunetti et al. (2006). Then, monthly temperature normals over the same period available in the test station are calculated from this reconstructed series and subtracted to the test station normals to adjust them to the 1961-1990 interval. Moreover, all station normals are compared with those of neighbouring sites to highlight the largest discrepancies and to identify and correct possible errors. Differently, for precipitation the missing values in each station over the 1961-1990 period are estimated with the same method used for the quality-check.

For both variables the calculated normals are integrated with a small fraction of series for which the temperature and/or precipitation normals are already available (where referred to a different period they are rescaled to the 1961-1990 one).

Finally, the number of stations included in the databases and used to calculate the 30-arc-second monthly climatologies for the Italian Alpine region is 811 for temperature (Figure 1) and 2526 for precipitation (Figure 2).

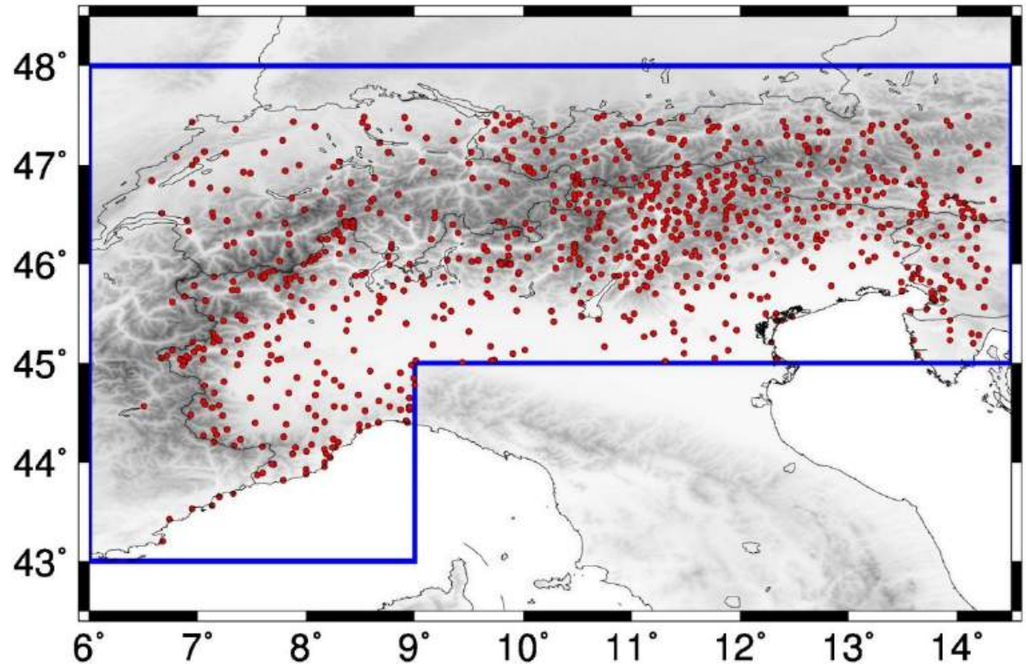

Figure 1. Spatial distribution of the stations included in the final database for temperature climatology reconstruction. The figure also shows the orography of the region.

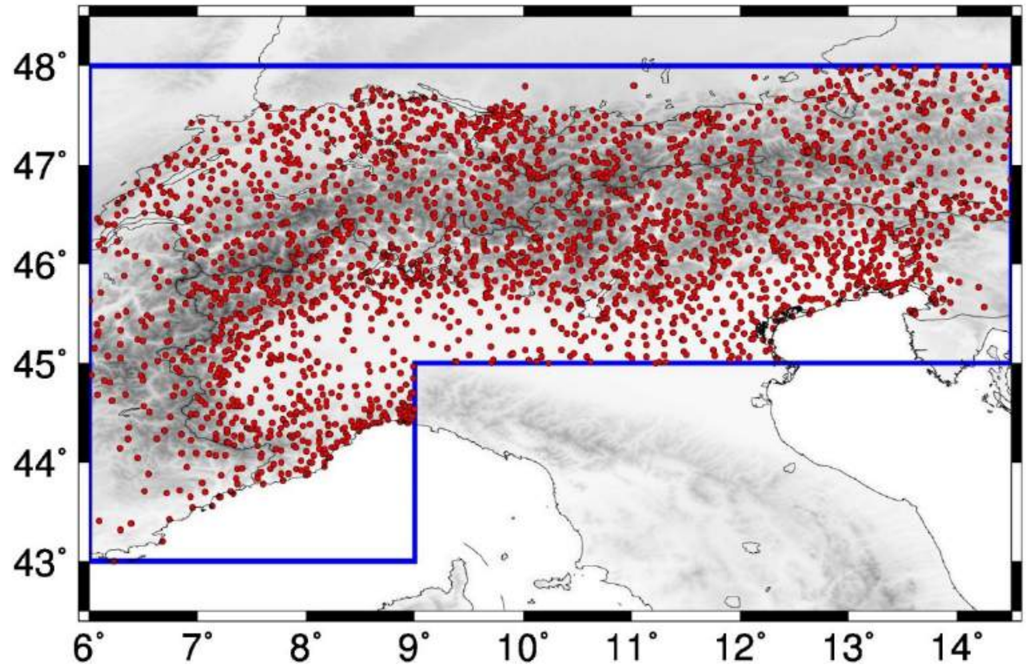

Figure 2. Spatial distribution of the stations included in the final database for precipitation climatology reconstruction. The figure also shows the orography of the region.

Interpolation method for climatologies

Temperature and precipitation climatologies are computed on the 30-arc-second-resolution (about 800m) GTOPO30 digital elevation model (DEM) (USGS, 1996) by means of a local weighted linear regression (LWLR) of temperature/precipitation *versus* elevation. This method is considered one of the most suitable approaches for areas with complex orography like the Alpine region.

The LWLR evaluates the relationship between the meteorological variable and elevation at a local level. In particular, temperature/precipitation normal X at the grid-cell (λ, ϕ) is obtained as a function of its elevation h by the following relation:

$$X(\lambda, \phi) = a(\lambda, \phi) + b(\lambda, \phi) \cdot h(\lambda, \phi) \qquad (2)$$

where $a(\lambda, \phi)$ and $b(\lambda, \phi)$ are the local regression coefficients.

Specifically, for any grid-cell only a cluster of stations is considered and greater weights are given to the nearest stations with topographic similarity to the target grid-cell. The stations selected for the regression (35 for temperature and 15 for precipitation) are those with the highest weights within a radius of 200 km from the given cell. More specifically, the weight of the i^{th} station involved in the linear regression yielding the estimation of the climatology of the point (λ, ϕ) is the product of weighting Gaussian functions depending on position, height, distance from the sea, slope steepness and slope orientation:

$$w_i(\lambda, \phi) = w_i^r(\lambda, \phi) \cdot w_i^h(\lambda, \phi) \cdot w_i^{dsea}(\lambda, \phi) \cdot w_i^{slope}(\lambda, \phi) \cdot w_i^{facet}(\lambda, \phi) \qquad (3)$$

where:

$$w_i^{var}(\lambda, \phi) = e^{-\left(\frac{\Delta_i^{var}(\lambda, \phi)^2}{c_{var}}\right)} \qquad (4)$$

where *var* is the geographical feature which is being considered, $\Delta_i^{var}(\lambda, \phi)$ the absolute difference between the values of this variable at the grid-point (λ, ϕ) and that at the i^{th} station location and c_{var} is a coefficient which regulates the decreasing rate of the weighting and it can be expressed in terms of the value of $\Delta_i^{var}(\lambda, \phi)$ which gives a weighting factor equal to 0.5 ($\Delta_{\frac{1}{2}}^{var}$). All the details on how $\Delta_{\frac{1}{2}}^{var}$ is defined and obtained for each geographical variable are reported in Brunetti et al. (2014) for temperature and Crespi et al. (2018) for precipitation. It is important to underline that for temperature, the most appropriate $\Delta_{\frac{1}{2}}^{var}$ factors are obtained iteratively, for each month and for each geographical feature by searching for the value that gives the lowest error at station locations. Differently, for precipitation, the most appropriate $\Delta_{\frac{1}{2}}^{var}$ are locally optimised over a grid of 1°x1° resolution for each month and then interpolated onto the finest resolution grid.

An advantage of LWLR scheme is that it is possible to define a prediction interval for each grid-cell estimation. The procedure consists in estimating the variance of the temperature/precipitation of a grid-point at elevation h_{new} as Daly et al. (2008):

$$s^2\{X_{h_{new}}\} = s^2\{\tilde{X}_{h_{new}}\} + MSE \qquad (5)$$

where $s^2\{\tilde{X}_{h_{new}}\}$ is the variation in the possible location of the expected temperature/precipitation for a given elevation and MSE is the mean square error of the observed station temperature/precipitation values compared to those obtained by the regression model (i.e. the variation of the individual station values about the regression line). The prediction interval (with confidence α) for the grid-point with elevation h_{new} is then defined as:

$$X_{h_{new}} \pm t_{\frac{1-\alpha}{2}, df} \cdot s\{X_{h_{new}}\} \qquad (6)$$

where t is the value of a Student distribution with *df* degrees of freedom (equal to the number of stations selected for the linear regression) corresponding to a cumulative probability (1-α)/2.

The climatologies

The performances of the LWLR for temperature and precipitation are evaluated in terms of the ability to reconstruct the 1961-1990 observed temperature/precipitation monthly normals at station sites. Specifically, the monthly normals of all the stations contained within the study domain (607 for temperature and 1177 for precipitation – area for which the climatologies are calculated – see Figure 3 for temperature and precipitation, respectively) are estimated and then compared to the observed values. The reconstruction is performed in each case by means of the leave-one-out approach, i.e. by removing the station whose normals are being estimated, in order to avoid "self-influence" of the station data to reconstruct. The results of the comparison between estimated and observed values are listed in Table 1 for temperature and Table 2 for precipitation, where the monthly accuracy is expressed in terms of mean error (BIAS), mean absolute error (MAE) and root mean square error (RMSE).

	1	2	3	4	5	6	7	8	9	10	11	12
BIAS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAE	1.0	0.8	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.8	1.0
RMSE	1.2	1.1	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.9	1.0	1.2

Table 1. Accuracy of the monthly temperature climatologies obtained from leave-one-out validation of the LWLR interpolation for the stations included in the study domain. All the values are expressed in °C. The errors are evaluated as the difference between simulated and observed values.

	1	2	3	4	5	6	7	8	9	10	11	12
BIAS	-0.1	0.2	0.3	0.7	0.8	0.4	0.1	0.2	0.1	0.5	0.5	0.1
MAE	7.9	8.1	9.7	12.1	12.7	10.8	8.9	9.9	9.8	11.1	11.6	7.1
RMSE	11.2	11.8	14.2	17.9	17.9	14.6	12.1	13.3	13.9	15.8	16.6	10.4

Table 2. Accuracy of the monthly precipitation climatologies obtained from leave-one-out validation of the LWLR interpolation for the stations included in the study domain. All the values are expressed in mm. The errors are evaluated as the difference between simulated and observed values.

The gridded monthly normals are shown for temperature and precipitation in Figure 3 and Figure 4, respectively.

Yearly temperature and precipitation climatologies for the whole Italian territory are available from the ISAC-CNR web site (http://www.isac.cnr.it/climstor/CLIMATE_DATA/).

1951-2012 monthly temperature and precipitation datasets for three National Parks

Interpolation methods are highly vulnerable to fluctuations in stations' spatial coverage. This vulnerability is even more evident in mountain areas where stations are located at different elevations and absolute temperature and precipitation values present strong spatial gradients. In fact, in presence of a gap in a high elevation station the interpolated value can be biased by interpolating between adjacent valley stations.

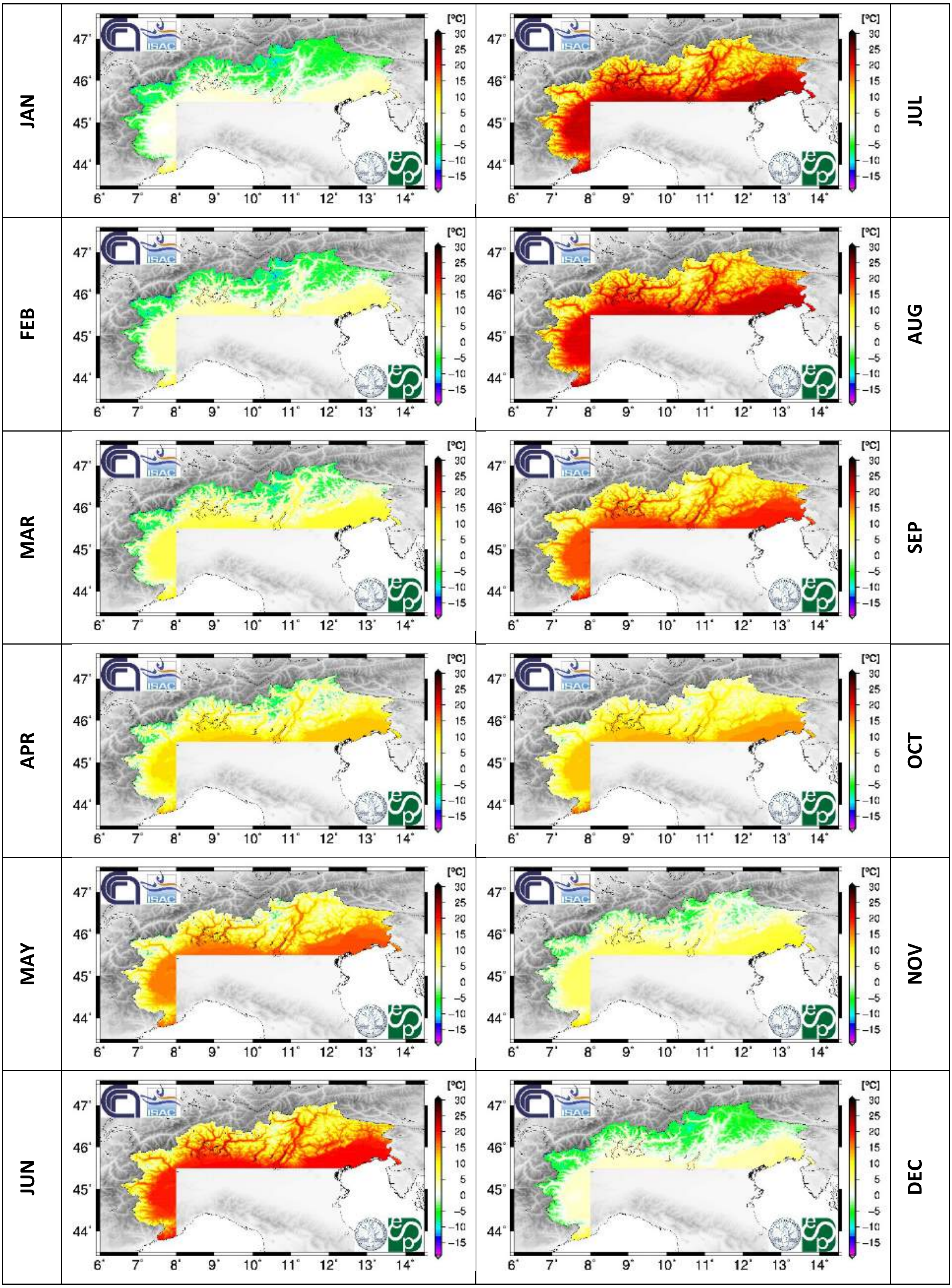

Figure 3. 1961-1990 monthly temperature climatology

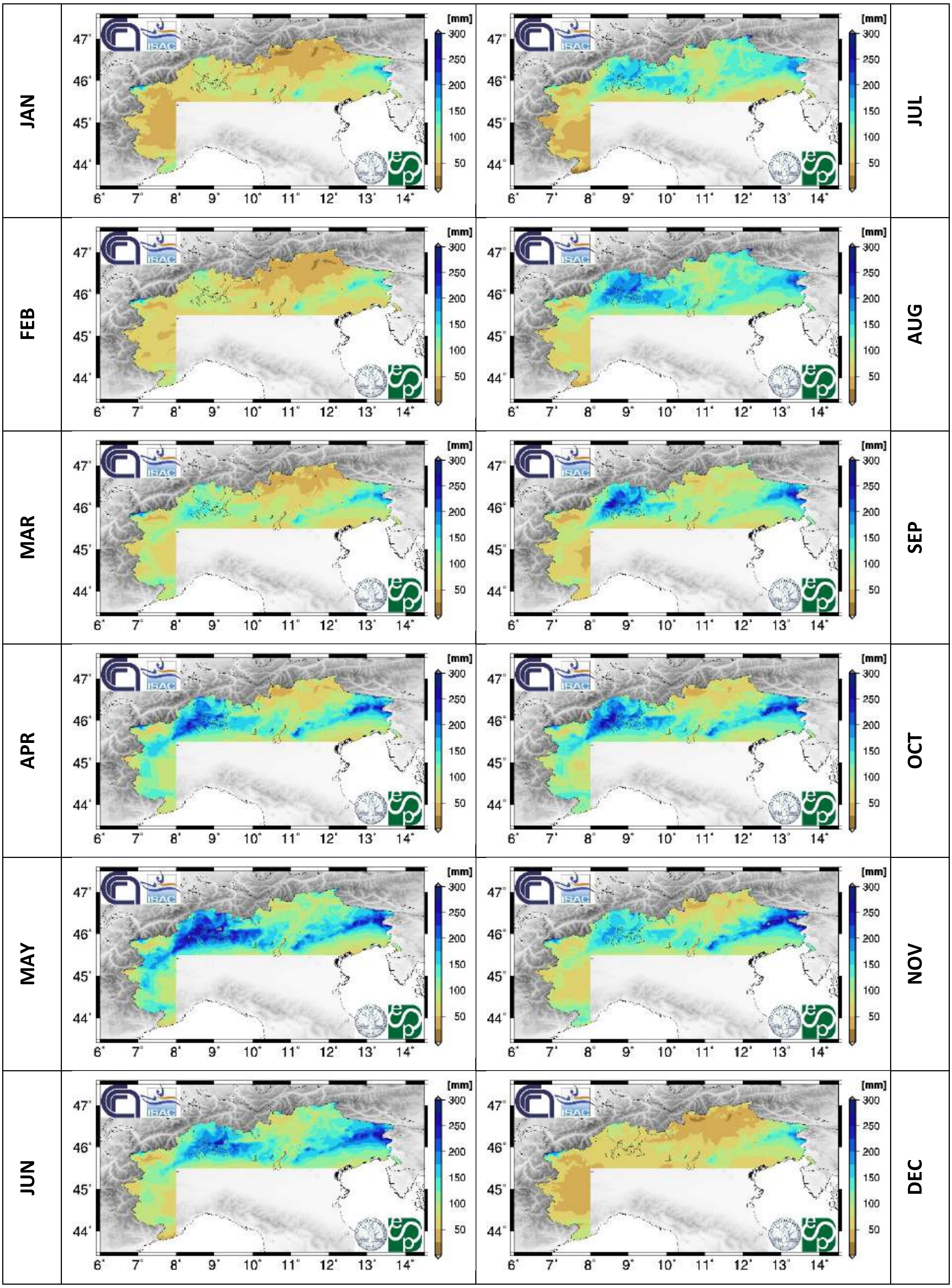

Figure 4. 1961-1990 monthly precipitation climatology

For this reason, the most widely adopted methodology to reconstruct gridded datasets is the anomaly method (New et al., 2000; Mitchell and Jones, 2005), based on the assumption that the spatio-temporal structure of the signal of a meteorological variable over a specific area can be described by the superimposition of two fields: the normals over a given reference period (i.e. the climatologies) and the departures from them (i.e. the anomalies). The formers are strongly linked to the geographical features of the territory and they can manifest remarkable spatial gradients. The most relevant aspect for their description is the availability of high-density observational dataset integrated by interpolation methods which describe the relationship between the meteorological variable and the physiographical characteristics of the Earth's surface (Daly et al., 2002, 2008; Daly, 2006). Differently, the latter are linked to climate variability and change and they are generally characterized by higher spatial coherence where the priority for their description lies in data quality and in the availability of long records.

Stations' database description

The databases considered for the interpolation of the 1951-2012 temperature and precipitation anomalies are independent from those used to calculate the monthly climatologies and they are collected mostly within and in the surroundings of the three boxes centred on the three parks. Specifically, they encompass 160 series for temperature (white symbols in Figure 5) and 732 series for precipitation (white symbols in Figure 6). Before performing the interpolation, all station records are subjected to detailed quality-checks and homogenization procedures in order to remove non-climatic signals (Brunetti et al., 2006). Then, the checked station series are converted into monthly anomalies (additive anomalies for temperature and multiplicative anomalies for precipitation) with respect to the 1961-1990 reference period (the same to which the above discussed climatologies are referred to).

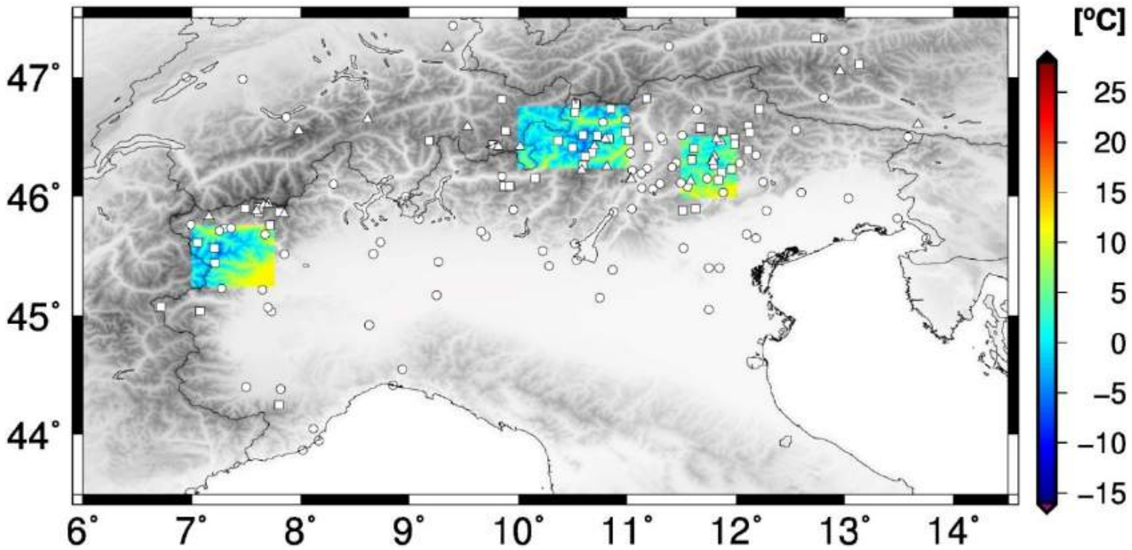

Figure 5. 30-arc-second-resolution temperature annual climatologiy for the 1961-1990 period for Gran Paradiso, Stelvio and Paneveggio - Pale di San Martino National Parks, together with the stations' network used for the 1951-2012 gridded temperature dataset, represented with dots for elevations lower than 1000m, with squares for elevations higher than or equal to 1000m and lower than 2000m and triangles for elevations higher than or equal to 2000m.

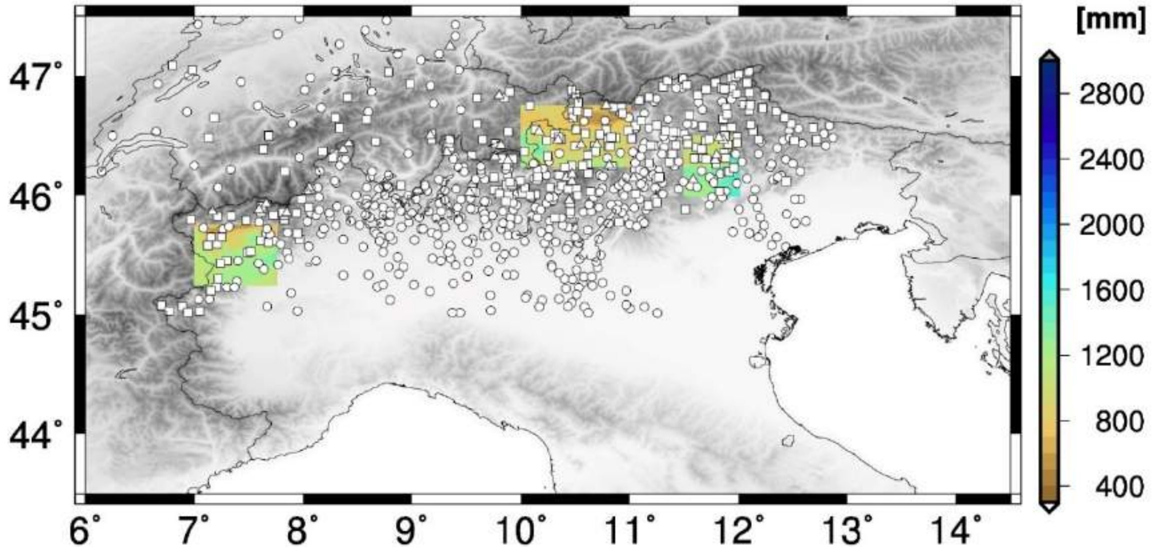

Figure 6. 30-arc-second-resolution annual precipitation climatologiy for the 1961-1990 period for Gran Paradiso, Stelvio and Paneveggio - Pale di San Martino National Parks, together with the stations' network used for the 1951-2012 gridded precipitation dataset, represented with dots for elevations lower than 1000m, with squares for elevations higher than or equal to 1000m and lower than 2000m and triangles for elevations higher than or equal to 2000m.

Interpolation method for anomalies and conversion into absolute values

In order to join the information arising from the climatologies with the one of the anomalies, temperature and precipitation anomaly records are firstly interpolated from station sites onto the same high-resolution grid for which climatologies have been computed. This step is based on an improved version of the method presented by Brunetti et al. (2006). Specifically, each monthly grid-point value is calculated if there is at least 1 series at a distance lower than 50km and 3 series at a distance lower than 100km. The 9 stations with the highest weights are retained and their weighted mean is calculated after discarding the minimum and maximum value. If there are less than 5 available series, the weighted mean is calculated without discarding the minimum and maximum value (but this does not happen for any grid-point). The weight of each station i is defined by means of the product of Gaussian functions (Equation 4) depending on the distance $w_i^{rad}(\lambda,\phi)$ and elevation difference $w_i^{ele}(\lambda,\phi)$ between the ith station and the considered grid-point, with an additional angular waiting factor taking into account the anisotropy in the spatial distribution of the stations about the grid-point and defined as follows:

$$w_i^{ang}(\lambda,\phi) = 1 + \frac{\sum_{l=1}^{n} w_l^{rad}(\lambda,\phi)\cdot w_l^{ele}(\lambda,\phi)\left[1-cos\theta_{(\lambda,\phi)}(i,l)\right]}{\sum_{l=1}^{n} w_l^{rad}(\lambda,\phi)\cdot w_l^{ele}(\lambda,\phi)} \qquad (7)$$

where $\theta_{(\lambda,\phi)}(i,l)$ is the angular separation of stations i and l with the vertex of the angle in the grid-point (λ,ϕ) and n is the number of stations considered to estimate the grid-point value (equal to 9 in our case). The weight halving distance (i.e., the distance for which the corresponding weight is equal to 0.5) is calculated as the mean distance from the grid-point of all stations within 100 km from it (this permits to regulate the weight decreasing rate taking into account the temporal variability in station availability), while the elevation weight halving elevation difference is set to 1000 m.

After obtaining the 1951-2012 monthly temperature and precipitation gridded anomalies, they are converted into absolute values by superimposing the corresponding 1961-1990 monthly normals. Specifically, for temperature the monthly normals are added to the monthly anomalies while for precipitation they are multiplied.

Bibliography

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

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. DOI: 10.1002/joc.3764.

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

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

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

Daly C., Halbleib M., Smith J.I., Gibson W.P., Doggett M.K., Taylor G.H., Curtis J., Pasteris P.A. (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.

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

New M., Hulme M., Jones P.D. (2000), *Representing twentieth century space-time variability*, Part 2: development of 1901–96 monthly grids of surface climate, J. Climate, 13, 2217–2238.

USGS (United States Geological Survey) (1996), *GTOPO30 Documentation*, Available at http://eros.usgs.gov/ #/Find_Data/Products_and_Data_Available/gtopo30_info.