4. 1961-90 high resolution temperature climatologies for Italy

4.1 Region of interest: geographical borders and characteristics

The study area for the temperature climatologies is Italy and the extra-Italian territories that encompass the Alps.

Such a geographical region is situated between Longitude 6.5° East and 18.8° East and between Latitude 35.5° North and 47.3° North. Within this area, we did not consider the seas, African regions, Corsica, Malta or regions situated east of Lon 14.2° East if north of Lat 42° N (from now on Lon means Longitude and Lat means Latitude). The total considered area is approximately 400,000 km^2 and it encompasses France, Italy, Switzerland, Austria, San Marino, Vatican City, Slovenia and Croatia.

In the studied area there are two high mountain ridges (the Alps: 1,300 km long, the highest peak is Monte Bianco, 4,810 m; the Apennines, 1,300 km long, the highest peak is Corno Grande, Gran Sasso, 2,912 m), many valleys (Valtellina, Adige Valley, Val D'Aosta and so on), many long rivers (Po and so on), parts of the Mediterranean Sea (subdivided into Ligurian Sea, Tyrrhenian Sea, Ionian Sea and Adriatic Sea) that hug the peninsular area, two main islands (Sicily and Sardinia), many isles (e.g. Eolie, Tremiti, Pantelleria, Lampedusa, Stromboli, Capri and so on), many lakes (Maggiore, Como, Garda, Iseo, Bracciano, Bolsena and so on), a great plain (Po Plain), a few metropolis (Milan, Turin, Genoa, Florence, Naples, Palermo and so on) and other orographic peculiarities that will be described later.

The Climate of Italy is very complex. According to Koppen classification, in Italy there are regions which belong to Mediterranean climate (*Csa*), Mediterranean mild climate (*Csb*), humid subtropical climate (*Cfa*), oceanic climate (*Cfb*), humid continental climate (*Dfb*), cold continental climate (*Dfc*), Tundra climate (*ET*) *(Peel et al., 2007)*.

The same area considered for temperature climatologies was considered for precipitation climatologies and solar radiation climatologies.

Fig.24 The Mediterranean area seen from satellites (Wikipedia, 2010).

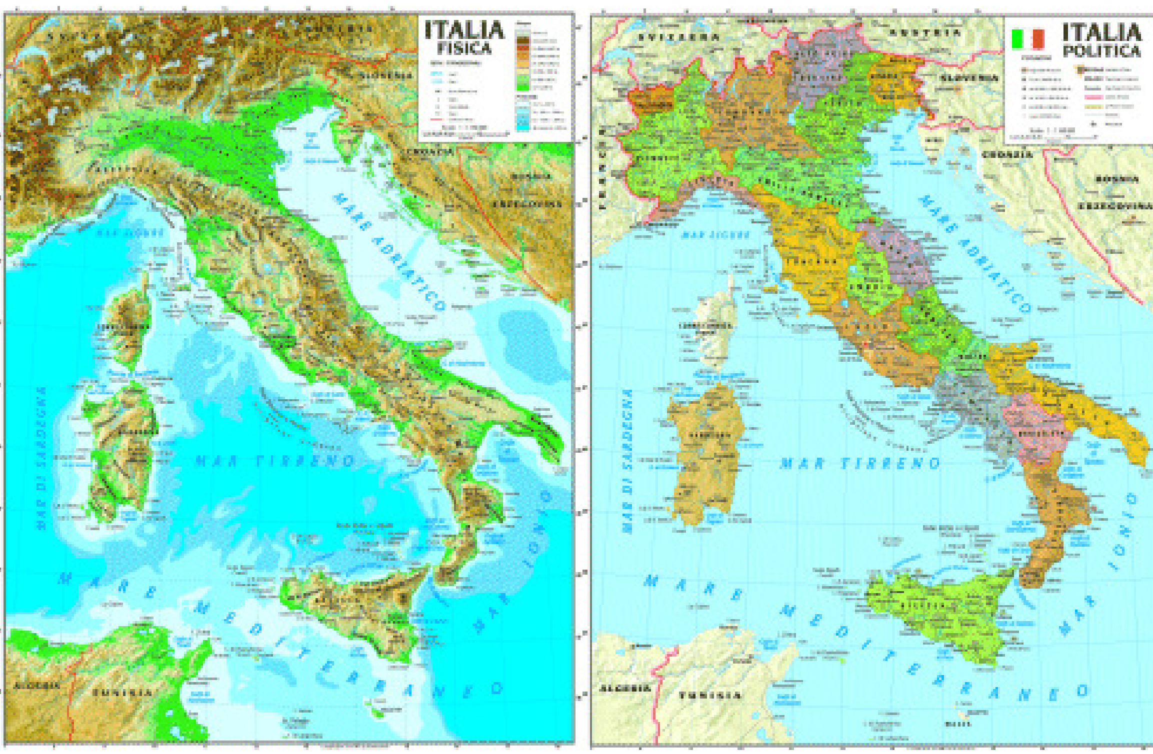

Fig.25 Left: geographic map of Italy; right: political map of Italy (Ortelio website, 2010).

4.2 The temperature records database

4.2.1 The data search: providers and first quality checks

From now on T_M means mean temperature, T_N means minimum temperature, T_X means maximum temperature.

The first step of this data rescue was dedicated to set up an adequate temperature database. In Italy, there is no official data provider; this situation complicated our work.

The data collection was performed with support of quality-check rules. In particular we accepted:

- records with at least 15 years of data in 1850-2010 period:
  - a year was considered only if all monthly data were available;
  - a month was considered only if a record has more than 25 daily data (23 for February);
- if T_N, T_X and T_M were available, we accepted only records where $(T_N+T_X)/2 = T_M$.

In case of two or more stations located very close with the same name, we collected the station with data nearest to 1961-1990 period and fully-reported metadata.

We obtained the temperature data from a list of previous projects and data providers.

- As far as **previous projects** are concerned the main data source was the *ECSN HRT-GAR* Project (European Climatic Support Network High Resolution Temperature climatologies for the Greater Alpine Area project; *http://www.zamg.ac.at/forschung/klimatologie/klimamodellierung/ecsn_hrt-gar/; Hiebl et al., 2009*). Within this project, a dataset of 1,734 monthly T_M normals of the 1961-1990 period was collected for a large European region centred over the Alps. Out of these 1,734 stations, 732 were considered in this research in order to become the basis of our data rescue for the northern part of Italy. Because these 732 station data are provided in monthly normals only, we searched for the corresponding time series. After the complete data search (see next pages), the *ECSN HRT-GAR* project was mainly a data source of climatic normals of the countries surrounding Italy.

- A relevant number of climatic normals was also recovered from the *DBT-ENEA* database (Archivio Climatico *DBT*, DataBase delle Temperature, Ente Nazionale Energia e Ambiente, *Petrarca et al., 1999*; *http://clisun.casaccia.enea.it/Pagine/Index.htm*). This database consists of 729 stations (T_N-T_M-T_X) data that were collected from the main Italian data providers (see the following discussion). Besides the 1961-1990 normals, the information on the period covered by the data is also provided. Such information was used in this research both to filter the data and to convert them into 1961-90 data (see Chapter 4.2.4).

Other **datasets of climatic normals** that were used are the following ones.

- The Italian Annals of 1926-55 monthly T_N-T_M-T_X data *(Servizio Idrografico, 1966)*. This monthly data are available only in paper formats. In this research, we digitalised the monthly averages of 612 stations, which are available on the annals together with the monthly data.

- The Atlante Climatico della Val d'Aosta *(Mercalli et al., 2003)* and 2 other records which were provided by the Società Meteorologica Italiana (*SMI*): 14 stations (T_N-T_M-T_X data).

- A similar publication by *SIAS* Sicily (*http://www.sias.regione.sicilia.it/*, Servizio Informativo Agrometeorologico Siciliano): 55 stations (T_N-T_M-T_X).

- Data tables from the website of *ARPA* Emilia-Romagna (Agenzia Regionale per la Protezione dell'Ambiente Emilia Romagna, *http://www.arpa.emr.it/*): 67 stations (T_N-T_M-T_X).

- Data tables from the website of the Italian Air Force (AMI, Aeronautica Militare Italiana, *http://www.meteoam.it/):* 110 stations (T_N-T_M-T_X).

- Data tables from the website of the *SCIA-APAT* Project (*SINANET*, Sistema di raccolta dati Climatologici di Interesse Nazionale, *http://www.scia.sinanet.apat.it/*): 59 stations (T_N-T_M-T_X data).

For all these data, as we did for the *ECSN HRT-GAR* ones, we collected the data again, where possible, in order to have time records also and not climatic normals only.

As far as **time records** are concerned we used the following data sources:

- Secular homogenised records set up by the University of Milan and *ISAC-CNR*. They include the records presented in *Brunetti et al. (2006b)*: 67 T_M stations and 48 T_N-T_X stations.

- Secular homogenised records set up by the University of Milan in the frame of the Kyoto Lombardia Project (*http://www.kyotolombardia.org/)*: 6 stations (T_M data).

- Italian Air Force records from previous researches set up by the University of Milan and *ISAC CNR*: about 100 homogenised station records (T_N-T_M-T_X data).

- The data collection of some sections of the former Italian former Servizio Idrografico e Mareografico, whose network is now managed by the Italian regions. We collected these data by different providers. In particular:
- 37 stations (Toscana) by Idropisa (*http://www.idropisa.it/*, T_N-T_M-T_X data);
- 31 stations (Abruzzo) by Regione Abruzzo (*http://www.regione.abruzzo.it/xIdrografico/index.asp*, T_N-T_M-T_X data);
- 68 stations (Calabria) by Protezione Civile Calabria (*http://www.protezionecivilecalabria.it/*, T_M data);
- 23 stations (Marche) by Protezione Civile Marche (*http://www.protezionecivile.marche.it/*, T_N-T_M-T_X data);
- 12 stations (Molise) by private communications (T_M data);
- 81 stations (Puglia) by Protezione Civile Puglia (*http://www.protezionecivile.puglia.it/*, T_N-T_M-T_X data);
- 29 stations (Liguria) by *ARPA* Liguria (*http://www.arpal.org/*, T_N-T_M-T_X data);
- 8 stations (Lombardia) by *ARPA* Lombardia (*http://ita.arpalombardia.it/ita/index.asp*, T_N-T_M-T_X data);
- 102 stations (Piemonte) by *ARPA* Piemonte (*http://www.arpa.piemonte.it/*, T_N-T_M-T_X data);
- 74 stations (Sardegna) by *ENAS* (Ente Acque della Sardegna, the former Ente Autonomo del Flumendosa, *http://www.enas.sardegna.it/* , T_M data);
- 58 stations (Bolzano) by Provincia Autonoma di Bolzano (*http://www.provincia.bz.it/*, T_N-T_M-T_X data);
- 53 stations (Trento) by Meteo Trentino (*http://www.meteotrentino.it/*, T_N-T_M-T_X data).

- A collection of records from the Servizio Idrografico set up by Italian colleagues Paola Nola (University of Pavia), Renzo Motta (University of Torino) and Marco Carrer (University of Padua) for researches on the impact of climate change on Alpine forests (e.g., *Nola et al. (1996)*): 183 stations (T_N-T_M-T_X) data.

- The <u>*UCEA* database</u> (Ufficio Centrale di Ecologia Agraria, *http://www.cra-cma.it/*): 126 stations (T_N-T_M-T_X data).

- The database of the Italian energy board, i.e. <u>*ENEL*</u> (Ente Nazionale Energia eLettrica, *http://www.enel.it/it-IT/*). This database is organised in regional sections. In particular we collected:
- 49 stations for Lombardia (T_N-T_M-T_X data);
- 47 stations for Piemonte (T_N-T_M-T_X data).

- The <u>*NCDC-GSOD* global dataset</u>: (National Climatic Data Center, Global Surface Observation of the Day, *http://www.ncdc.noaa.gov/oa/about/whatsnew.htm*): approximately 300 stations (T_N-T_M-T_X data).

- Some other <u>minor sources</u> as <u>monographic books</u>.

The total number of stations that was considered was approximately 4,000. It included however a number of duplicates, as well as stations, that could not be used because they had insufficient data.

After a first check on coordinates, elevation, and names, we obtained a database of 1,524 T_M and 1,155 T_N-T_X. At this point, we did not perform cross-stations quality checks. Nevertheless, we rejected more than 60 T_M stations and more than 20 T_N-T_X stations because of imprecise metadata or evidently wrong geographic locations.

4.2.2 Geographic, elevation, and other deeper metadata checks

Because of different coordinate systems, of hand-made written errors in reports or in providing metadata, and of old metadata not updated correctly, the stations can be wrongly located.

First, we assigned to every station a longitude, latitude and elevation value, which corresponds to the grid cell of the *USGS* digital elevation model *(USGS website)* in which the station is situated. We wrote a simple Fortran code to label the 1,524 T_M with the cited parameters. The T_N-T_X dataset is a subset of the T_M dataset, thus we performed geographic quality checks only on T_M dataset and we consequently accepted or rejected the corresponding T_N-T_X records.

A fast comparison between metadata elevations and digital elevation model elevations showed an average discrepancy of 89.4 m. We selected the stations where this discrepancy in absolute value is higher than 200 meters. This happens in 252 times. If we consider that the temperature dependence on elevation ranges between approximately 3°C/1km and 7°C/1km at the latitudes of the area under investigation, we can easily see how discrepancies in elevation of more than 200 meters can be important. By means of Google Earth™ (*http://earth.google.com/intl/it/*), we manually checked these 252 stations and we corrected 87 elevation values and 73 longitude and/or latitude values.

In the next page we show an example of this re-collocation of a "suspect" station (see fig. 26), here we show a direct comparison:

Provider's metadata (*ENEL* Piemonte) / corresponding *DEM*'s grid cell:
Diga Rochemolles Station 45.132° N ; 6.767 °E ; 1999 m / 45.129° N ; 6.771 °E ; 2272 m
Absolute elevation's difference between station and *DEM*: 273 m

Real station collocation from Google Earth:
Diga Rochemolles Station 45.131° N ; 6.764 °E ; 1953 m / 45.129° N ; 6.763 °E ; 1975 m
Absolute elevation's difference between station and *DEM*: 22 m

Fig.26 Yellow placeholder: Diga Rochemolles; purple placeholder: station position as labelled by provider's metadata; green placeholder: real station's position (Google Earth™)

According to the provider's metadata, the station would be collocated not far from the real position for longitude and latitude, but this slight difference would have caused an important elevation discrepancy as referred to *DEM*. In fact, the station would be collocated in a grid cell characterised by an elevation of 2,272 m, i.e. 273 meters higher than the real station elevation. Whilst, after the re-collocation, it is in a grid cell characterised by an elevation of 1,975 m, only 22 meters higher than the real elevation.

During this check, we rejected 6 stations because the geographic and elevation metadata were completely wrong and consequently not correctable: *Alessano, Castelnuovo Garfagnana, Mongiana, Nus St.Bathelemy, San Cassiano 2, Venegono Inferiore*.

After the geographic and elevation check, the average absolute discrepancy between metadata and real elevation was 76.5 m. Thus, the improvement was 12.9 m. This check cannot be performed automatically, therefore it is not possible to check every single station, but our improvement (approximately 15%) will reduce the statistical errors of the model (see next paragraphs for details).

4.2.3 Dataset conversion to the 1961-1990 period

Following the methodology adopted by (*Hiebl et al., 2009)*, we corrected the inhomogenities due to nationally different methods of the estimation of daily means.

At this point, we calculated the monthly averages of the 1,518 T_M stations and of the 1,155 T_N-T_X stations for the data period.

The dataset covers a wide temporal interval, i.e. from 1851 to 2008, thus it is necessary to transform the monthly averages to the common 1961-1990 reference period (see the next pages for details). The quality checks were postponed after the conversion (see Chapter 4.2.4).

For the Italian territory, 1° Lon x 1° Lat anomaly grids for mean, minimum and maximum temperature, are available from *Brunetti et al. (2006b)*. The dataset used in order to realise such anomaly grids is the secular dataset also used in *Brunetti et al. (2006b)*, the temporal range of the grids is 1851-2010; thus it is suitable for our purpose.

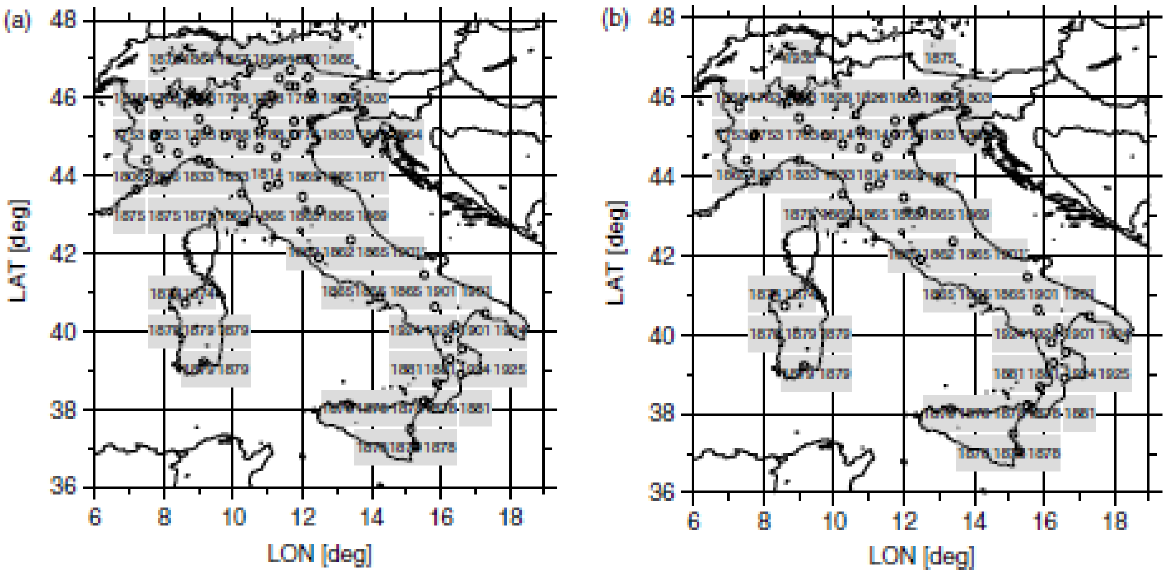

Fig:27-28 Anomaly grid start years. Left: for T_M. Right: for T_N-T_X (Brunetti et al., 2006b)

By means of a Fortran code, we assigned to each station a radial Gaussian weight which was set up to convert the station data to 1961-1990 data using the nearest 3 grid points in the anomaly grids. The high spatial coherence of the anomalies allows us to limit to 3 surrounding points. We tried an inverse distance weighting (with radial Gaussian weights) model using 5 or 7 grid points also, but results were quite identical. It is important to underline that we used a different anomaly grid for each variable (T_N, T_X, T_M) and either for each month, thus we used 36 anomaly grids.

To summarize, we proceeded like it is showed in the example below.

Station / coordinates / January T_M 1926-1955 station data:

Asti / 44.901 °N ; 8.170 °E / -0.6 °C

Station / coordinates / January T_M anomaly from 1926-1955 to 1961-1990 :

Asti / 44.901 °N ; 8.170 °E / -1.2 °C

Station / coordinates / January T_M converted data to 1961-1990:

Asti / 44.901 °N ; 8.170 °E / 0.6 °C

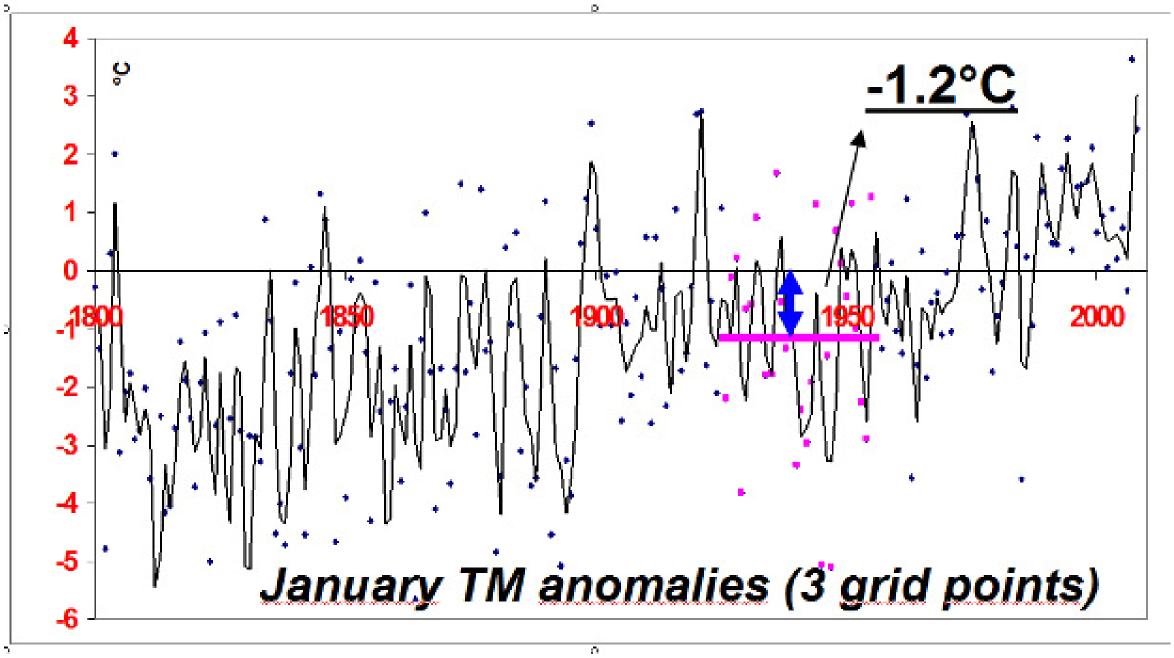

Fig.29 Anomaly conversion for Asti, T_M, January, 3 nearest anomaly grid points

The anomaly value of -1.2 °C means that in 1926-1955 Asti was 1.2 °C colder than in 1961-1990, thus the anomaly must be subtracted to get 1961-1990 data.

This methodology was used for all the datasets. Finally, we had 1,518 monthly 1961-1990 T_M station data and 1,155 monthly 1961-1990 T_N-T_X station data.

4.2.4 Quality checks on the 1961-1990 temperature dataset

Once our data were all converted into 1961-1990 monthly "clinos" (climate normals), we performed an elevation monthly linear regression de-trending for each variable in order to have a sea-level dataset (i.e. we removed the elevation effects from the data). Then we performed data quality checks by means of the following rules, month by month, for T_N-T_X-T_M data, using a Fortran code.

Data quality check rules :

- for each station we selected the nearest 10 surrounding stations;
- we averaged the nearest 10 surrounding station values with an inverse distance Gaussian weighting model;
- we compared the station data with the weighted average from the 10 nearest surrounding stations.

A station was rejected if :

- a single monthly value (it is enough for one variable only) differs, in absolute values, more than 5.0 °C from the averaged value of the 10 nearest stations subset ;
- all the monthly values (it is enough for one variable only) differ, in absolute values, more than 3.0 °C from the averaged values of the 10 nearest stations subset ;
- all the monthly values (for all 3 variables) differ, all with the same minus or plus sign, more than 2.5 °C from the averaged values of the 10 nearest stations subset.

The majority of the stations removed from the dataset were part of *DBT-ENEA*, Provincia di Bolzano and 1926-1955 *SIMN* datasets. It must be said that these datasets provided us data in which were not subjected to homogenization procedures.

All other stations were accepted, with the exception of 3 single special cases (*Diga Sampeyre, Castel Volturno and Fabriano*) which were rejected manually.

In the end, we rejected 25 T_M stations and 17 T_N-T_X stations after the data quality check.

4.2.5 1961-1990 T_N-T_M-T_X dataset used for the Italian climatologies

After the checks, the dataset was reduced to **1,493** T_M stations and **1,138** T_N-T_X stations.

The data densities are approximately: 1/268km^2 for T_M, 1/351km^2 for T_N and T_X. Our temperature dataset for Italy shows a higher density if we compare such data densities with other hi-resolution datasets: examples given, *Pan et al. (2004)* obtained a T_M density of approximately 1/16,810km^2 for China, *Hancock et al. (2006)* obtained a T_M density of approximately 1/3,843km^2 for Australia, *New et al. (2002)* obtained a T_M density of approximately 1/11,626km^2, for the Globe, *Ninyerola et al (2007)* obtained a T_M density of approximately 1/357km^2 for Spain and *Daly et al. (2009)* obtained a T_M density of approximately 1/963km^2 for the conterminous USA.

For a territory with a complex orography as Italy, the vertical distribution of station records is a critical task. Out of 1,493 T_M stations, 159 are "high-mountain" stations (more than 1,500 m), 357 are "mountain" stations (between 800 m and 1,499 m), 470 are "hill" stations (between 300 and 799 m), 507 are "plain" or "coast" stations (less than 300 m).

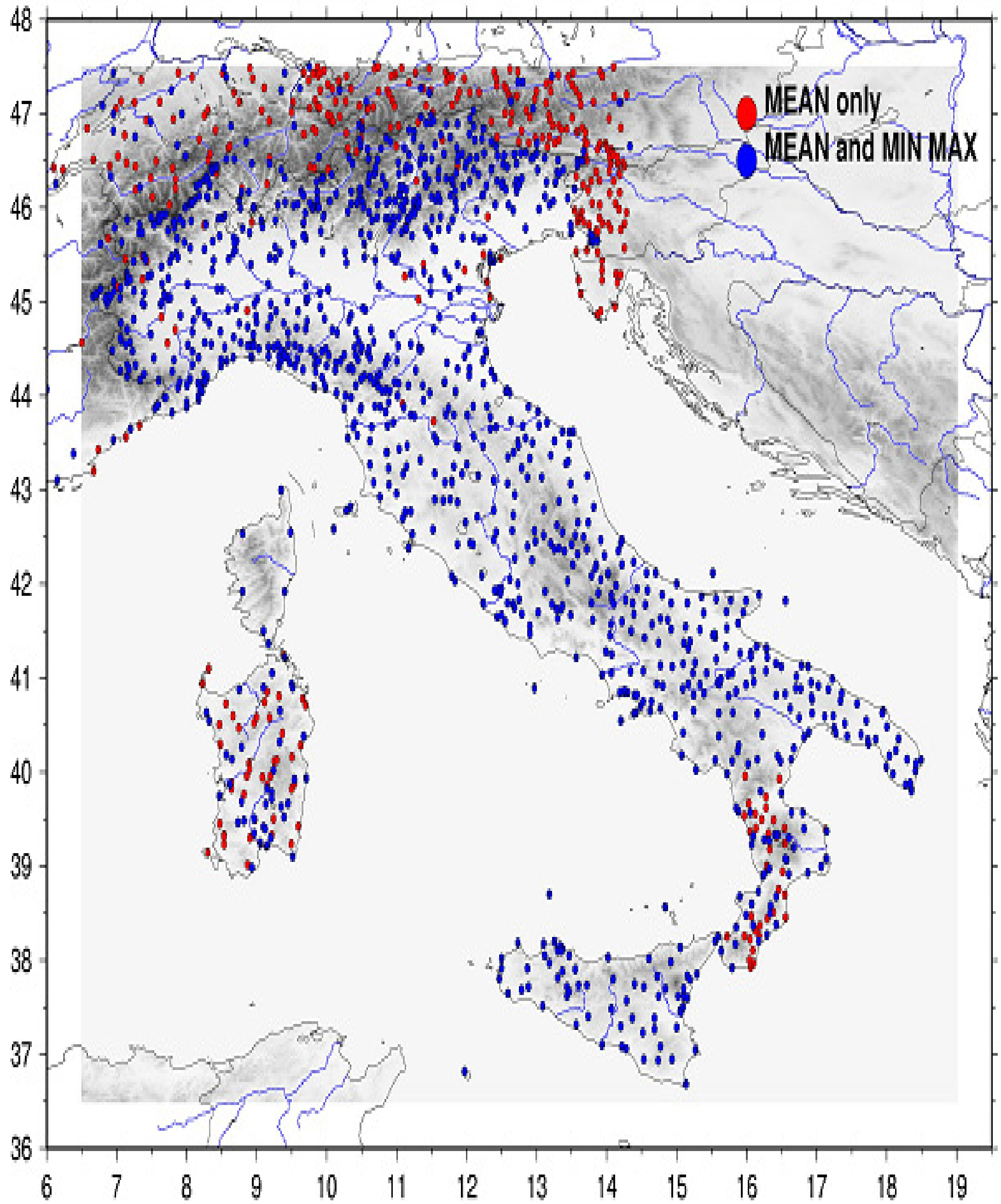

Fig.30 Temperature dataset distribution: T_M-T_N-T_X stations (blue dots) and T_M stations only (red dots)

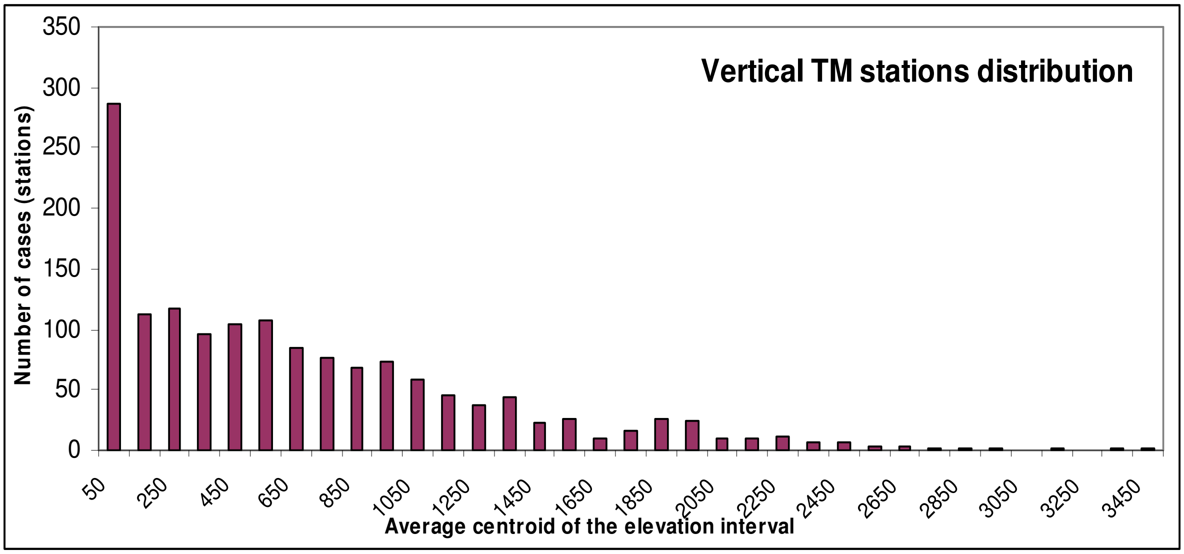

Fig.31 Vertical distribution of mean temperature stations

Eventually, we list the number of accepted stations in the final dataset by provider source: 7 from secular homogenised records dataset of the University of Milan and *ISAC-CNR*, 88 from *AMI* (65 from the homogenized dataset, 23 from the website), 47 from *ARPA* Emilia-Romagna, 20 from *ARPA* Liguria, 6 from *ARPA* Lombardia, 49 from *ARPA* Piemonte, 12 from Atlante della Val d'Aosta, 64 from *DBT-ENEA* database, 64 from *ENEL* (39 from *ENEL* Lombardia, 25 from *ENEL* Piemonte), 257 from *ECSN HRT-GAR* database, 31 from MeteoTrentino, 149 from Nola-Motta dataset, 62 from *ENAS*, 7 from Idropisa, 10 (Molise) from private communications, 29 from Regione Abruzzo, 237 from 1926-55 Hydrographic Servie Annals, 54 from Protezione Civile Calabria, 5 from Protezione Civile Marche, 73 from Protezione Civile Puglia, 51 from Provincia Autonoma di Bolzano, 48 from *SIAS* Sicily, 8 from *SCIA-APAT* website, 70 from *UCEA*, 1 from a monographic book, 42 from World *NCDC-GSOD* database.

Approximately 4 months were spent getting this updated version of the initial database. Various labels were assigned to each station (see next paragraphs for details about rasters): a list number, the name of provider, the name of station, the original data period, the number of complete years, latitude, longitude; corresponding *DEM* grid cell parameters: elevation, latitude, longitude elevation, slope, aspect, distance from the coast, land cover class; binary yes-no parameters: "lake", "sea", "Po plain"; other parameters: "*NCEL*", "macro-aspect"; correlated physical variables: 12 monthly anomaly radiation anomalies; temperature 1961-90 values: 12 monthly T_N values, 12 monthly T_M values, 12 monthly T_X values.

4.3 Data analysis and 1961-1990 T_N-T_X-T_M models for Italy

4.3.1 The spatialization model for temperatures: *MLR* plus local and global improvements plus residual *GIDW*: motivations of the choice.

The main motivation behind the temperature models consists in the facts that our model should represent the spatial climate variability and that every parameterization of the deterministic part of the climate signal ought to have a strong physical explanation. The same basic idea lies behind the temperature climatologies for the *GAR* which were realized with the contribution of this project *(Hiebl et al., 2009)*.

We applied the same methodology to the models for T_N, T_X and T_M. First, the deterministic part of the signal was studied by means of a *MLR* (vs. longitude, latitude and elevation). Then we locally studied the residuals and we introduced many different gridded independent variables in order to capture the secondary deterministic temperature variability. In the end, we gridded the stochastic residuals with a geographical inverse distance weighting (with Gaussian weights and with semi-variograms that decided the radial search distance limits).

We preferred a *MLR* for the deterministic part rather than a kriging or a splines model because this technique, and a simpler *LR* methodology, were already tested in the *GAR* area, in the northern part of Italy, and in the North-Eastern part of Italy in the frame of this PhD project *(Hiebl et al., 2009; Brunetti et al., 2009b; Brunetti et al., 2010*). Furthermore, kriging is preferably only used with an external drift variable (i.e. elevation) and, because of the geographical complexity of Italy, a tri-variate external drift is preferable. We could use a tri-variate *TPS* as well, but it would yield similar results with a stronger computational effort. *PRISM* is not the best choice if the vertical distribution of the stations is not very homogeneous and if the station density itself is not very high. In fact, after some tests that led to too high *MAEs* for our purposes, we decided to use *MLR* instead of *PRISM* for temperatures.

Local improvements were already used in *Hiebl et al. (2009)*, but here we introduced some raster independent variables: "macro-aspect", an *LR* for the top-valley effect, a solar radiation anomaly versus temperature transfer function, and other variables.

The stochastic part was studied by a *GIDW* (similar to *PRISM*) that balanced between the statistical semi-variogram and the a-priori knowledge hypothesis on local climate features.

From a geostatistical point of view, similar models were applied to temperatures by *Hiebl et al. (2009)*, *Ninyerola et al (2002)*, *Ninyerola et al. (2007)*, *Lennon et al. (2005)* and so on.

4.3.2 Step 1: *MLR* versus longitude, latitude, and elevation

• Elevation-Longitude-Latitude grid: *USGS GTOPO30 DEM*

We wrote a simple Fortran code to assign to each station three parameters: elevation, longitude, and latitude. The code searches for the grid cell whose central longitude and latitude are nearest to the station geographical coordinates; next, the code assigns the parameters of the grid cell to the station.

The raster for elevation, longitude, and latitude is the *USGS GTOPO30* (see fig. 32) Digital Elevation Model (U.S. Geological Survey; *USGS website*) with the horizontal resolution of 30-arc-second. Such a resolution is approximately 1 km^2 at Italian latitudes and this is the resolution of the temperature models. *GTOPO30* is a global *DEM* realised by *EROS* (Earth Research Observation and Science) and it is projected in *WGS84* geographical coordinates.

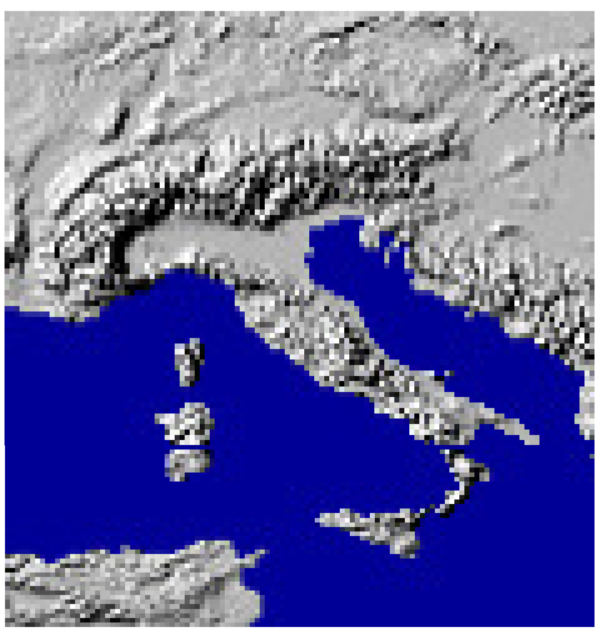

Fig.32 DEM USGS GTOPO30: southern part of w020n90 and northern part of w020n40 (USGS website)

USGS GTOPO30 was also used for precipitation climatologies and for solar radiation climatologies in the framework of this PhD thesis.

• Mean Temperature

We performed monthly *MLR* on 1961-1990 temperature data; each month was subjected to a different *MLR* that has the following basic equation:

$$TM_S(LAT_S, LON_S, ELEV_S) = a \cdot LAT_S + b \cdot LON_S + c \cdot ELEV_S + d \quad \textbf{(54)}$$

where the subscript $_S$ means "station data"; TM_S (or equivalently T_{MS}, are the stations regressed versus their corresponding metadata (i.e. latitude, LAT_S, longitude, LON_S, elevation, $ELEV_S$), the coefficients *a, b, c* are the latitude, longitude and elevation lapse rates, *d* is the T_M interception.

We did not use all the 1,493 T_M stations because we excluded from *MLR* a subset made of the stations located at less than 15 km from the sea and the stations at less than 4 km from the lakeshores. This is because the sea and the lake effect can lead to biased values of the elevation-longitude-latitude lapse rates. Thus we used 1,143 T_M stations.

The *MLR* for the 12 months produced these coefficients:

	a (°C / °LAT)	b (°C / °LON)	c (°C / km)	d (°C)
JAN	-1.11	-0.15	-3.88	54.83
FEB	-0.88	-0.13	-4.78	46.57
MAR	-0.64	-0.10	-5.61	39.17
APR	-0.46	-0.04	-6.10	34.18
MAY	-0.43	0.02	-6.14	36.46
JUN	-0.54	-0.01	-6.19	45.56
JUL	-0.69	-0.08	-6.11	55.76
AUG	-0.75	-0.04	-5.95	57.35
SEP	-0.70	-0.02	-5.56	51.41
OCT	-0.76	-0.04	-4.84	49.29
NOV	-0.98	-0.05	-4.32	53.13
DEC	-1.15	-0.14	-3.78	57.89

Tab.1 Monthly coefficients of the MLR model for mean temperatures

Then we used these monthly coefficients to model T_M for Italy (we obtained the grids using dedicated Fortran codes):

$$TM_M(LAT_{DEM}, LON_{DEM}, ELEV_{DEM}) = a \cdot LAT_{DEM} + b \cdot LON_{DEM} + c \cdot ELEV_{DEM} + d \qquad \textbf{(55)}$$

where $_M$ means modelled and $_{DEM}$ refers to the *DEM*'s values.

We show, for the intermediate steps, only January and July T_M maps. Maps were realised with the free tool *GMT*™ (Generic Mapping Tools, version 4.5.2) which were created by the School of Ocean and Earth science and technology, University of Hawaii (*http://gmt.soest.hawaii.edu/*).

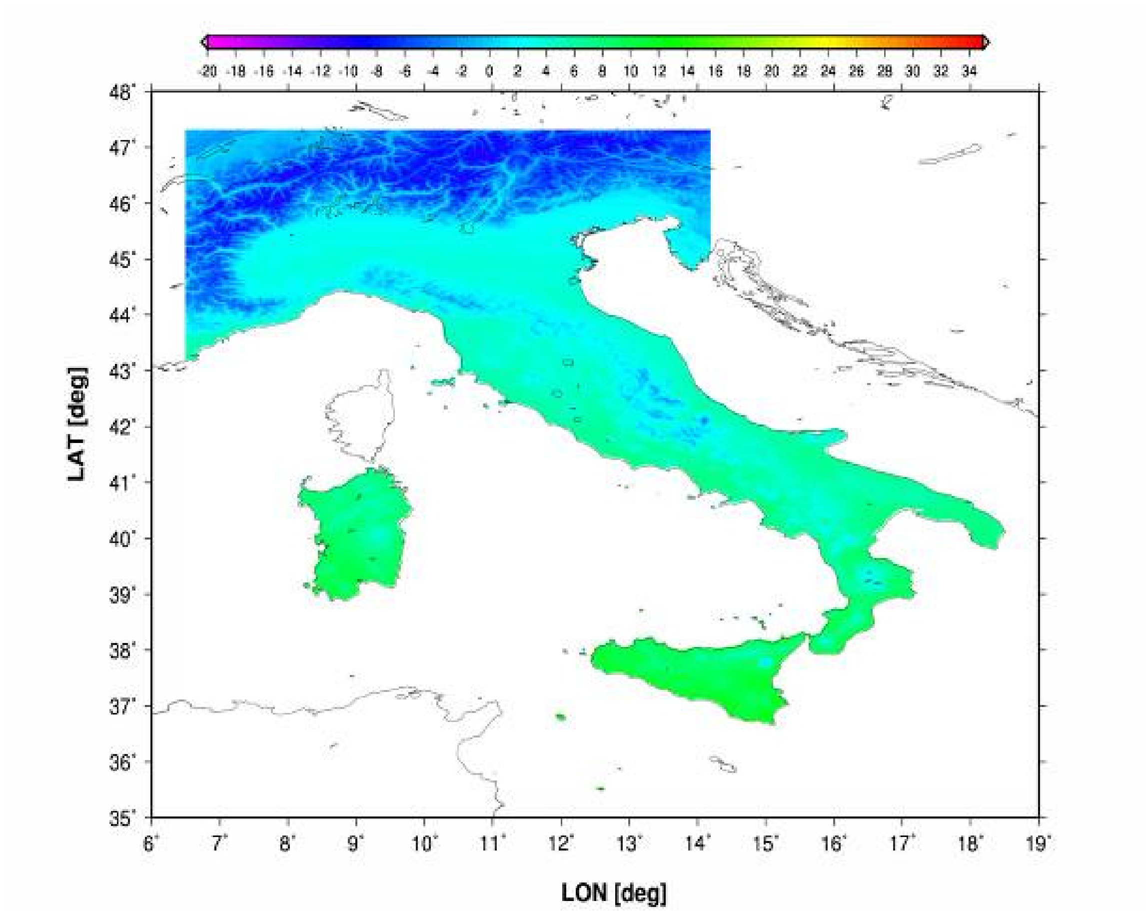

Fig.33 January 1961-90 mean temperature map after MLR model in °C

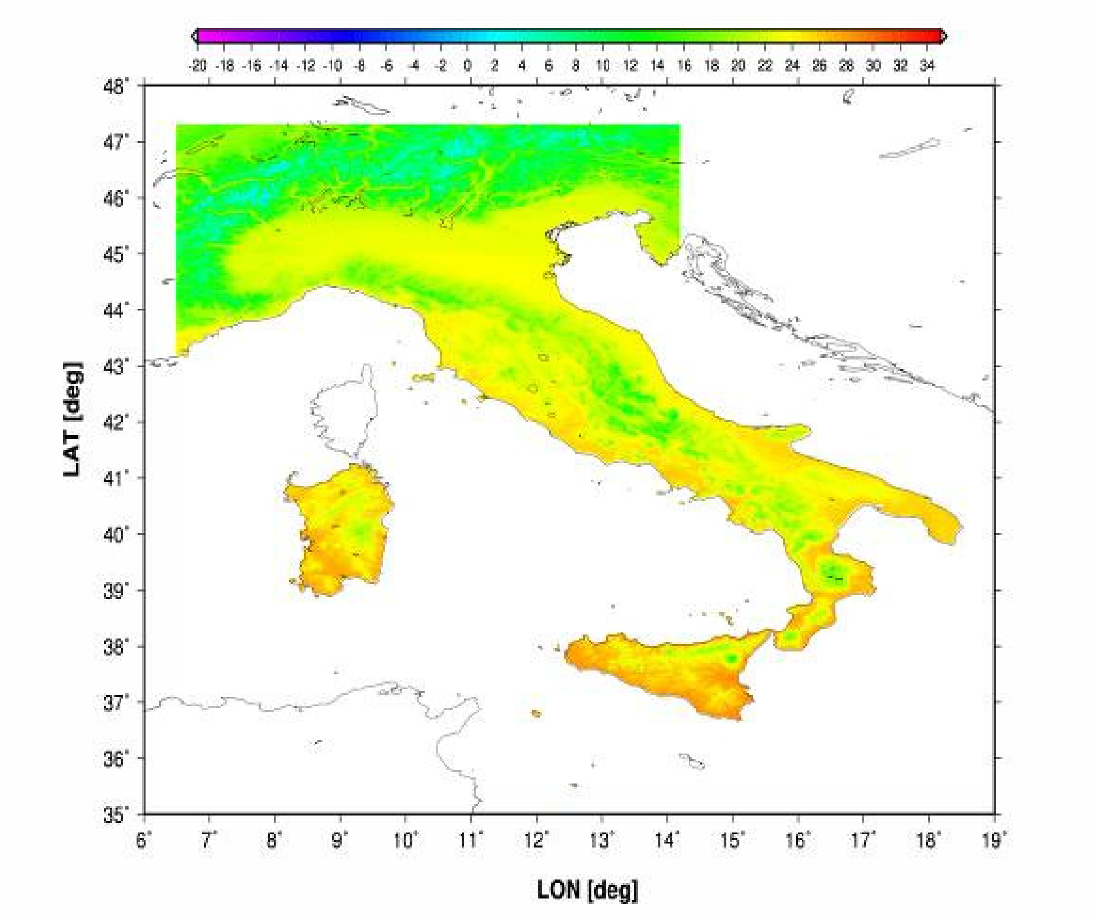

Fig.34 July 1961-90 mean temperature map after MLR model in °C

• Minimum Temperature

We performed monthly *MLR* on 1961-1990 temperature data. Each month was subjected to a different *MLR* that can be written as equation 54:

$$TN_S(LAT_S, LON_S, ELEV_S) = a \cdot LAT_S + b \cdot LON_S + c \cdot ELEV_S + d \quad \textbf{(56)}$$

We did not use all the 1,138 T_N stations because we excluded from the *MLR* a subset made of the stations located at less than 15 km from the sea coasts and the stations at less than 4 km from the lakeshores, thus we used 858 T_N stations.

The *MLR* for the 12 months produced the coefficients shown in table 2.

	a (°C / °LAT)	b (°C / °LON)	c (°C / km)	d (°C)
JAN	-1.16	-0.11	-3.81	53.65
FEB	-0.97	-0.12	-4.65	46.95
MAR	-0.71	-0.08	-5.17	37.75
APR	-0.49	-0.02	-5.50	30.55
MAY	-0.44	0.02	-5.39	31.56
JUN	-0.51	-0.01	-5.38	38.61
JUL	-0.62	-0.06	-5.28	46.39
AUG	-0.68	-0.05	-5.12	48.63
SEP	-0.70	-0.03	-4.85	46.35
OCT	-0.81	-0.03	-4.38	46.72
NOV	-0.98	-0.01	-4.09	49.16
DEC	-1.17	-0.06	-3.71	54.34

Tab.2 Monthly coefficients of the MLR model for minimum temperatures

Then, we used these monthly coefficients to model T_N for Italy:

$$TN_M(LAT_{DEM}, LON_{DEM}, ELEV_{DEM}) = a \cdot LAT_{DEM} + b \cdot LON_{DEM} + c \cdot ELEV_{DEM} + d \qquad \textbf{(57)}$$

• Maximum Temperature

We performed monthly *MLR* on 1961-1990 temperature data. Each month was subjected to a different *MLR* that can be written as equation 54:

$$TX_S(LAT_S, LON_S, ELEV_S) = a \cdot LAT_S + b \cdot LON_S + c \cdot ELEV_S + d \qquad \textbf{(58)}$$

As for T_N, we used only 858 T_X stations, withholding "sea stations" and "lake stations" from calculations of the coefficients.

The *MLR* for the 12 months produced the coefficients shown in table 3.

	a (°C / °LAT)	b (°C / °LON)	c (°C / km)	d (°C)
JAN	-0.97	-0.04	-3.86	51.52
FEB	-0.66	0.01	-4.77	39.72
MAR	-0.42	-0.01	-5.90	33.24
APR	-0.28	0.05	-6.59	30.37
MAY	-0.29	0.15	-6.88	34.60
JUN	-0.41	0.09	-6.99	44.84
JUL	-0.56	0.05	-6.96	55.02
AUG	-0.62	0.12	-6.79	56.15
SEP	-0.53	0.15	-6.30	47.87
OCT	-0.57	0.12	-5.34	44.00
NOV	-0.88	0.06	-4.47	51.49
DEC	-1.06	-0.06	-3.80	56.22

Tab.3 Monthly coefficients of the MLR model for maximum temperatures

Then, we used these monthly coefficients to model T_X for Italy:

$$TN_M(LAT_{DEM}, LON_{DEM}, ELEV_{DEM}) = a \cdot LAT_{DEM} + b \cdot LON_{DEM} + c \cdot ELEV_{DEM} + d \qquad \textbf{(59)}$$

- **Evaluation of the Residuals after *MLR***

We evaluated the statistical parameters *ME*, *MAE*, and *RMSE*. We calculated them by using the residuals of all the stations, including the "Lake stations" and the "Sea stations" i.e.:

$$RESTM_{MLR} = TM_S - TM_M(LON_{DEM}, LAT_{DEM}, ELEV_{DEM}) \qquad \textbf{(60)}$$

$$RESTN_{MLR} = TN_S - TN_M(LON_{DEM}, LAT_{DEM}, ELEV_{DEM}) \qquad \textbf{(61)}$$

$$RESTX_{MLR} = TX_S - TX_M(LON_{DEM}, LAT_{DEM}, ELEV_{DEM}) \qquad \textbf{(62)}$$

Where we compared the 1961-1990 station T_N-T_M-T_X data versus the modelled temperature T_N-T_M-T_X for the corresponding grid cell in the *USGS DEM*. In this case, we used the station elevation from metadata and in the following residual evaluations we used the parameters of the grid cell where the station is located.

Let us show the residual maps for January and July for T_M and let us notice that a positive residual means that the model underestimates the temperature, vice versa if the residual is negative, the model overestimates the temperature.

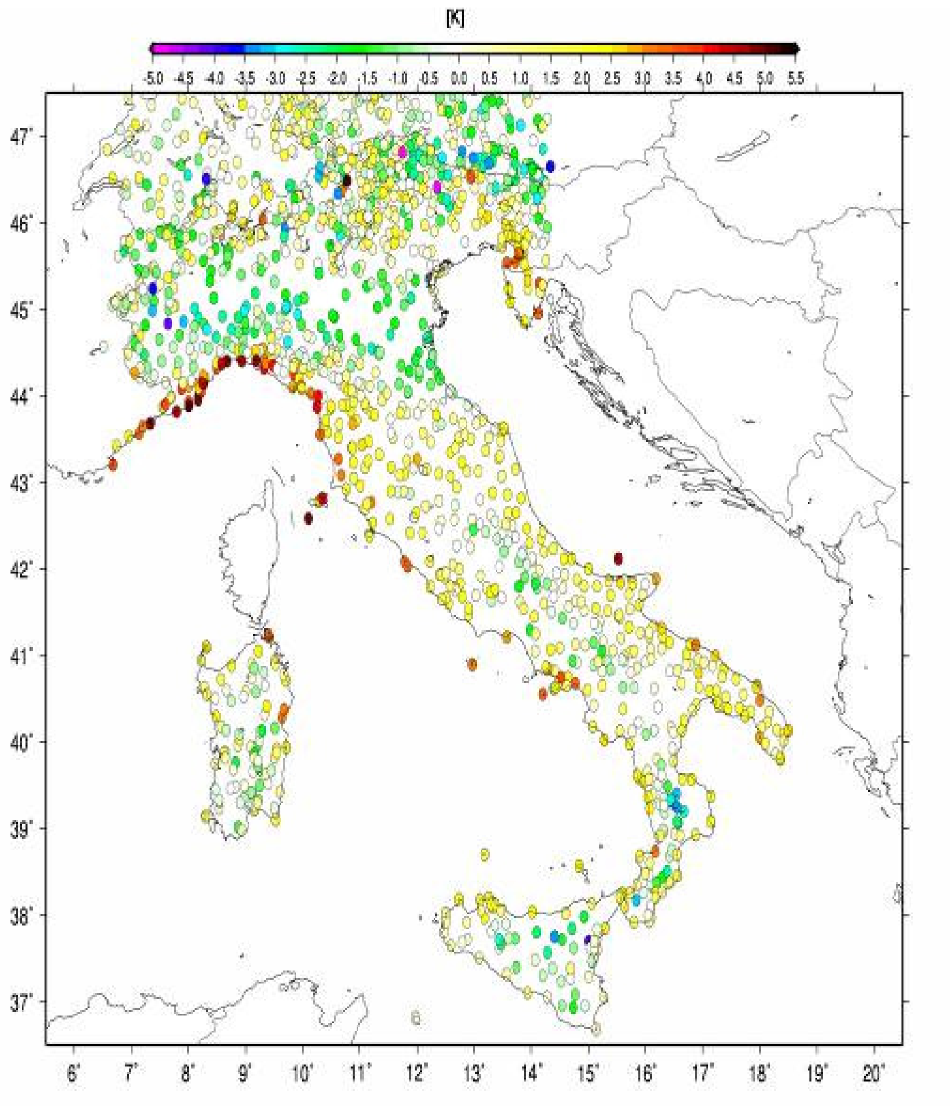

Fig.35 January map of mean temperature residuals after MLR model

For January, many improvements should be introduced. As we can see, the coast areas are generally underestimated, the Po plain is overestimated, the Apennines in southern Italy are overestimated, Sicily and Sardinia are not satisfactorily modelled and so the small isles.

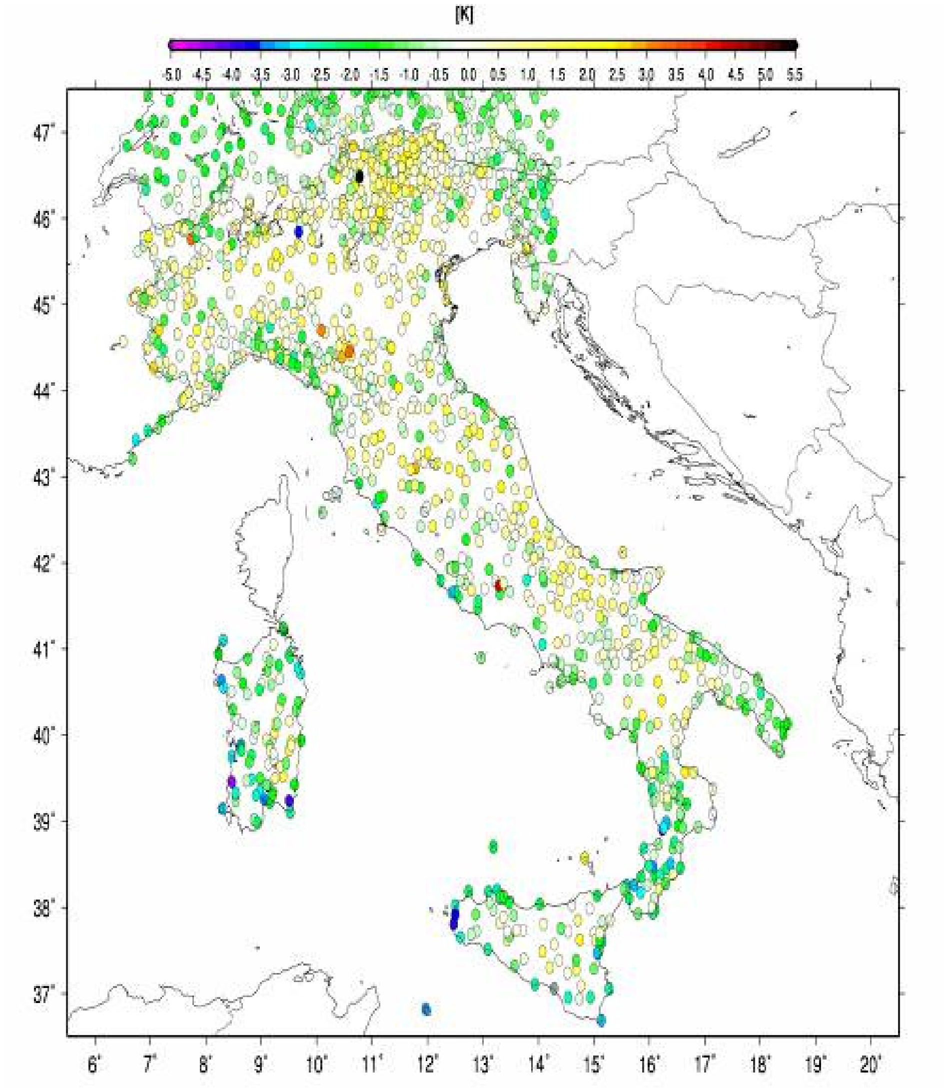

Fig.36 July map of mean temperature residuals after MLR model

Also for July, many improvements should be introduced: the coastal areas are generally overestimated, the Adige Valley is underestimated, over the Alps the model is biased, Sicily and Sardinia are not satisfactorily modelled and so the small isles, Puglia is overestimated. Thus, as we expected, a *MLR* only model is not enough to reproduce local temperature features.

The statistical parameters show errors higher than our goal thresholds; monthly *MAE* and *RMSE* are higher than 1.0 °C. Furthermore, *ME* are higher in winter for T_N and *ME* are higher in summer because the model was calculated withholding a subset and this bias reflects the opposite sea effect in winter and in summer.

Nevertheless, using only a three variable *MLR* leads to an average *MAE* lower than 1.0 °C and *MAE* is lower than 1.0 °C in 10 out of 12 months.

T_N	ME	MAE	RMSE	T_M	ME	MAE	RMSE	T_X	ME	MAE	RMSE
JAN	0.38	1.43	1.83	JAN	0.36	1.25	1.57	JAN	0.33	1.37	1.72
FEB	0.27	1.25	1.61	FEB	0.22	0.96	1.22	FEB	0.18	1.15	1.45
MAR	0.16	1.07	1.37	MAR	0.05	0.77	0.98	MAR	-0.05	0.97	1.28
APR	0.09	1.01	1.29	APR	-0.06	0.76	0.95	APR	-0.22	1.03	1.32
MAY	0.05	1.04	1.31	MAY	-0.14	0.79	0.98	MAY	-0.38	1.15	1.48
JUN	0.05	1.12	1.41	JUN	-0.19	0.88	1.10	JUN	-0.47	1.29	1.67
JUL	0.06	1.24	1.55	JUL	-0.21	0.97	1.22	JUL	-0.55	1.41	1.83
AUG	0.12	1.22	1.53	AUG	-0.16	0.91	1.13	AUG	-0.49	1.35	1.73
SEP	0.19	1.15	1.44	SEP	-0.01	0.77	0.97	SEP	-0.26	1.13	1.43
OCT	0.27	1.14	1.47	OCT	0.16	0.81	1.03	OCT	0.02	1.03	1.31
NOV	0.33	1.18	1.54	NOV	0.29	0.97	1.24	NOV	0.24	1.11	1.41
DEC	0.38	1.38	1.77	DEC	0.38	1.26	1.59	DEC	0.37	1.37	1.71
AVG	0.19	1.19	1.51	AVG	0.06	0.92	1.17	AVG	-0.11	1.20	1.53

Tab.4 Monthly statistical error values in °C after the MLR model de-trendings.

4.3.3 Step 2 (Local improvement): sea effect

+ Distance from the coast grid: SeaDist Grid from *USGS GTOPO30 DEM*

We wrote a Fortran code that calculates, in kilometres, the distance of every grid cell from the nearest sea. Such raster was calculated by means of a weighted product between the straight line distance between the grid cell and the nearest "sea" grid cell and a mathematical formula that includes the orographic obstacles (mountains, ridges, hills), which the described straight line encounters from the grid cell to the "sea" grid cell. This sea distance parameterization first appeared in *Brunetti et al. (2009)*.

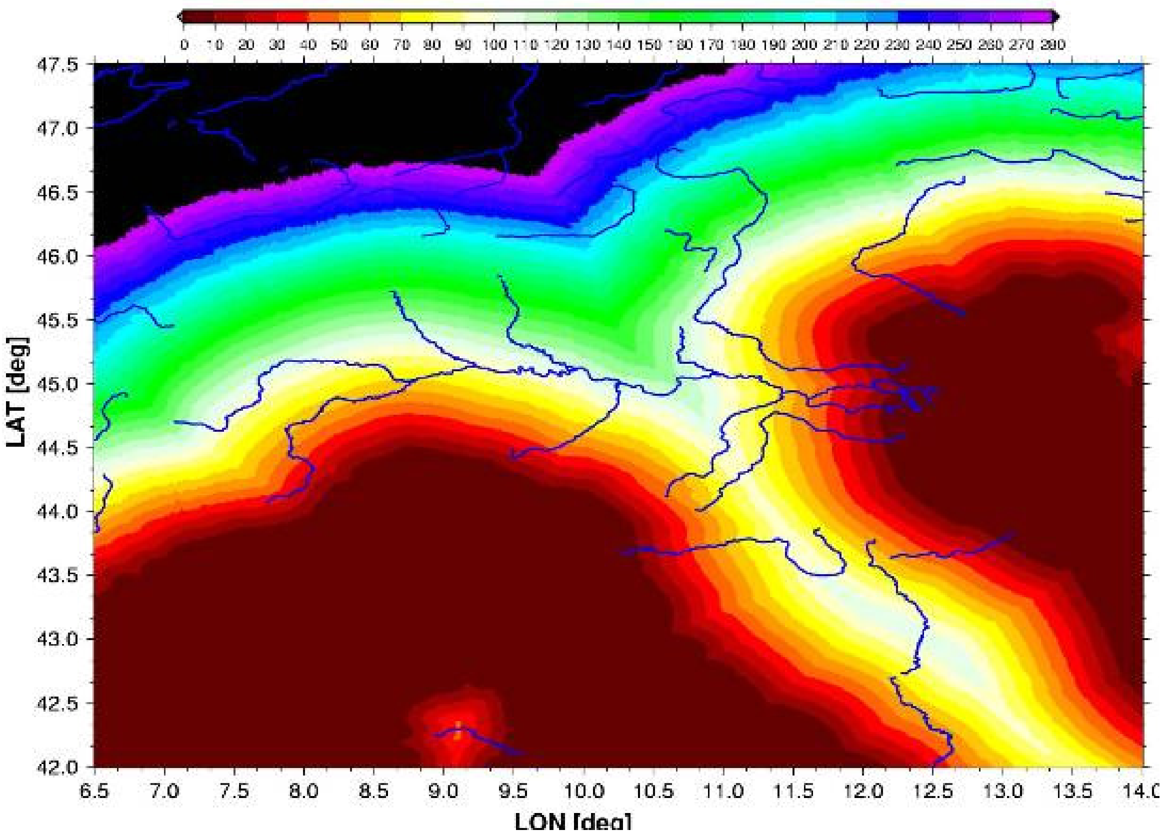

Fig.37 Northern and Central Italy Sea-distance raster (km)

• Mean Temperature

First, we assigned to each station its distance from the sea from the sea distance grid. Second, we divided the "sea" zone in three sub regions. The stations located at less than 15 km from the sea on the Italian Peninsula were labelled as "Peninsula stations", the stations in Sicily or Sardinia located at less than 10 km from the sea were labelled as "Island stations", the stations on the small isles (i.e. Elba and Arcipelago Toscano, Ponza, Capri and Ischia, Tremiti Isles, Lampedusa, Pantelleria, the Eolie Isles, the Egadi Isles, Ustica, Vulcano and Stromboli and so on) were labelled as "Isle stations". The model was thus used only by using 295 T_M stations.

The 10 km and 15 km thresholds were decided after a first evaluation of the *MLR* residuals versus the sea distance: the sea effect is not felt anymore over these thresholds, as we can see in fig. 38.

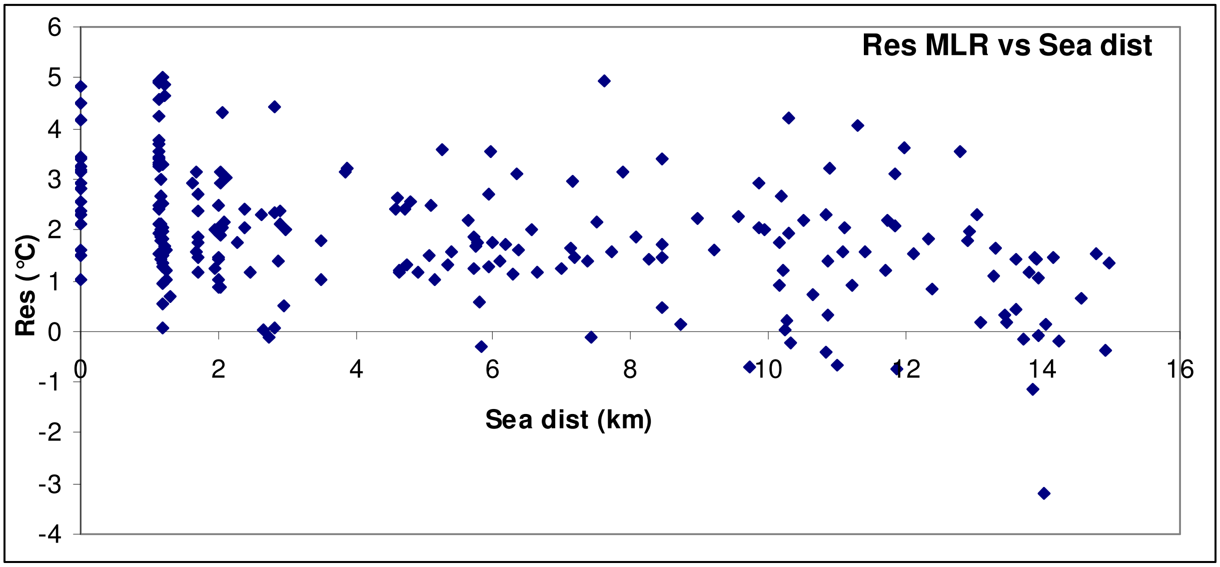

Fig.38 January T_M sea effect for Peninsula stations vanishes over approximately 15 km

We chose a logarithmic model for evaluating the sea effect for Peninsula region:

$$SEA_M(SEADIST_{DEM}) = a \cdot \ln(SEADIST_{DEM}) + b \quad for \quad 1km \leq (SEADIST_{DEM}) \leq 15km \qquad \textbf{(63)}$$

$$SEA_M(SEADIST_{DEM}) = b \quad for \quad (SEADIST_{DEM}) < 1km \qquad \textbf{(64)}$$

And a similar model for Island region:

$$SEA_M(SEADIST_{DEM}) = a \cdot \ln(SEADIST_{DEM}) + b \quad for \quad 1km \leq (SEADIST_{DEM}) \leq 10km \qquad \textbf{(65)}$$

$$SEA_M(SEADIST_{DEM}) = b \quad for \quad (SEADIST_{DEM}) < 1km \qquad \textbf{(66)}$$

We applied such a model to the gridded sea distance raster and we obtained the monthly *a, b* coefficients in this way:

$$b = \overline{RES(1km)}_{MLR} \cdot \overline{\left(\frac{RES(2km)_{MLR}}{RES(1km)_{MLR}}\right)}_{year} \qquad \textbf{(67)}$$

That is we calculated the monthly average residual after *MLR* for the sea stations in the 1 km belt. Then we calculated the yearly ratio between the average residual after *MLR* for the sea stations in the 2 km belt and the average residual after *MLR* for the sea stations in the 1 km belt. We did not calculate *b* only considering stations in the first kilometre belt because the number of the stations in the first km was too low and it would have provided an unrealistic evaluation of *b*.

On the other hand, *a* was obtained by imposing, for Peninsula areas:

$$SEA_M = 0 \qquad for \qquad (SEADIST_{DEM}) = 15km \tag{68}$$

And by imposing, for Island areas:

$$SEA_M = 0 \qquad for \qquad (SEADIST_{DEM}) = 10km \tag{69}$$

The coefficients found for the Peninsula region (left), and for the Island region (right) are:

(°C)	a	b			a	b
JAN	-1.04	2.81		JAN	-0.67	1.55
FEB	-0.71	1.94		FEB	-0.41	0.94
MAR	-0.30	0.82		MAR	0.03	-0.08
APR	-0.08	0.22		APR	0.29	-0.66
MAY	0.17	-0.47		MAY	0.56	-1.30
JUN	0.22	-0.60		JUN	0.77	-1.77
JUL	0.30	-0.82		JUL	0.98	-2.26
AUG	0.23	-0.63		AUG	0.77	-1.77
SEP	-0.04	0.11		SEP	0.23	-0.54
OCT	-0.33	0.91		OCT	-0.30	0.69
NOV	-0.84	2.28		NOV	-0.57	1.31
DEC	-1.04	2.81		DEC	-0.65	1.49

Tab.5 Monthly coefficients for T_M sea effect for Italian Peninsula (left) and for Sicily and Sardinia (right)

Whilst for Isles we used a simpler model:

$$SEA_M = c \tag{70}$$

Where *c* is calculated as the 75% of the monthly residual of the station (if more than one, the averaged residuals) on the considered small isle. Thus, we used 7 different *c* values for each month, depending on the 7 group of isles: Tremiti, Pantelleria, Ponza, Elba with Arcipelago Toscano (3 stations), Lampedusa, Pantelleria, Ustica with Isles North to Sicily.

The coefficients found for the Isles are:

c (℃)	Elba	Tremiti	Ponza	Capri	Ustica	Pantelleria	Lampedusa
JAN	2.94	3.43	2.35	2.44	1.25	0.04	0.01
FEB	2.05	2.56	1.51	1.72	0.70	-0.06	-0.05
MAR	0.82	1.93	0.45	0.81	-0.03	-0.55	-0.55
APR	0.02	1.01	-0.24	0.42	-0.54	-0.64	-0.95
MAY	-0.40	0.81	-0.67	0.17	-1.03	-1.17	-1.88
JUN	-0.62	0.50	-0.73	-0.15	-1.29	-1.89	-2.70
JUL	-0.32	0.76	-0.73	-0.22	-1.44	-2.52	-3.47
AUG	-0.08	1.23	-0.38	0.25	-0.99	-2.09	-2.60
SEP	0.44	1.33	0.32	0.99	0.02	-0.83	-0.77
OCT	1.57	2.45	1.12	1.47	0.74	-0.00	0.42
NOV	2.56	3.21	1.95	2.06	1.21	0.26	0.64
DEC	3.00	3.96	2.48	2.26	1.31	-0.06	0.13

Tab.6 Monthly coefficients for T_M sea effect for seven groups of Italian isles

Let us show the T_M sea effect in January and in July:

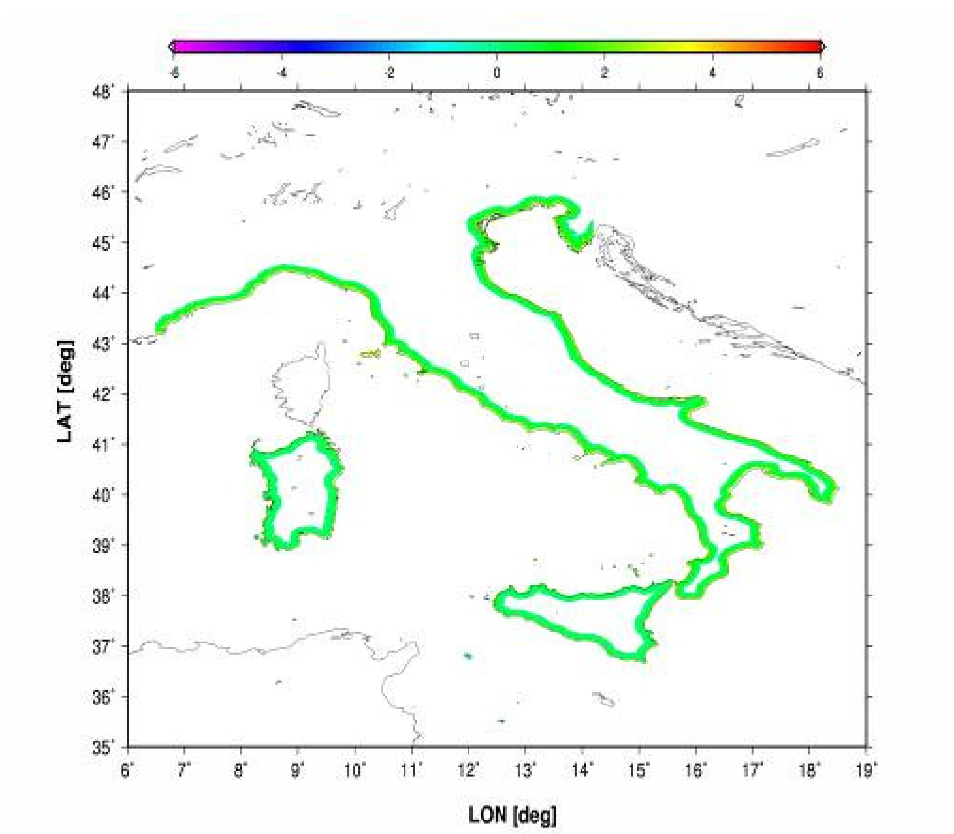

Fig.39 Overall sea (warming) effect in January for T_M in ℃

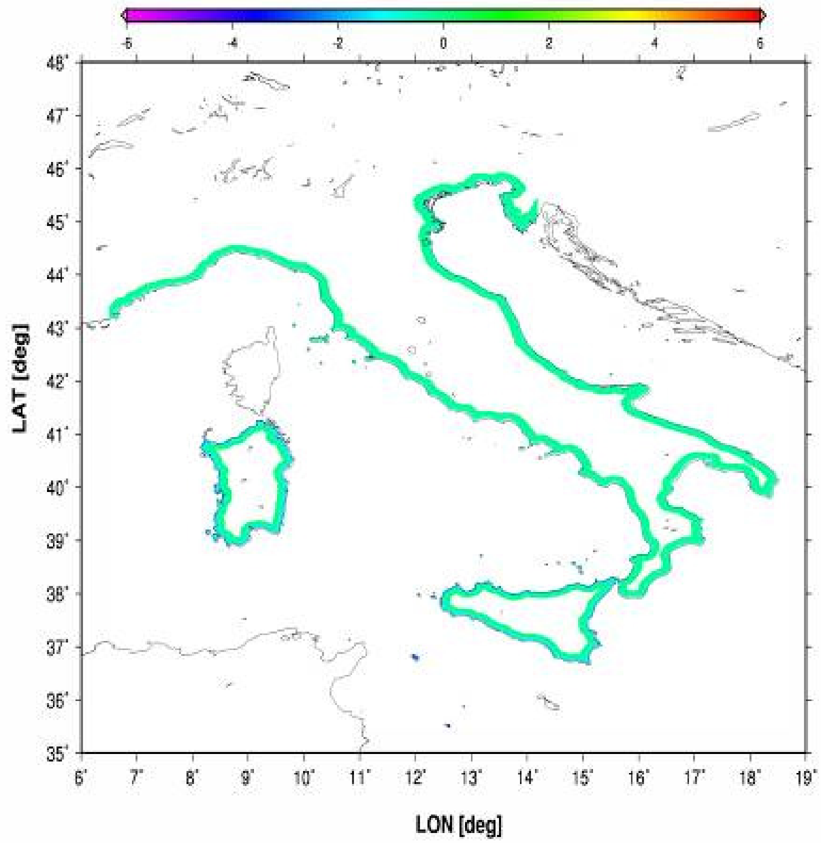

Fig.40 Overall sea (cooling) effect in July for T_M in °C

As expected, the sea causes a warming effect in January (the effect is stronger in Peninsula than in Great Islands, while in the Isles it is strongest) and oppositely a cooling effect in July (in this case the effect is stronger in Great Islands than in Peninsula, but it is once again strongest in the Isle). The maximum modelled warming effect can be found in Tremiti Isles in December (3.96 °C). The maximum modelled cooling effect, in absolute values, can be found in Pantelleria in July (-3.47 °C).

In the end, we added back the sea effects to the modelled *MLR* T_M as:

$$TM2_M = TM_M(LAT_{DEM}, LON_{DEM}, ELEV_{DEM}, SEADIST_{DEM}) = TM1_M + SEA2_M \quad \textbf{(71)}$$

If we define:

$$TM_M(LAT_{DEM}, LON_{DEM}, ELEV_{DEM}) = TM1_M \quad \textbf{(72)}$$

$$SEA_M(SEADIST_{DEM}) = SEA2_M \quad \textbf{(73)}$$

• Minimum Temperature

For T_N we used the same methodology used for T_M, the only differences lie in the number of stations (i.e., 238 stations) for calculating the coefficients and the coefficients themselves.

The coefficients found for the peninsula region (left), and for the island region (right) are:

(°C)	a	b			a	b
JAN	-0.61	1.41		JAN	-1.19	3.22
FEB	-0.39	0.91		FEB	-0.96	2.61
MAR	-0.23	0.54		MAR	-0.70	1.90
APR	-0.14	0.33		APR	-0.58	1.56
MAY	-0.02	0.05		MAY	-0.52	1.41
JUN	0.01	-0.03		JUN	-0.59	1.59
JUL	0.07	-0.18		JUL	-0.61	1.65
AUG	-0.10	0.24		AUG	-0.66	1.79
SEP	-0.38	0.88		SEP	-0.71	1.93
OCT	-0.57	1.32		OCT	-0.77	2.10
NOV	-0.64	1.47		NOV	-0.96	2.60
DEC	-0.62	1.43		DEC	-1.11	3.01

Tab.7 Monthly coefficients for T_N sea effect for Italian peninsula (left) and for Sicily and Sardinia (right)

The coefficients found for the isles are:

c (°C)	Elba	Tremiti	Ponza	Capri	Ustica	Pantelleria	Lampedusa
JAN	3.88	4.49	3.53	2.73	2.17	0.66	0.88
FEB	3.09	3.77	2.85	2.01	1.69	0.45	0.88
MAR	2.21	3.18	2.08	1.30	1.35	0.43	0.95
APR	1.72	2.30	1.50	1.05	1.11	0.51	0.97
MAY	1.38	2.11	1.22	0.95	0.77	0.15	0.52
JUN	1.30	1.91	1.20	0.78	0.71	-0.27	0.07
JUL	1.66	2.38	1.39	0.82	0.85	-0.45	-0.21
AUG	1.90	2.80	1.75	1.26	1.34	-0.06	0.61
SEP	2.09	2.96	2.33	1.74	1.97	0.70	1.69
OCT	2.93	3.94	2.88	2.02	2.26	1.12	2.08
NOV	3.70	4.38	3.32	2.49	2.46	1.20	2.04
DEC	4.01	4.88	3.61	2.49	2.27	0.68	1.20

Tab.8 Monthly coefficients for T_N sea effect for seven groups of Italian isles

On T_N, the sea causes a warming effect in all months in the great islands, whilst in the peninsula, it tends to be a cooling one in summer. For the small isles, with rare exceptions, the modelled effect is a warming effect. The maximum warming effect is in December for Elba (4.01 °C).

• Maximum Temperature

For T_X hold the same considerations made for T_N.

The coefficients found for the peninsula region (left), and for the island region (right) are:

(°C)	a	b			a	b
JAN	-0.58	1.58		JAN	-0.60	1.40
FEB	-0.20	0.56		FEB	-0.47	1.08
MAR	0.29	-0.78		MAR	0.07	-0.16
APR	0.60	-1.62		APR	0.49	-1.13
MAY	1.07	-2.89		MAY	0.94	-2.18
JUN	1.22	-3.32		JUN	1.28	-2.95
JUL	1.46	-3.97		JUL	1.65	-3.81
AUG	1.39	-3.78		AUG	1.46	-3.38
SEP	0.93	-2.53		SEP	0.70	-1.62
OCT	0.41	-1.12		OCT	-0.07	0.16
NOV	-0.44	1.19		NOV	-0.44	1.01
DEC	-0.64	1.74		DEC	-0.57	1.32

Tab.9 Monthly coefficients for T_X sea effect for Italian peninsula (left) and for Sicily and Sardinia (right)

The coefficients found for the isles are:

c (°C)	Elba	Tremiti	Ponza	Capri	Ustica	Pantelleria	Lampedusa
JAN	2.08	1.88	1.02	1.89	0.28	-0.39	-0.68
FEB	1.07	0.90	0.09	1.25	-0.20	-0.20	-0.56
MAR	-0.55	0.28	-1.21	0.22	-1.26	-1.05	-1.48
APR	-1.69	-0.68	-2.04	-0.31	-2.03	-1.32	-2.30
MAY	-2.23	-0.93	-2.68	-0.79	-2.74	-2.11	-3.82
JUN	-2.60	-1.37	-2.77	-1.25	-3.17	-3.05	-4.91
JUL	-2.33	-1.45	-2.95	-1.47	-3.54	-3.95	-5.98
AUG	-2.11	-0.94	-2.63	-0.97	-3.15	-3.51	-5.10
SEP	-1.20	-0.92	-1.83	-0.01	-1.82	-1.88	-2.69
OCT	0.28	0.37	-0.75	0.67	-0.70	-0.68	-0.73
NOV	1.50	1.55	0.47	1.41	-0.01	-0.38	-0.44
DEC	2.07	2.50	1.19	1.76	0.29	-0.59	-0.71

Tab.10 Monthly coefficients for T_X sea effect for seven groups of Italian isles

On T_X, the sea causes a warming effect almost everywhere in winter and a cooling one in summer. The maximum cooling effect, in absolute values, is just on the coast area in July for the peninsula (-3.97 °C); the maximum warming effect is in December for the Tremiti Isles (2.51 °C).

- **Improvements and evaluation of the Residuals after sea effect**

In order to compare the modelled climatologies before the modelling of the sea effect (see fig. 33 and fig. 34) and after the sea effect improvements, let us show the modelled T_M for January and July after *MLR* and sea effects (see fig. 41 and fig. 42).

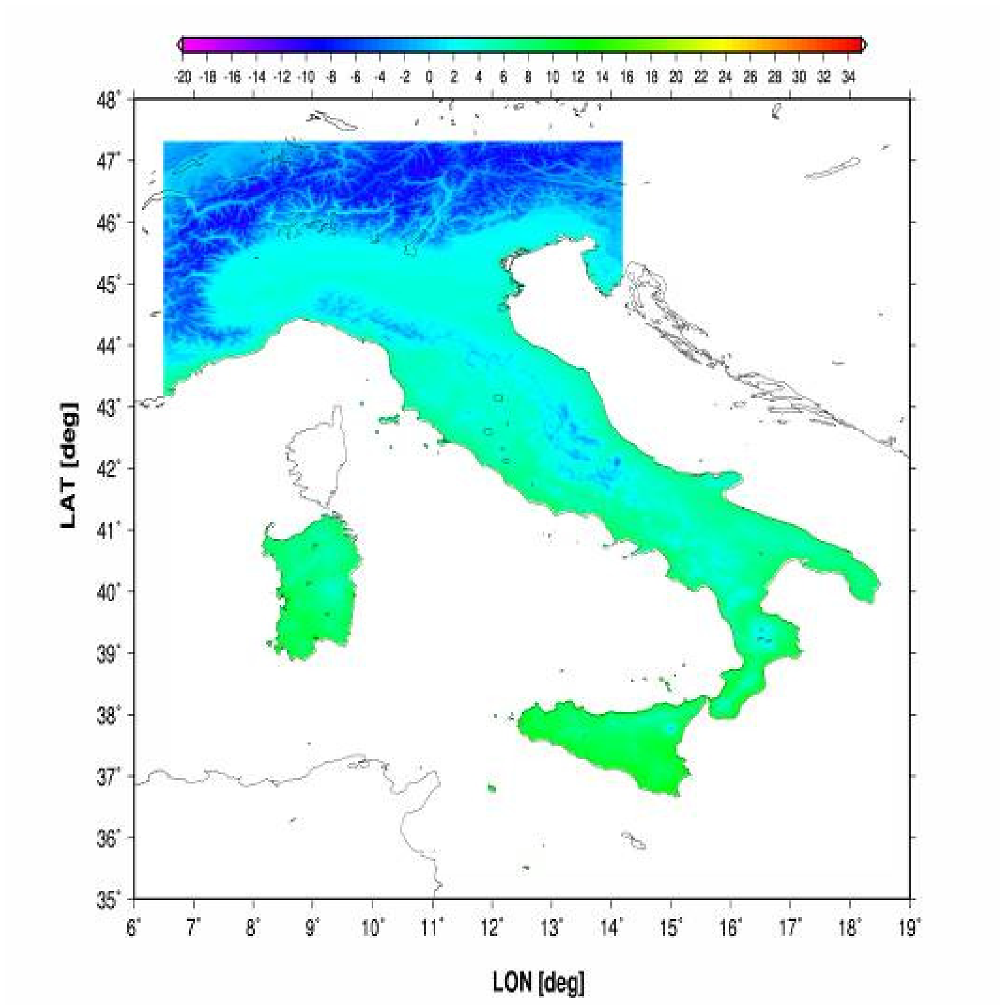

Fig.41 January 1961-90 mean temperature map after MLR model plus sea effect improvements (°C)

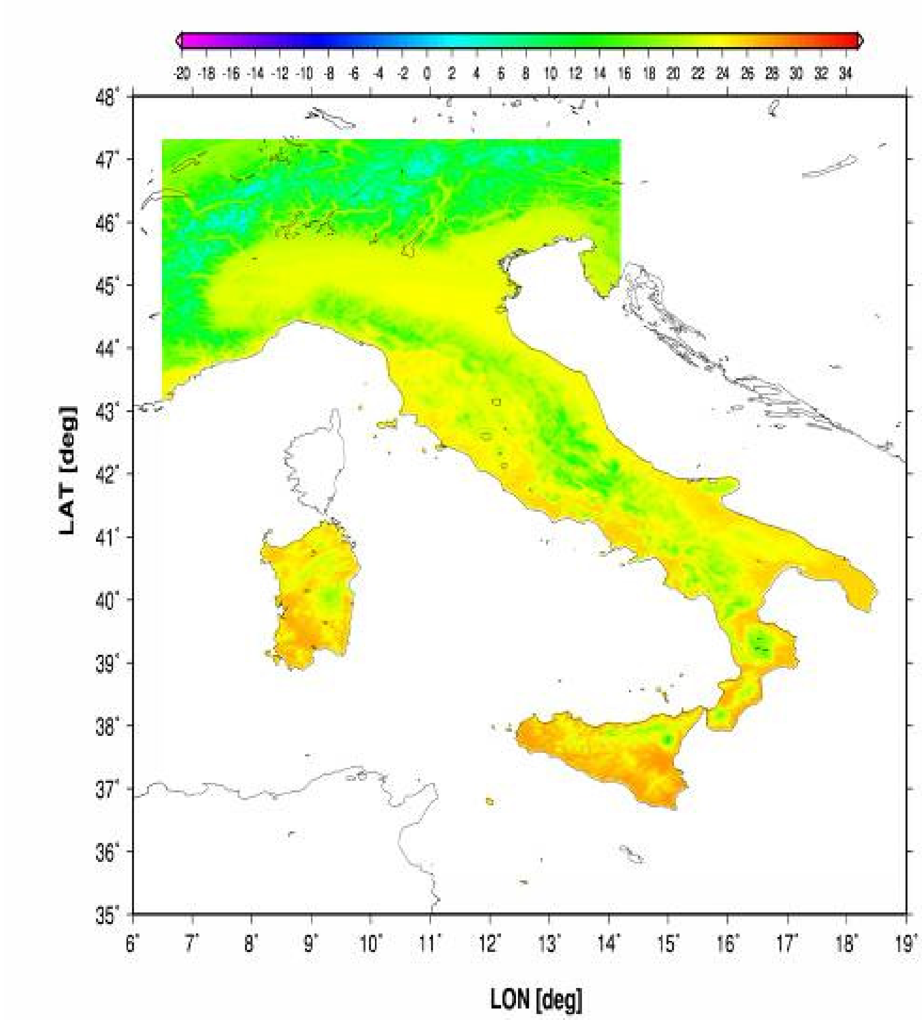

Fig.42 July 1961-90 mean temperature map after MLR model plus sea effect improvements (°C)

We evaluated the statistical parameters by means of *ME*, *MAE* and *RMSE* calculated using the residuals of all the stations, including the stations not labelled by "sea" labels, i.e.:

$$RESTM_{MLR+SEA} = RESTM_{MLR} - SEA_M\,(SEADIST_{DEM}) \qquad \textbf{(74)}$$

Where the modelled sea effect were subtracted from the residuals after the *MLR*.

We showed the formula for T_M, the same holds for T_N and T_X.

The statistical parameters show meaningful improvements, as it can be seen in the tab. 11 shown in the next page.

	TM	ME	MAE	RMSE		ME	MAE	RMSE	TX	ME	MAE	RMSE
	JAN	0.05	1.29	1.62		0.09	1.08	1.35	JAN	0.16	1.24	1.58
	FEB	0.00	1.17	1.49		0.04	0.87	1.11	FEB	0.11	1.10	1.41
	MAR	-0.03	1.03	1.31		-0.02	0.75	0.96	MAR	0.02	0.97	1.28
	APR	-0.06	1.00	1.27		-0.06	0.75	0.95	APR	-0.04	0.95	1.24
	MAY	-0.08	1.03	1.31		-0.07	0.75	0.94	MAY	-0.06	0.98	1.27
	JUN	-0.09	1.12	1.42	JUN	-0.09	0.82	1.03	JUN	-0.10	1.06	1.36
	JUL	-0.08	1.23	1.54	JUL	-0.08	0.90	1.12	JUL	-0.09	1.14	1.46
	AUG	-0.05	1.19	1.49	AUG	-0.06	0.86	1.07	AUG	-0.07	1.13	1.43
	SEP	-0.01	1.10	1.37	SEP	-0.01	0.76	0.96	SEP	0.00	1.06	1.35
	OCT	0.04	1.05	1.32	OCT	0.07	0.76	0.97	OCT	0.11	1.05	1.34
	NOV	0.05	1.07	1.34	NOV	0.07	0.84	1.06	NOV	0.11	1.03	1.33
	DEC	0.07	1.24	1.55	DEC	0.11	1.08	1.35	DEC	0.18	1.22	1.55
MLR+SEA	AVG	-0.02	1.13	1.42	AVG	0.00	0.85	1.07	AVG	0.03	1.08	1.38
MLR	AVG	0.19	1.19	1.51	AVG	0.06	0.92	1.17	AVG	-0.11	1.20	1.53
IMPROVEMENT	AVG	0.17	0.06	0.09	AVG	0.06	0.07	0.10	AVG	0.08	0.12	0.15

Tab.11 Monthly statistical error values after MLR model plus sea effect de-trendings and comparisons.

The introduction of the sea effect in the models partially removes the bias caused by the calculation of the *MLR* models without considering the sea and the lake stations, and it reduces the *MAE* for any variable. The reductions of *MAE* are consistent: 0.06 °C for T_N, 0.07 °C for T_M, 0.12 °C for T_X and even bigger reductions for the *RMSE*.

After the *MLR* plus sea effect, the average *MAE* for T_M is 0.92 °C and 5 out of 12 months show a *RMSE* lower than the 1.0 °C threshold.

In *Hiebl et al. (2009),* the *MLR* used the distance from the sea (calculated with a different shape weighting function) as the fourth independent variable, whilst the sea effects are usually de-trended singularly (e.g. *Goodale et al., 1998*).

4.3.4 Step 3 (Local improvement): lake effect

• Distance from the lake grid: Lake grid from *GLC2000* Land Cover

The Global Land Cover 2000 project, by the Institute of Environment and Sustainability of the Joint Research Center in Ispra, Italy, is a global hi-res (approximately

1 km² at Italian latitudes) gridded land cover dataset projected in *WGS84* coordinates *(Belward et al., 2002)*. The *GLC2000* is subdivided into many land cover classes, as we can see in the European region map showed in fig. 43.

Fig.43 GLC 2000 land cover map (JRC GEM website)

The *GLC2000* has 23 classes for European area: 10 different tree cover classes, 3 shrub classes, 2 herbaceous classes, 3 classes related to cultivated terrains, 2 crop classes, 1 bare area class, 1 water bodies class, 1 snow and ice class, 1 artificial surface class.

We preferred the *GEM GLC2000* land cover because metadata and methodology are physically stronger than other land cover gridded dataset ad *CORINE* (*CORINE report, 1994*, *http://www.eea.europa.eu/publications/COR0-landcover)* realised by the European

Environment Agency or the *PELCOM* (Pan-European Land Use and Land Cover Monitoring *(Champeaux et al., 2000)*. Even though the *GLC2000* is related to 2000 and we deal with 1961-1990, we compared *PELCOM* (related to 1993) and found no significant differences between the two land covers with some exceptions for artificial areas. However, the land cover grids are difficult to be compared because of different land cover classes.

With the help of a Fortran code and a nearest neighbour technique, we converted the *GLC2000* grid to match the *USGS GTOPO30* grid coordinates and we created a "lake grid", using the land cover class "water bodies" and introducing some geographical borders, in order to isolate Leman Lake, the northern Italian lakes, and three central Italy lakes (Bolsena, Bracciano, Trasimeno). Then we wrote a dedicated Fortran code to label with "1" the grid points at less than 2 km from the cited lakes, with "2" the grid cells between 2 km and 4 km and with "0" the other grid points.

• Mean Temperature

First, we selected a subset of stations located at less than 5 km from the cited lakes. We used the Google Earth tool to calculate the distances from the lakeshores. Then we analyzed their residuals and we found that the 3 lakes in Central Italy (Bolsena, Trasimeno and Bracciano) cause no significant effects, whilst the northern lakes behave like the sea, thus causing a warming effect in winter and a cooling effect in summer.

We got 20 T_M "lake stations" for the northern Italy lakes and the Leman Lake. We averaged their monthly residuals and we used these averages as the lake effect modelled in the first 2 km belt from the lakeshores. Then we used the half of these values in the belt between 2 km and 4 km. That is:

$$LAKE_M(LAKEGRID_{DEM}) = \left(\overline{RESTM_{MLR+SEA}}\right)_{LAKESTATIONS} \quad for \quad LAKEGRID_{DEM} \leq 2km \qquad \textbf{(75)}$$

$$LAKE_M(LAKEGRID_{DEM}) = \frac{1}{2}\left(\overline{RESTM_{MLR+SEA}}\right)_{LAKESTATIONS} \quad for \ 2km < LAKEGRID_{DEM} \leq 4km \qquad \textbf{(76)}$$

Where the subscript "*LAKESTATIONS*" refers to 18 T_M "lake stations", because we withheld Peschiera and Desenzano which felt also the Po Plain cold pool effect (see next pages), thus we calculated the modelled coefficients using 18 stations.

Here we are the coefficients used for the lake effect on T_M:

(°C)	0-2km	2-4km
JAN	1.30	0.65
FEB	0.79	0.39
MAR	0.40	0.20
APR	0.25	0.12
MAY	0.12	0.06
JUN	0.12	0.06
JUL	0.06	0.03
AUG	0.02	0.01
SEP	0.11	0.05
OCT	0.38	0.19
NOV	0.85	0.42
DEC	1.34	0.67

Tab.11 Monthly coefficients for T_M lake effect for Leman lake and northern Italy lakes

As we can see, lakes cause a warming effect in winter and a null effect in summer for T_M. Because of the smaller water surface and volume, the lake effect is lower than the corresponding sea effect, yet similar.

In the end, we added back the lake effects to the modelled *MLR* + Sea Effect T_M as:

$$TM3_M = TM2_M + LAKE_M \quad \textbf{(77)}$$

This method is valid for every step.

• Minimum Temperature

We employed the same procedure used for mean temperatures, but we used only 15 T_N stations to calculate the coefficients.

Here they are the coefficients used for the <u>lake effect</u> on T_N:

(°C)	0-2km	2-4km
JAN	2.20	1.10
FEB	1.82	0.91
MAR	1.43	0.71
APR	1.25	0.62
MAY	1.19	0.59
JUN	1.24	0.62
JUL	1.25	0.62
AUG	1.16	0.58
SEP	1.19	0.59
OCT	1.32	0.66
NOV	1.56	0.78
DEC	2.05	1.02

Tab.12 Monthly coefficients for T_N lake effect for Leman lake and northern Italy lakes

On minimum temperatures, the lake effect is a warming effect for any month, as we expected.

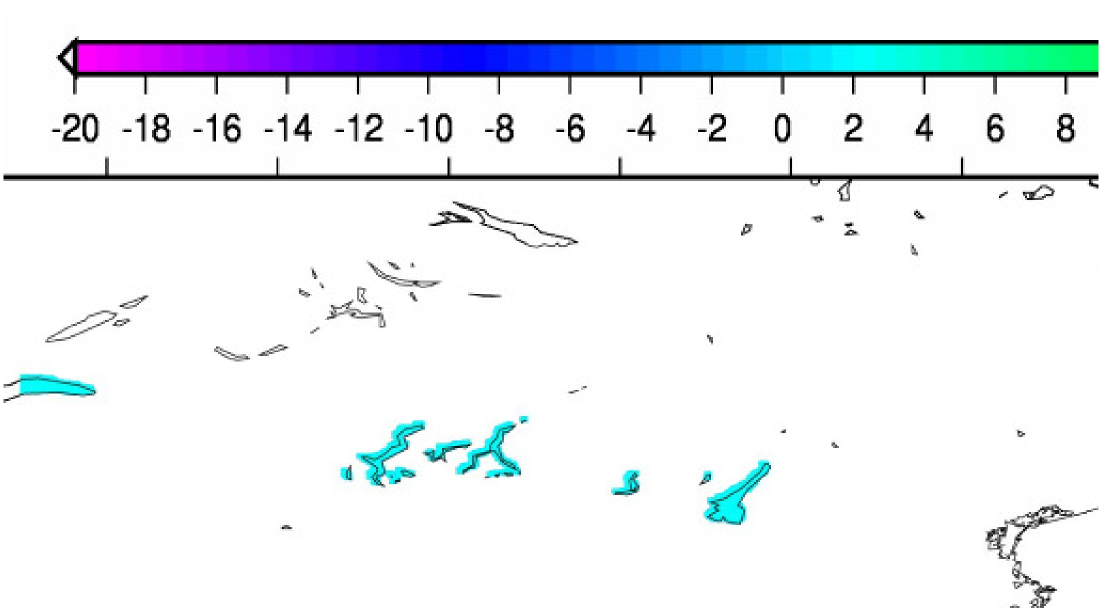

Fig.44 lake effect (°C) for T_N in January, particular of the Italian map.

• Maximum Temperature

The same considerations made for T_N holds for T_X.

Here they are the coefficients used for the lake effect on T_N:

	0-2km	2-4km
JAN	0.49	0.24
FEB	-0.19	-0.09
MAR	-0.48	-0.24
APR	-0.64	-0.32
MAY	-0.87	-0.43
JUN	-0.93	-0.46
JUL	-1.17	-0.58
AUG	-1.20	-0.60
SEP	-0.99	-0.49
OCT	-0.58	-0.29
NOV	0.21	0.10
DEC	0.72	0.36

Tab.13 Monthly coefficients for T_X lake effect for Leman lake and northern Italy lakes

The lake effect is a cooling effect for T_X but in winter, where it is a warming effect.

• Improvements and evaluation of the Residuals after lake effect

Once again, residuals were calculated as:

$$RESTM_{MLR+SEA+LAKE} = RESTM_{MLR+SEA} - LAKE_M(LAKEGRID_{DEM}) \quad \textbf{(78)}$$

The same holds for T_N and T_X.

MLR+SEA+LAKE	AVG	-0.04	1.11	1.41		AVG	-0.01	0.85	1.07		AVG	0.03	1.07	1.38
IMPROVEMENTS	AVG	-0.02	0.02	0.01		AVG	-0.01	0.00	0.00		AVG	0.00	0.01	0.00

Tab.14 Yearly averaged statistical error values after lake effect and comparisons (°C)

The lake effect introduces small improvements, but it is an important refinement because it is rarely modelled in spatial climate studies.