Spatial distribution of the average air temperatures in Italy: a quantitative analysis

Pierluigi Claps[*1], Pietro Giordano [2] and Giovanni Laguardia [3]

[1] Dipartimento di Idraulica, Trasporti ed Infrastrutture Civili – Politecnico di Torino

C.so Duca degli Abruzzi 24, 10129. Torino, Italy. (claps@polito.it)

[2] Presidenza del Consiglio dei Ministri - Dipartimento della Protezione Civile, Previsione, Valutazione, Prevenzione e Mitigazione dei rischi naturali

Via Vitorchiano 2, 00189 Rome, Italy. (Pietro.Giordano@protezionecivile.it)

[3]European Commission – DG Joint Research Centre, Institute for Environment and Sustainability

Via E. Fermi 1, T.P.261, 21020 Ispra (VA), Italy. (giovanni.laguardia@jrc.it)

Short Title: *Spatial distribution of the average air temperatures in Italy*

[*] Correspondence to: Pierluigi Claps, Dipartimento di Idraulica, Trasporti ed Infrastrutture Civili-Politecnico di Torino. C.so Duca degli Abruzzi 24, 10129. Torino, Italy.

ph: +39 011 5645656; fax: +39 011 5645698; E-mail: claps@polito.it

ABSTRACT

Annual and monthly average temperatures are analyzed with statistical methods to characterize the temperature regime in Italy. Data from 738 weather stations, with homogenous spatial cover throughout Italy, are used to estimate the annual and monthly air temperature normals. Geographic and morphologic parameters, computed around the points of measure, are considered as explicative variables within a geo-regression model. Morphometric variables are defined using the USGS GTOPO30 digital elevation model, which has a resolution of 1 km. On the basis of a stepwise regression analysis, the significant variables are found to be elevation, latitude, distance from the sea, and a measure of terrain concavity. The relationship between the average annual air temperature and the mentioned variables explains 92% of the variance and produces a standard error of 0.89°C. The temperature regime (normalized mean temperature for each of the 12 months) is reproduced with a two-harmonic Fourier series, with parameters estimated using stepwise regression. Analyses of the reconstruction errors demonstrate that the results are quite satisfactory for many technical purposes, particularly for large-scale climatic characterization.

KEYWORDS: air temperature; spatial distribution; regression analysis; kriging; Fourier series; Italy.

INTRODUCTION

Large scale spatial variability of climatic variables, such as temperature or rainfall, has recently received considerable attention (*e.g.,* Zheng and Basher, 1996; Nalder and Wein, 1998; Prudhomme and Reed, 1999). Considering temperature, the basic variable to investigate is the *temperature normal*, defined as a "period average, computed for a uniform and relatively long period comprising at least three consecutive ten-year periods" (WMO, 1983). As regards the spatial distribution of air temperature, of particular interest are the studies concerning its relations with physiographic parameters (see *e.g.,* Zheng and Basher, 1996; Agnew and Palutikof, 2000; Ninyerola et al., 2000; Gyalistras, 2003). These studies start from the premise that elevation and latitude already explain much of the spatial variability of temperature but proceed to evaluate other factors that can have an additional significant influence on it, including the position of the site with respect to seas and continents and, on small scales, terrain attributes (aspect and morphology of the relief), atmospheric factors (humidity, precipitation, and wind), and maritime factors (configuration and aspect of coasts and effects of sea currents).

Within this framework a few authors attempted to reproduce the spatial variability of climatological variables with the combined use of regression analysis and Geographic Information Systems (GIS), intended as a technique used to quickly compute geographical and morphological parameters (*e.g.*, Prudhomme and Reed, 1999; Ninyerola et al., 2000). Compared to studies which address the maximum efficiency of statistical spatial interpolation techniques (*e.g.,* Nalder and Wein, 1998; Lapen and Hayhoe, 2003), the approaches that investigate the role of a given predictor on the spatial variability of a climatic variable can help to improve the characterization of other climatic variables. For example, the regionalization of flood peak indices can certainly benefit from the results presented by Prudhomme and Reed (1999) on the spatial analysis of extreme precipitation. Following this approach, the study presented here is devoted to the selection of meaningful spatial variables to predict average annual and monthly air temperatures in Italy. The areal extent of the region under study (about 300,000 km^2) makes this objective of interest for all those applications that require temperature estimates over medium-large areas and include, among others, assessment of monthly average evapo-transpiration, relations between rainfall, temperature and vegetation density, or heating and cooling energy requirements.

To date, quantitative analyses of the mean temperature variability in Italy are available only for limited areas. Gentilli (1959) and Paiero (1968) considered relatively small regions and used simple regression relationships between temperature and altitude. More recently, Claps and Sileo (2001) investigated the effects of elevation and latitude on the annual and monthly temperature normals in peninsular Southern Italy. The purpose of the present study is to assess the role of additional physiographic factors on the reconstruction of average temperature over the entire country using linear stepwise regression and kriging. Specific attention is devoted to the representation of the monthly temperature curve, or 'regime', introduced to combine the 12 relations usually considered between monthly means and geographic parameters (see *e.g.*, Zheng and Basher, 1996) into a single set of regression equations.

DEFINITION OF GEOGRAPHIC AND MORPHOMETRIC PARAMETERS

Italy extends for about 10°30’ of latitude, which corresponds to approximately 1,200 km. The coasts of Sicily are only a few hundred kilometers from Africa, while the Northern border is part of continental Europe. This implies a great diversity of environmental and geographic conditions that have a strong influence on the spatial distribution of temperature. Analyzing national temperature maps (e.g., Ministero dei Lavori Pubblici, 1969), the most evident characteristic observed is the common tendency for the isotherms to follow the contour lines of the mountains (the Alps and the Apennines).

The sea influence is also important, yet complex. Its buffering effect is not uniform along the coast of Italy. Along the Adriatic sea coast (refer to Figure 1) its effect is less evident compared to the coastal areas of the Tyrrhenian and the Ionian seas, which are deeper and more open to the influence of the lower Mediterranean atmospheric circulation. Moreover, for the Adriatic sea, the mitigating effect decreases in the northerly direction.

It is important to consider the influence exerted by the Alpine chain, which protects the Padana plain from cold northern winds. The Alps modify the thermal conditions of the cold air masses that come from the north, determining an increase in their temperature and producing föhn winds (e.g., American Meteorological Society, 1959). This effect can contribute to raise the mean air temperature values in some large alpine valleys, particularly in the spring months. In the same way, the Apennines chain generates climatic differences between the Tyrrhenian and Adriatic coastlines because of the prevailing west-to-east wind direction.

Some other factors exert a smaller influence on temperature or cause significant effects only in limited areas. A significant example is the presence of large lakes, such as the pre-Alpine lakes, which mitigate the temperature variations in the neighboring areas. Heat islands, relative to large cities, are the extreme case of local effects, because they produce an increase in temperatures in the areas surrounding the cities. This can increase the model prediction errors near the larger built-up areas. All of these local effects are not accounted for in this analysis.

Definition of physiographic parameters

Geographic and morphologic parameters that can influence the spatial distribution of mean temperatures are defined here following the criteria suggested by Prudhomme and Reed (1999) for the spatial mapping of extreme rainfall. The topographic and geographic factors selected with respect to the variability of the mean temperature are (see Figure 2):

- Minimum absolute distance from the sea, d_{min};
- Angle α formed by the minimum distance vector and the South (Figure 2a);

- Distance d_i from the sea in the eight cardinal directions, *i* (Figure 2a);
- Azimuthal angle β_i of the horizon in the eight cardinal directions (Figure 2b).

The latter variable plays the role of an *obstruction* factor, according to the original definition of Faulkner and Prudhomme (1998), and is defined as the angle subtended by the highest topographical barrier in the i^{th} direction, such that $tg\beta_i = \Delta H / \Delta X$, with ΔH as the difference of elevation between the barrier and the station, and ΔX as the distance between the two points. Note that d_{min} is not necessarily one of the eight distances d_i taken in the cardinal directions, because it represents the minimum distance between the station point and the coastline.

The above variables were computed for each of the considered stations (see Figure 1) using a geographical representation of the Italian coastal line and a Digital Terrain Model (DTM). The DTM used is a 1 km resolution model named GTOPO30 that is distributed by the U.S. Geological Survey (USGS, 2006). At present, DTMs with higher resolution are available for the Italian territory. Our decision to use the 1 km resolution DTM was based on a trade-off between advantages (better resolution for computing concavities and obstructions) and disadvantages (need for careful validation of the stations' position and aspect, substantial increase of the computational load) with respect to the objective of the analysis.

The previously described variables were subsequently transformed and averaged in different ways, producing three parameters: a distance measure *Ds*, an aspect variable *As*, and a concavity index *C*. Definitions for these parameters were the result of the evaluation of the relations between several combinations of the basic variables and the spatial patterns displayed by temperature data.

The distance measure *Ds* is the geometric mean of the distance from the sea in the eight cardinal directions:

$$Ds = \sqrt[8]{d_1 * d_2 * .... * d_8} \ . \qquad (1)$$

It discriminates among places that have the same minimum distance, but different average distances, from the sea. In directions for which the distance should be computed across the Alps, a fixed distance of 1,000 km was defined, which excluded consideration of the Northern European coastline. In practical terms, this parameter represents an index of continentality of the measurement site.

The second parameter is a combined measure of *aspect* (orientation) and sea proximity:

$$As = \frac{10 \cdot cos\alpha}{1 + d_{min}} \qquad (2)$$

The inverse of the minimum distance from the sea reduces the influence of the aspect of the sea proximity on inland stations. The reason for this definition is that the aspect characteristics were found to be significant only for near-coast stations, producing visible cold ‘anomalies’ in stations of the Adriatic Sea compared to stations located on the Tyrrhenian and the Ionian coastlines.

The third index is a *concavity* index, obtained by weighting the azimuthal angle β_i in the eight directions:

$$C = \sqrt[8]{\prod_{i=1}^{8} 10^{2\,tg\beta_i}} \;, \qquad (3)$$

Only a few authors have considered the effect of orographic barriers on average temperatures. Gentilli (1959) considered the shape of the terrain and noticed that, at the same elevation, prediction in areas with concave topography resulted in negative temperature anomalies (*i.e.*, cooler terrain) because of cold air stagnation in the concave areas. He defined a qualitative *topographic index* in an attempt to improve temperature-elevation regressions. Faulkner and Prudhomme (1998), and less explicitly Ninyerola et al. (2000), referred to ‘obstruction’ factors (with respect to wind directions) in their analyses, but no explicit consideration of a concavity effect on air temperature was found in the literature.

Temperature data and digital terrain model

Temperature data recorded in 738 stations distributed throughout Italy (Figure 1) were used in this study. Monthly and annual minimum and maximum temperature normals, computed as the average of the daily measurements, were published in a coordinated collection (Petrarca et al., 1999), which is the most comprehensive systematic temperature database available in Italy. In this database, about 85% of the stations have at least 30 years of observation, and the remaining 94 stations have at least 20 years of record. Monthly and annual normals were therefore obtained by averaging minimum and maximum values.

In datasets with uneven record lengths, some caution should be applied in the data preparation for regressions, as discussed for instance by Robeson and Janis (1998). We did not make corrections to the data based on the shorter series, considering that the small biases mentioned by Robeson and Janis (1998) would not significantly affect the statistical connections between temperature and the predictors. Moreover, even considering the temperature trends detected in Italy (e.g., Nanni et al., 1998, Brunetti et al., 2000), a 10-year discrepancy would not produce significant heterogeneity on the mean. However, quantitative evaluation of the possible effects of this decision will be presented in the discussion of results.

The spatial distribution of the stations is quite even, but the stations are not evenly distributed with elevation. Only 7% of the stations are located above 1200 m a.s.l, while 12% of the Italian territory lies above that elevation (see Table 1). Stations at elevations less than 100 m a.s.l. are 28% of the total, but the areas below that elevation cover 23% of the country surface.

ANALYSIS OF THE MEAN ANNUAL TEMPERATURE

The sample of annual normals (T_a) for the set of 738 stations was related to the geographic predictors described in the previous section, by means of a stepwise regression procedure. This procedure produces multiple regression models of increased complexity, according to the number of independent variables considered (Table 2). In addition to the coefficient of determination (R^2) and the test of coefficients significance (*Tstat*), Table 2 reports the percent of stations that fall in each error class (<1°C up to >3°C), so as to provide a further indicator of the improvement achieved with the introduction of each new variable.

Apart from elevation and latitude, the value of geographic information in the explanation of T_a is questionable. The role of the sea, represented by variables *As* and *Ds*, is significant, as indicated by *Tstat*, but their inclusion in the regression model does not produce significant increases in R^2 (see Table 2). Considering also the results of the cross-validation procedure (described below), the best model for annual normals remains based only on elevation (E) and latitude (L), as follows:

$$\hat{T}_a = 43.59 - 0.0054\,E - 0.6601\,L, \tag{4}$$

where elevation is expressed in m a.s.l., and latitude is in hexadecimal degrees. This model explains 92.4% of the variance and gives a standard error of estimate of 0.894°C. The estimated average thermal gradient is 1°C for 184 m of altitude and 0.66°C for each degree of latitude, which is in good agreement with values in the literature (e.g., Pinna, 1977).

Cross-validation of the annual temperature estimates

According to an approach commonly adopted in the evaluation of the spatial variability of geophysical variables, a cross-validation assessment was made to verify the performance of the regression model of Equation 4. Following Wilks (1995), the idea is to remove iteratively one station from the complete sample, and to re-evaluate the model without that station to have the possibility to compute the estimation error. Using these at-station errors, a global root mean square error (*RMSE*) is then computed, which allows one to rate the performance of the regression model. This technique resembles the Jackknifing procedure, except for the use of the results, being Jackknifing typically used to evaluate the variance of parameter estimates.

In our application, the *RMSE* was evaluated on the results of the five different regression models presented in Table 2, to evaluate the increase in efficiency produced by the introduction of each new explicative variable. It is evident from Figure 3 that the use of more than two variables (*E* and *L*) produces negligible decrease in *RMSE*, confirming that even the high values attained by *Tstat* for *Ds* and *As* cannot justify their inclusion in the regression model.

Spatial analysis of the annual temperature residuals

To verify the quality of the results of estimates obtainable with Equation 4 we tested the residuals of the regression model for normality, by plotting their Cumulative Frequency (CF) on normal probability paper, and also against the presence of unaccounted autocorrelation. Residuals' CF was fairly linear in the probability paper, as could be expected in consideration of the low sample skewness coefficient (which was -0.1). Therefore the assumption of normality was upheld. The autocorrelation function values were found within the 5% confidence bands, demonstrating also incorrelation in the residuals. However, observing the residuals on a map, some coherent spatial patterns of anomalies are still apparent. It was then decided to carry out additional analyses to account for the presence of spatial correlation, using a kriging approach (e.g., Cressie, 1993). The objective was to improve the prediction of annual normals from Equation 4 by adding a correction factor e_{reg}, obtained by kriging, for every point of the 1 km spatial grid covering Italy. Therefore, the prediction model of T_a assumes the form:

$$T_a = \hat{T}_a + e_{reg} \qquad (5)$$

where $\hat{T}_a$ is estimated through Equation 4 and e_{reg} is the residual component estimated by kriging.

The kriging procedure measures the spatial correlation in the errors of Equation 4 through the computation of the sample semi-variogram, which is the function $\gamma(h)$ that measures one half of the average squared difference of data values separated by the lag distance h, and the subsequent fitting of the coefficients of a theoretical semi-variogram function (see Figure 4). The theoretical model here selected is an exponential function with a vertical offset (nugget). The model, represented as a solid line in Figure 4, shows a nugget of 0.35 and a range of 80 km. The latter value represents the distance at which the correlation effects become negligible, while the former indicates the part of the estimation error that will remain unexplained even after kriging.

To better qualify this result we can consider that the average distance between a given station and the nearest one is 13 km and that a circle with diameter 80 km contains, on average, 36 stations. In essence, this demonstrates the existence of a spatial correlation that should be accounted for in the prediction of temperature values. On the other hand, considering that the sill (total spatial variance) is 0.76, the significant amount assumed by the nugget means that about half of the total variance remains unexplained. The mean annual temperature estimate obtained by means of Equation 5 explains 97.9% of the variance and gives a standard error of estimate of 0.472°C. Since the estimated mean annual temperature has a role in the reconstruction of the monthly values, as shown in the application section, it was judged reasonable to sum up the kriging correction factor to the regression model, in order to minimize the prediction error.

Figure 5 reproduces the map of residuals computed with the kriging model in the 1-km grid of Italy. In the figure, structures of spatial organization of the error anomalies are quite evident, confirming that there is matter for further investigation, which possibly requires additional, or more detailed, geographical predictors.

Summarizing, the determination of the average annual temperature at a given point *(i,j)* in the Italian 1-km grid can be achieved as follows:

1. evaluation of parameters E and L for the given pixel having position (i,j) in the digital terrain model matrix;
2. evaluation of $\hat{T}_a$ from Equation 4;
3. selection of the kriged residual term $e_{reg}(i,j)$ from the map in Figure 5;
4. evaluation of the final predicted temperature by means of Equation 5.

MONTHLY TEMPERATURES ANALYSIS

The temperature regime is meant to be the curve of the 12 monthly mean temperatures within the year. Rather than estimating the monthly temperature normals by separate regression equations (e.g., Zheng and Basher, 1996) the whole parameterized curve of the temperature regime is estimated here at each grid point in the Italian territory. In this way we try to build a single model that can demonstrate possible causal relations between some geographical predictors and the within-year variability of temperatures. Such relations are not evident when regression models are built for individual months (see *e.g.* Ninyerola et al., 2000, with regard to the role of continentality in the 12 mean monthly air temperatures).

Examining the temperature regime in different areas of Italy one recognizes that the coldest month is January for all stations. In some peninsular and coastal locations, because of the sea influence, the mean temperature in January is only slightly different from that in February, while in continental areas the mean in the two months differ considerably, reaching differences of 4°C. This difference tends to decrease again for mountain sites above 1,000 m a.s.l.

With regard to the warmest months, the most ‘inland’ areas of Italy (i.e., the Padana plain) have mean temperatures in July that are markedly higher than those of August. This difference decreases in central Italy and almost disappears in the south and in the islands. The annual range of average temperatures varies over the Italian territory between 13°C and 23°C, increasing with the latitude and the distance from the sea, and decreasing with the elevation.

Fitting the temperature regime by Fourier series

The sequence of 12 monthly temperature normals *T(j)* can be well approximated by means of sinusoidal curves obtained by Fourier series:

$$T(j) = A_o + \sum_{i=1}^{N} A_i \cos\left(\frac{2\pi}{\tau_i} j + \frac{\tau}{\tau_i}\phi_i\right) \tag{6}$$

where j = month of the year (1 to 12); A_0 = mean of $T(j)$; τ (= 12) fundamental period of the cycle; N = number of the harmonics; A_i = amplitude, ϕ_i = phase, and τ_i period of the i^{th} harmonic. For the estimation of the Fourier series parameters the cosine argument in Equation 6 can be decomposed to a polynomial form:

$$T(j) = A_0 + \sum_{i=1}^{N}\left[b_i * \cos\left(\frac{2\pi n_i}{12} j\right) + c_i * \sin\left(\frac{2\pi n_i}{12} j\right)\right] \tag{7}$$

where $b_i = A_i \cos(n_i\phi_i)$, $c_i = -A_i \sin(n_i\phi_i)$ and A_0 are parameters that can be estimated by least squares, and where $n_i = \tau/\tau_i$. The amplitude and phase of the i^{th} harmonic can then be obtained as:

$$\begin{cases} A_i = \dfrac{b_i}{\cos(n_i \phi_i)} \\ \phi_i = \dfrac{1}{n_i} \arctan\left(-\dfrac{c_i}{b_i}\right) \end{cases} \tag{8}$$

Having selected a model to estimate the spatial variability of the mean annual temperatures, only the monthly deviations from the annual mean were analyzed, considering two alternatives:

(a) a non-dimensional temperature regime *t(j)*, where

$$t(j) = T(j) / T_a, \tag{9}$$

with T_a as the mean annual temperature. From Equation 9 one obtains that in Equation 7 A_0=1 (see Figure 6a);

(b) a zero-mean temperature regime, where *t(j)* is obtained as

$$t(j) = T(j) - T_a. \tag{10}$$

The above relation produces A_0=0 (see Figure 6b).

Five stations that have substantially different geographic features (described in Table 3) were selected to compare these two alternatives. The shape of the zero-mean temperature regime curve (alternative *b*) is less variable, from one station to another, than that of the other curve. Giordano (2002) showed that the model for representation of the alternative *b* was more accurate than the one for the first alternative. The assessment was made on the quality of reconstruction of the regime curves in all of the considered stations. Results from the alternative *b* were better in terms of R^2 of the estimation of the Fourier coefficients and of the RMSE computed with the estimated curves. So, the zero-mean temperature regime (alternative *b*) was used in the subsequent analyses.

The first attempt to reconstruct the zero-mean temperature regime was made with a one-harmonic Fourier series, with τ_1 =12 months. Parameters A_1 and ϕ_1, estimated for each temperature station, were correlated with the stations' geographical and morphological parameters using the stepwise procedure. Results are reported in Tables 4 and 5. The most efficient models found for amplitude and phase are:

$$\begin{aligned} A_1 &= 4.15 - 0.00044 \cdot E + 0.1209 \cdot L + 0.00184 \cdot Ds - 0.1704 \cdot As - 0.4413 \cdot C \\ \phi_1 &= 1.61 - 0.000035 \cdot E + 0.0197 \cdot L + 0.000114 \cdot Ds - 0.00954 \cdot As \end{aligned} \tag{11}$$

Compared to the model for annual normals, here the introduction of the new geographic variables can be recognized to be effective, producing a significant increase in R^2, even though the final value is not particularly high. All of the values of the *Tstat* assume values greater than 2, which indicates that the parameter coefficients are all significant and the variables are to be kept in the model.

Even though the model results are fair, significant errors occur in correspondence to the highest and lowest values of the curve, in particular when these values persist for two or more consecutive months. To improve the representation, a second Fourier wave with $\tau_2 = 6$ months was introduced:

$$t(j) = A_o + A_1 \cos\left(\frac{2\pi}{12} j + \phi_1\right) + A_2 \cos\left(\frac{4\pi}{12} j + 2\phi_2\right) \tag{12}$$

In this case, least squares regressions of the four parameters on all of the stations produce estimates of A_1 and ϕ_1 identical to those obtained considering only one harmonic. The analysis on the second harmonic parameters shows that A_2 can be considered constant in space, with an average value of 0.75 and a spatial standard deviation of 0.204. On the other hand, ϕ_2 is strongly related to geographic parameters (Table 6). The regression model for the second harmonic is:

$$\begin{aligned} A_2 &= 0.75 \\ \phi_2 &= 4.00 + 0.00016 \cdot E - 0.0458 \cdot L - 0.00044 \cdot Ds \end{aligned} \tag{13}$$

Influence of the parameters included in Equation 13 can be recognized in terms of the sign of the variation they induce on the phase. In fact, an increase in *L* and *Ds* values (which means going northwise and away from the sea) produces an increase in the summer peak, while an increase in E (which means moving to high elevations) produces the opposite effect. This is in agreement with the actual observations discussed at the beginning of the section.

The final model for the temperature regime in Italy is then represented by Equations 10 and 12, with the four parameters obtained respectively by Equations 11 and 13.

Quality of the monthly temperatures estimates

To test the quality of the estimates obtained by the multiple regressions, monthly mean temperatures are reconstructed using the following combined model:

$$\hat{T}(j) = T_a + t(j) = T_a + A_1 \cos\left(\frac{2\pi}{12} j + \phi_1\right) + A_2 \cos\left(\frac{4\pi}{12} j + 2\phi_2\right) \quad (14)$$

with T_a obtained by means of Equation 5 and A_1, A_2, ϕ_1, ϕ_2 obtained through Equation 11 and Equation 13. The performance indices used for testing the estimates in each of the 738 stations are:

- Root mean square error (*RMSE*) of reconstruction of the 12 monthly mean values:

$$RMSE = \sqrt[2]{\frac{1}{12}\sum_{j=1}^{12}\left[T(j) - \hat{T}(j)\right]^2} \quad (15)$$

- Maximum Absolute Error (*MAE*) in the 12 months:

$$MAE = \max[abs(T(j) - \hat{T}(j))] \quad (16)$$

The *RMSE* values obtained on the whole sample of stations has a mean value 0.53°C, a maximum value 2.2°C, and a minimum value 0.12°C. The *MAE* obtained on the whole sample of stations has a mean value 0.99°C, a maximum value 3.8°C, and a minimum value 0.22°C. The use of just one harmonic for the temperature regime produces a mean RMSE equal to 0.74°C and a mean *MAE* equal to 1.39°C. This quantifies the improvement achieved with the introduction of the second harmonic. The above results compare well with those achieved by Zheng and Basher (1996) and Gyalistras (2003).

DISCUSSION AND CONCLUSIONS

The first result of the application presented above relates to the mean annual temperature regression model, which explains 92% of the variance and presents a RMSE of reconstruction (through cross-validation) of 0.89°C. These estimates are improved by application of the kriging technique, which reduces uncertainty by incorporating the spatial correlation effect recognized by the analysis of the variogram. The final standard error characterizing the mean annual temperature is of 0.472°C.

To comment on this result, we must first consider that in this application the temperatures time series were assumed to be stationary. This means that the whole available sample of stations has been analyzed, regardless of the fact that the record length is not the same for all stations. This choice could imply a structural inhomogeneity in the available sample, whose effects depend on the existence of a temperature trend in time. For Italy, Nanni et al. (1998) and Brunetti et al. (2000) estimated a positive trend of 0.7°C in a century (Northern Italy), which means an average of 0.07°C variation in a decade. Considering the overall RMSE resulting from the cross-validation (0.90°C) and the fact that this analysis relates to average temperature variability in space (and its dependence on geographic parameters) the assumption of stationarity can be considered acceptable.

Regarding monthly mean temperatures, the spatial variation of the within-year pattern (temperature regime) is found to depend linearly on ‘continentality’, concavity, and ‘coastal aspect’, in addition to the usual elevation and latitude predictors. The parameters of the two-harmonics Fourier series reproducing the regime have been related to the above predictors through a linear multivariate model. This made it unnecessary to have 12 different models, one for the mean temperature of each month.

What is interesting in this approach is that the results of the regression analysis have shown that some geographic parameters have an influence on the amplitude and phase of the temperature regime, and that their role is quite clearly intelligible. Similar analyses can reveal the importance of other parameters in different geographic contexts. Even more interesting could be the extension of this kind of analysis to the monthly mean of daily minimum or maximum temperature, that have a significant impact in the estimation of average net radiation in large areas.

Extension of the present study to other regions would involve a preliminary analysis on the candidate geographic predictors. They should be necessarily related to the size of the region considered and to its position with respect to seas and continents. In this sense, it is interesting to compare the parameters used here with those used in New Zealand (Zheng and Basher, 1996), whose geographic configuration is similar to Italy.

Further developments of this study involve the possibility to improve the quality of estimates. Improvements can derive from a more accurate morphological representation of the landscape (through a higher resolution DTM), and from additional investigations on factors that could explain the spatial correlation of residuals. In particular, local conditions, such as the effects of the great Alpine lakes, and the relationship between temperature anomalies and vegetation and precipitation (e.g., Zheng and Basher, 1996) can be explicitly investigated.

ACKNOWLEDGMENTS

Comments and suggestions received by three anonymous reviewers are acknowledged and appreciated. The authors also acknowledge the financial support of the National Research Council (CNR-GNDCI grant 03.00022.GN42) and of the Ministry of Education, University and Research (PRIN 2003 Project).

REFERENCES

Agnew, M.D., and Palutikof, J.P. (2000). "GIS-based construction of baseline climatologies for the Mediterranean using terrain variables." *Climate Research*, 14(2), 115-127.

American Meteorological Society (1959). *Glossary of Meteorology*, <http://amsglossary.allenpress.com/glossary/> (Feb., 18, 2006).

Brunetti, M., Maugeri, M., Nanni, T. (2000). "Variations of temperature and precipitation in Italy from 1866 to 1995." *Theoretical and Applied Climatology*, 65, 165-174.

Claps, P., and Sileo, C. (2001). "Temperature characteristics in Southern Italy." *L'Acqua*, (5), 23-31. (in Italian).

Cressie, C. (1993). *Statistics for spatial data*, Wiley-Inter-Science Publication.

Faulkner, D.S., and Prudhomme, C. (1998). "Mapping an Index of extreme rainfall across the UK." *Hydrology and Earth System Science*, 2(2-3), 183-194.

Gentilli, G. (1959). "Mountain temperatures in Tuscany." *Rivista Geografica Italiana*, LXVI-Fasc. IV. (in Italian).

Giordano, P. 2002. *Quantitative analysis of the average temperatures spatial distribution in Italy*. Unpublished M.Sc. Thesis. (in Italian).

Gyalistras, D. (2003). "Development and validation of a high-resolution monthly gridded temperature and precipitation data set for Switzerland (1951-2000)." *Climate Research*, 25(1), 55-83.

Lapen, D.R., and Hayhoe, H.N. (2003). "Spatial analysis of seasonal and annual temperature and precipitation normals in southern Ontario, Canada." *Journal of Great Lakes Research*, 29(4), 529-544.

Ministero dei Lavori Pubblici (1969). *Distribution of air temperature in Italy in the period 1926-1955*, Publication n.21, SIMN, Istituto Poligrafico dello Stato, II ed. (in Italian).

Nalder, I.A., and Wein, R.W. (1998). “Spatial interpolation of climatic Normals: test of a new method in the Canadian boreal forest.” *Agricultural and Forest Meteorology*, 92(4), 211-225.

Nanni, T., Lo Vecchio, G., Cecchini, S. (1998). “Variability of surface air temperature in Italy, 1870-1980.” *Theoretical and Applied Climatology*, 59, 231-235.

Ninyerola, M., Pons, X., Roure, J.M. (2000). “A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques.” *International Journal of Climatology*, 20(14), 1823-1841.

Paiero, P. (1968). “Average annual temperature estimation in the Giulian Prealps.” *Italia Forestale e Montana*, XXIV, 211-214. (in Italian).

Petrarca, S., Spinelli, F., Cogliani, E., Mancini, M. (1999). *Profilo climatico dell’Italia*, ENEA. (in Italian).

Pinna, M. (1977). *Climatologia*, UTET. (in Italian).

Prudhomme, C., and Reed, D.W. (1999). “Mapping extreme rainfall in a mountain region using geostatistical techniques: a case study in Scotland.” *International Journal of Climatology*, 19(12)**,** 1337-1356.

Robeson, S.M., and Janis, M.J. (1998). “Comparison of temporal and unresolved spatial variability in multiyear time-averages of air temperature.” *Climate Research*, 10(1), 15-26.

U.S. Geological Survey. *GTOPO30 Documentation*, <http://edc.usgs.gov/products/elevation/gtopo30/README.html> (Feb., 13, 2006).

Wilks, D.S. (1995). *Statistical methods in the atmospheric sciences*, Academic Press.

WMO (1983). *Guide to Climatological Practices*, 2nd ed., WMO-No. 100, World Meteorological Organization, Geneva.

Zheng, X., and Basher, R. (1996). “Spatial Modelling of New Zealand temperature normals.” *International Journal of Climatology*, 16(3), 307-319.

TABLE CAPTIONS

Table 1. Distribution of the temperature stations of the database with elevation and related percent of the area of the Italian territory.

Table 2**.** Synthesis of the results of the stepwise regression for the annual temperature normal. The columns named 'Error Distribution' report the percent of station presenting reconstruction errors (*e*) of different amounts, indicated in each column. Variables defined in text.

Table 3. Characteristics of the temperature stations considered in Figure 6.

Table 4. Synthesis of the results of the stepwise procedure for the amplitude A_1. Variables defined in text.

Table 5. Synthesis of the results of stepwise procedure for the phase ϕ_1.

Table 6. Synthesis of the results of the stepwise procedure for the second harmonic phase ϕ_2.

FIGURE CAPTIONS

Figure 1: Digital elevation model and temperature stations.

Figure 2: Morphological variables calculated for each station using the geographic representation of the coast-line and a Digital Terrain Model: (**a**) Angle α formed by the minimum distance vector and the South; (**b**) azimuthal angle β of horizon in a generic direction.

Figure 3: Root Mean Square Errors and R^2 evaluated through the cross-validation procedure applied to the annual normals. k is the number of variables used in the regression (see Table 2).

Figure 4: Sample semi-variogram of the residuals of regression (4) and estimated exponential with nugget model (solid line).

Figure 5: Spatial distribution of the kriged residuals.

Figure 6: Diagrams for the 5 stations reported in table 3. (a) Non-dimensional temperature regime; (b) Zero-mean temperature regime.

TABLES

Table 1. Distribution of the temperature stations of the database with elevation and related percentages of the area of the Italian territory.

Elevation (m a.s.l.)	# Stations	% Stations	% Area
E<100	209	28%	23%
100<E<800	400	54%	54%
800<E<1200	78	11%	11%
E>1200	51	7%	12%

Old:

Elevation (m a.s.l.)	**# stations**	**% Stations**	**% Area**
E<100	209	28%	23%
100<E<800	400	54%	54%
800<E<1200	78	11%	11%
E>1200	51	7%	12%

Table 2**.** Synthesis of the results of the stepwise regression for the annual temperature normal. The columns named 'Error Distribution' report the percent of station presenting reconstruction errors (*e*) of different amounts, indicated in each column. Variables defined in text.

Step	R^2	Constant		*E*		*L*		*Ds*		*As*		*C*		Error Distribution (% stations)			
		Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	e<1	1<e<2	2<e<3	e>3
1	0.646	15.77	155	-0.006429	-36.7	-	-	-	-	-	-	-	-	30.5	35.8	23.3	10.4
2	0.924	43.59	81	-0.005428	-64.8	-0.6601	-51.9	-	-	-	-	-	-	76.7	19.9	2.7	0.7
3	0.926	40.17	45.3	-0.005261	-58.1	-0.5745	-26.4	-0.001166	-4.82	-	-	-	-	76.4	20.9	2	0.7
4	0.927	40.33	45.7	-0.005216	-57.3	-0.5808	-26.7	-0.001004	-4.08	0.07015	3.12	-	-	76.6	20.6	2.2	0.7
5	0.927	40.22	45.3	-0.005271	-53.5	-0.5804	-26.7	-0.00109	-4.23	0.06722	2.97	0.1046	1.12	76.8	20.5	2	0.7

STEP	R^2	Constant		*E*		*L*		*Ds*		*As*		*C*		Error Distribution (% stations)			
		Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	e<1	1<e<2	2<e<3	e>3
1	0.646	15.77	154.57	-0.006429	-36.65									30.5	35.8	23.3	10.4
2	0.924	43.59	80.98	-0.005428	-64.76	-0.6601	-51.87							76.7	19.9	2.7	0.7
3	0.926	40.17	45.33	-0.005261	-58.14	-0.5745	-26.41	-0.001166	-4.82					76.4	20.9	2.0	0.7
4	0.927	40.33	45.70	-0.005216	-57.26	-0.5808	-26.74	-0.001004	-4.08	0.07015	3.12			76.6	20.6	2.2	0.7
5	0.927	40.22	45.29	-0.005271	-53.51	-0.5804	-26.73	-0.001090	-4.23	0.06722	2.97	0.1046	1.12	76.8	20.5	2.0	0.7

Table 3. Characteristics of the temperature stations considered in Figure 6. Latitude is in decimal degrees.

Province	Station	E (m a.s.l.)	L (°)	d_{min} (Km)	T_a (°C)
AO	Aosta	583	45.73	212	10.9
SO	Bormio	1225	46.46	304	7.7
AN	Ancona	103	43.61	1	14.7
SA	Battipaglia	72	40.6	10	15.9
CT	Catania	75	37.5	3	18.3

Province	**Station**	***E* (m a.s.l.)**	***L*(°)**	**d_{min} (Km)**	**T_a (°C)**
AO	Aosta	583	45.73	212	10.9
SO	Bormio	1225	46.46	304	7.7
AN	Ancona	103	43.61	1	14.7
SA	Battipaglia	72	40.60	10	15.9
CT	Catania	75	37.50	3	18.3

Table 4. Synthesis of the results of the stepwise procedure for the amplitude A_1. Variables defined in text.

Step	R^2	Constant		*E*		*L*		*Ds*		*As*		*C*	
		Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat
1	0.005	8.89	166	0.000171	1.86	-	-	-	-	-	-	-	-
2	0.398	-1.46	-3.08	-0.000219	-2.97	0.2455	21.9	-	-	-	-	-	-
3	0.461	4.11	5.48	-0.000524	-6.79	0.1061	5.77	0.001897	9.27	-	-	-	-
4	0.528	3.67	5.22	-0.000626	-8.58	0.1226	7.08	0.001476	7.53	-0.18275	-10.2	-	-
5	0.551	4.15	6	-0.000444	-5.74	0.1209	7.16	0.001839	9.18	-0.17041	-9.69	-0.4413	-6.09

STEP	R^2	Constant		*E*		*L*		*Ds*		*As*		*C*	
		Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat
1	0.005	8.89	166.04	0.000171	1.86								
2	0.398	-1.46	-3.08	-0.000219	-2.97	0.2455	21.91						
3	0.461	4.11	5.48	-0.000524	-6.79	0.1061	5.77	0.001897	9.27				
4	0.528	3.67	5.22	-0.000626	-8.58	0.1226	7.08	0.001476	7.53	-0.18275	-10.22		
5	0.551	4.15	6.00	-0.000444	-5.74	0.1209	7.16	0.001839	9.18	-0.17041	-9.69	-0.4413	-6.09

Table 5. Synthesis of the results of stepwise procedure for the phase ϕ_1.

Step	R^2	Constant		*E*		*L*		*Ds*		*As*		*C*	
		Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat
1	0.029	2.44	527	0.000037	4.7	-	-	-	-	-	-	-	-
2	0.731	1.23	44.3	-0.000008	-1.91	0.0288	43.8	-	-	-	-	-	-
3	0.773	1.63	38.1	-0.00003	-6.85	0.0188	18	0.000136	11.6	-	-	-	-
4	0.797	1.61	39.6	-0.000035	-8.43	0.0197	19.8	0.000114	10.1	-0.00954	-9.26	-	-
5	0.798	1.61	39.7	-0.000032	-7.06	0.0196	19.8	0.00012	10.2	-0.00931	-9	-0.0079	-1.85

STEP	R^2	Constant		*E*		*L*		*Ds*		*As*		*C*	
		Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat
1	0.029	2.44	526.84	0.000037	4.70								
2	0.731	1.23	44.29	-0.000008	-1.91	0.0288	43.82						
3	0.773	1.63	38.14	-0.000030	-6.85	0.0188	17.96	0.000136	11.63				
4	0.797	1.61	39.64	-0.000035	-8.43	0.0197	19.75	0.000114	10.06	-0.00954	-9.26		
5	0.798	1.61	39.66	-0.000032	-7.06	0.0196	19.75	0.000120	10.18	-0.00931	-9.00	-0.0079	-1.85

Table 6. Synthesis of the results of the stepwise procedure for the second harmonic phase ϕ_2.

Step	R^2	Constant		*E*		*L*		*Ds*		*As*		*C*	
		Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat
1	0.003	2	147	-0.000034	-1.47	-	-	-	-	-	-	-	-
2	0.621	5.3	55.4	0.00009	6.05	-0.0784	-34.6	-	-	-	-	-	-
3	0.674	4	27	0.000162	10.6	-0.0458	-12.6	-0.000444	-10.9	-	-	-	-
4	0.675	4.01	27	0.000164	10.6	-0.0462	-12.6	-0.000434	-10.5	0.00442	1.17	-	-
5	0.676	4.04	27	0.000175	10.5	-0.0463	-12.7	-0.000412	-9.5	0.00515	1.35	-0.0262	-1.67

STEP	R^2	Constant		*E*		*L*		*Ds*		*As*		*C*	
		Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat	Coefficient	Tstat
1	0.003	2.00	146.70	-0.000034	-1.47								
2	0.621	5.30	55.37	0.000090	6.05	-0.0784	-34.61						
3	0.674	4.00	26.95	0.000162	10.57	-0.0458	-12.56	-0.000444	-10.94				
4	0.675	4.01	26.97	0.000164	10.63	-0.0462	-12.62	-0.000434	-10.46	0.00442	1.17		
5	0.676	4.04	27.02	0.000175	10.46	-0.0463	-12.66	-0.000412	-9.50	0.00515	1.35	-0.0262	-1.67

FIGURES

Figure 1: Digital elevation model and temperature stations in Italy.

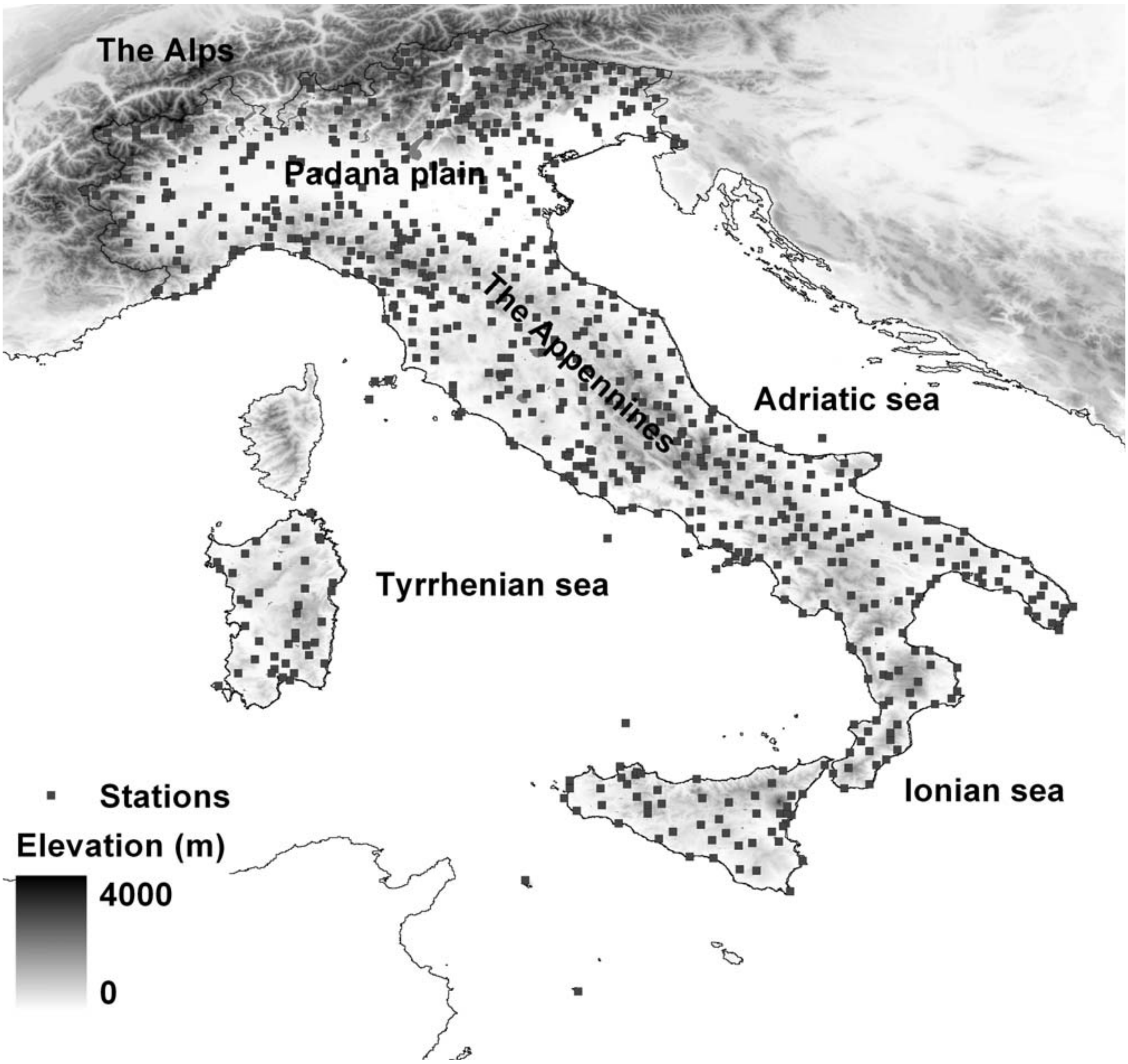
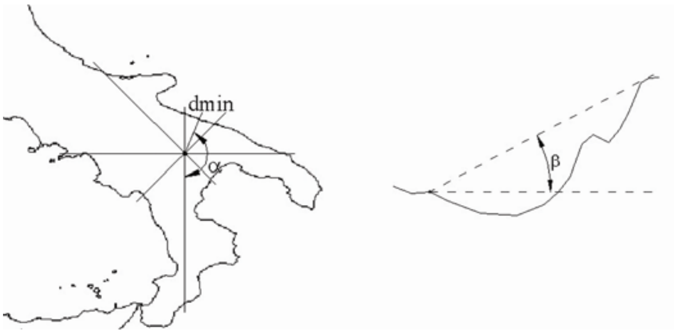


Figure 2: Morphological variables calculated for each station using the geographic representation of the coast-line and a Digital Terrain Model: (**a**) Angle α formed by the minimum distance vector and the South; (**b**) azimuthal angle β of horizon in a generic direction.



(a) **(b)**

Figure 3: Root Mean Square Errors and R^2 evaluated through the cross-validation procedure applied to annual normals. *k* is the number of variables used in the regression (see Table 2).

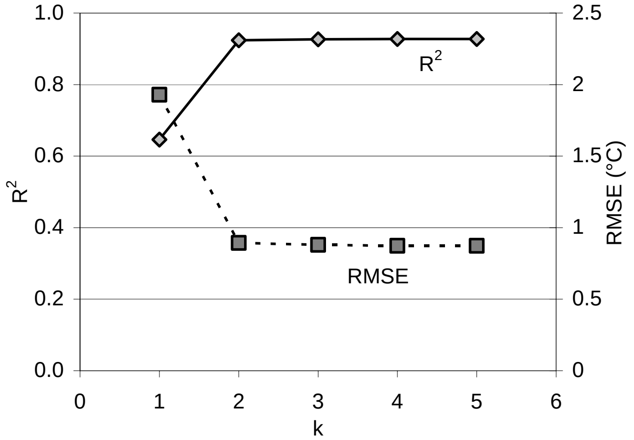


Figure 4: Sample of the residuals of regression (4) and estimated exponential with nugget model (solid line).

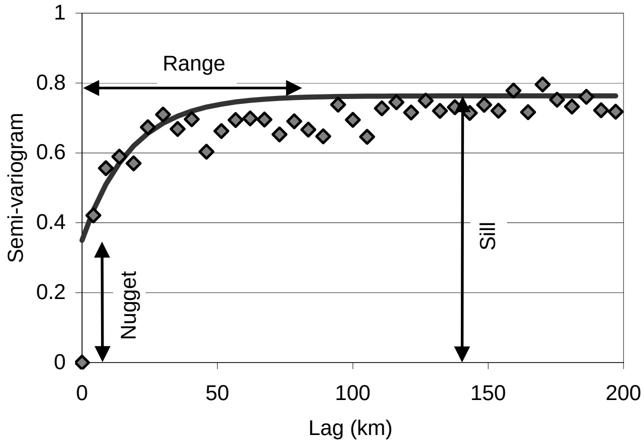


Figure 5: Spatial distribution of the kriged residuals.

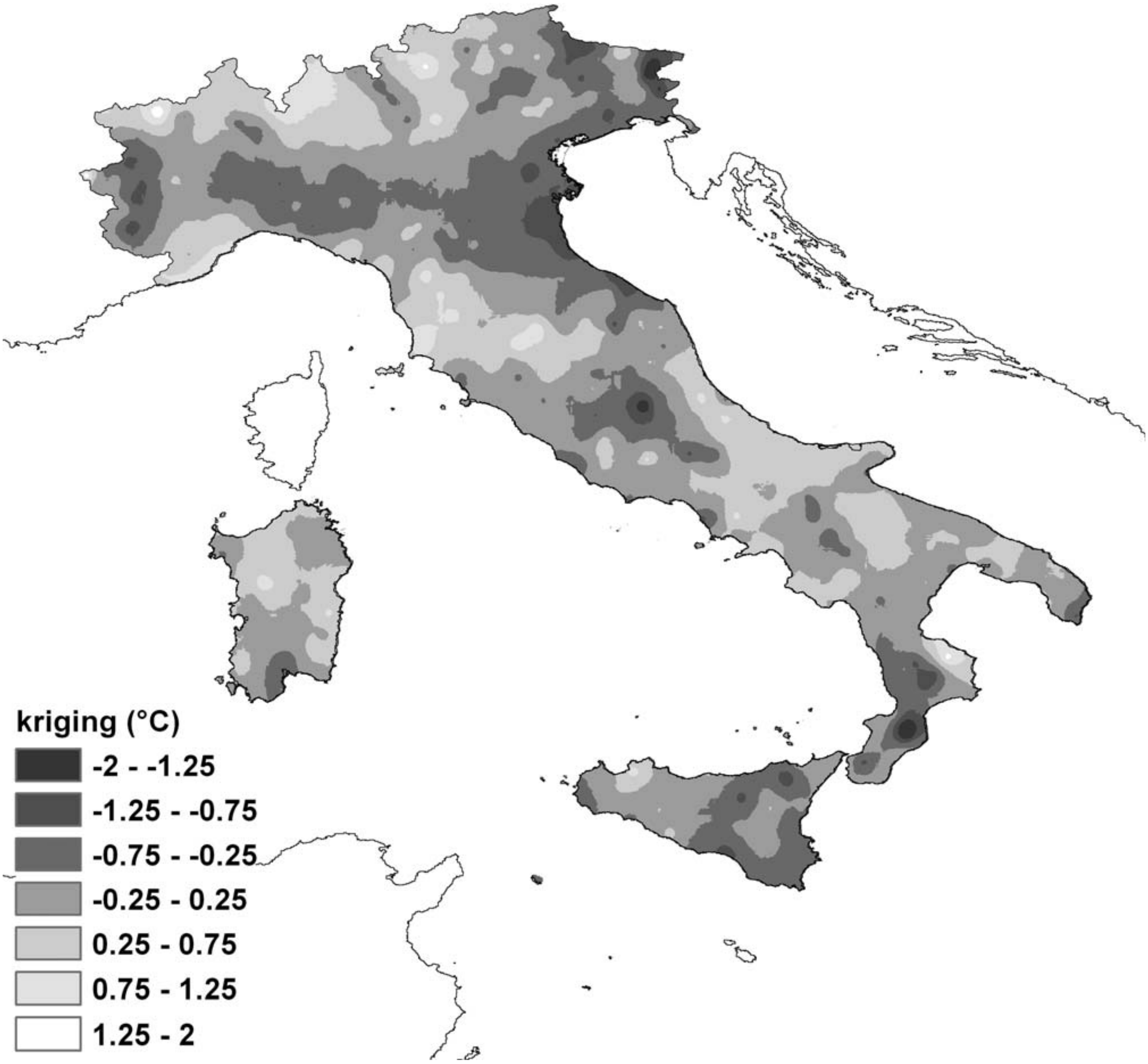
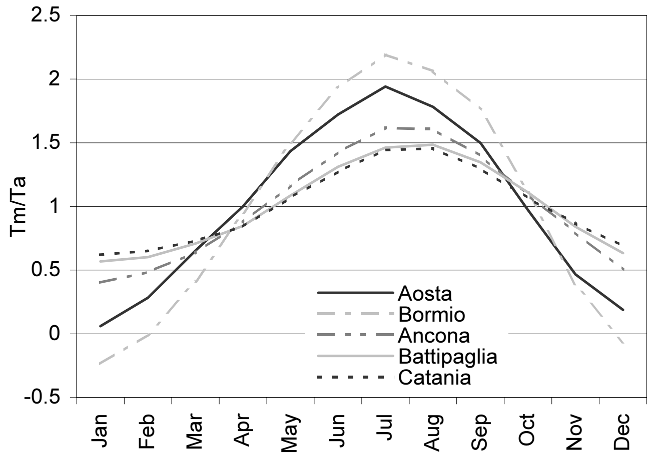
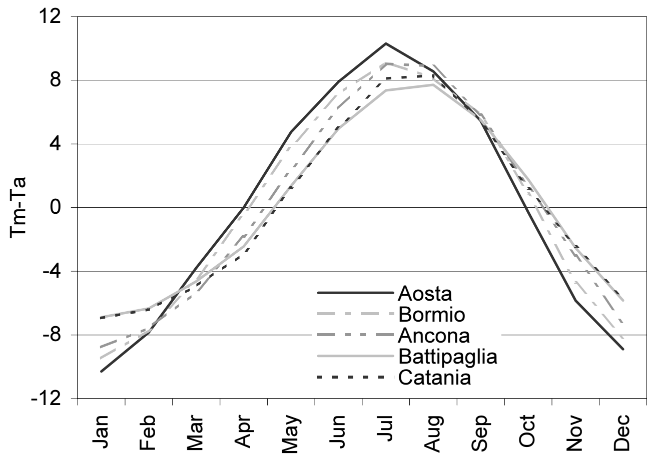


Figure 6: Diagrams for the 5 stations reported in Table 3. (a) Non-dimensional temperature regime; (b) Zero-mean temperature regime.



(a)



(b)

a