4.3.10 Step 9 (Interpolation of the Residuals): *GIDW*

- ***GIDW*: semi-variograms, radial, elevation and sea weights**

In order to best capture the residual variability still present, we decided to use a sort of inverse distance weighted (*IDW*) model and introduced other geographical weights (thus the name *GIDW*) such as elevation weighted thresholds and sea distance weights.

Here it is the radial Gaussian weight used for residual interpolation:

$$w_i^{rad}(x, y) = e^{-\left(\frac{d_i^2(x,y)}{c}\right)} \quad \textbf{(99)}$$

$$c = -\frac{\overline{d}^2}{\ln(0.5)} \quad \textbf{(100)}$$

Where *(x,y)* are the geographical coordinates of the grid cells and of the stations, *i* is referred to the *i-th* station used for the weighting process which depends on the stations that are within the search distance, *d* is the distance between the grid cell under investigation and the *i-th* stations, while $\overline{d}^2$ is the distance where $w = 0.5$.

To define the elevation weight, we used a moving window that depends on elevation and on the local averaged elevation. The model is not linear when it considers the difference in elevation between the grid cell and stations as a independent variable. The basic hypothesis is that at high elevations the difference in elevation should be less important than at low. We chose the moving threshold after analysis on elevation dependence of the residuals.

The sea distance weight is more than just a binary condition. If a grid cell is in the first 15 km belt from the coast, it would be influenced only by stations located in the 15 km belt (weight equals to 1). If a cell is out of such "sea coast belt" it would be influenced only by stations located out of the "sea coast belt". Nevertheless, we smoothed this yes or no condition in the belt between 30 km and 15 km from the sea's coast, with a linear regression model that makes the weights range from 0 to 1.

We used a Fortran code to calculate the weights and to assign the correct normalized weight to each grid cell:

$$w_i(x,y) = w_i^{rad}(x,y) \cdot w_i^{elev}(x,y) \cdot w_i^{Sea}(x,y) \quad \textbf{(101)}$$

$$IDS_M = \frac{\sum_i (w_i \cdot RES_i)}{\sum_i w_i} \quad \textbf{(102)}$$

where RES_i are the residuals after *MLR* plus the 7 improvements. i.e. from equation (98) and the super-script refer to the listed weights.

• Mean Temperature

First, we studied monthly semi-variograms (see equation (5)) where we plotted semi-variances versus distance, and where we used monthly T_M residuals from equation (98).

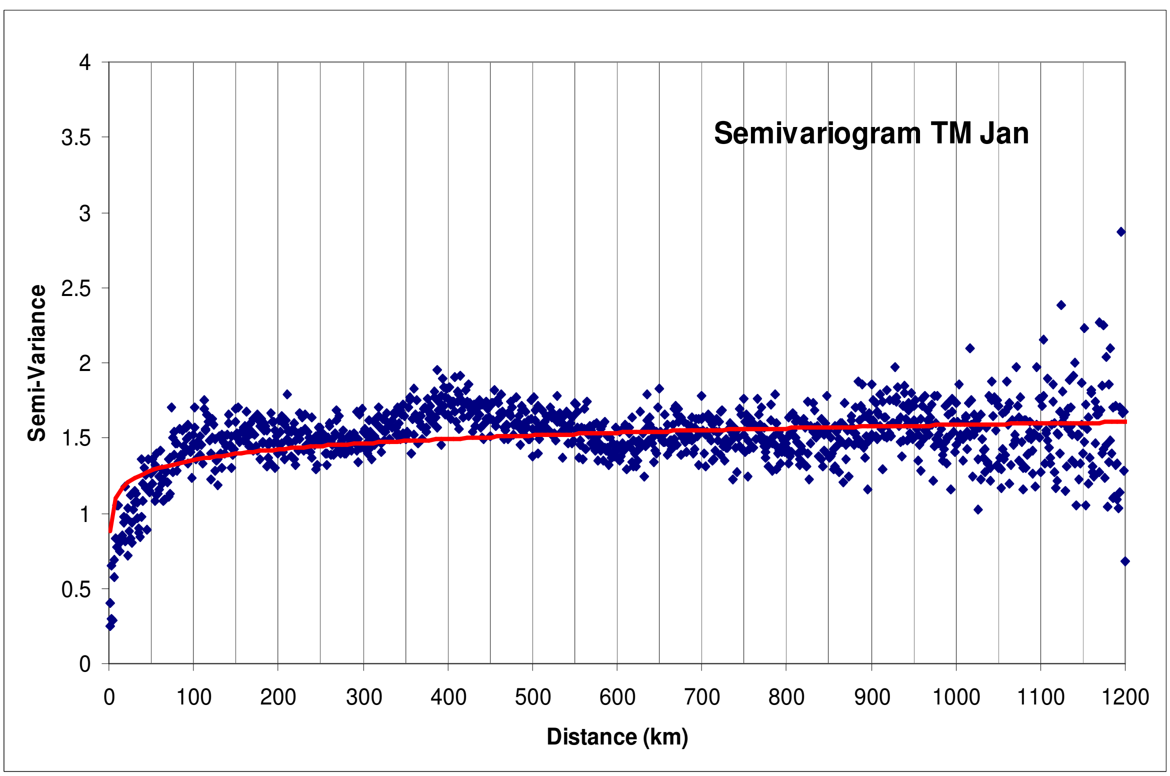


Fig.62 Logarithmic semi-variogram for January T_M residuals

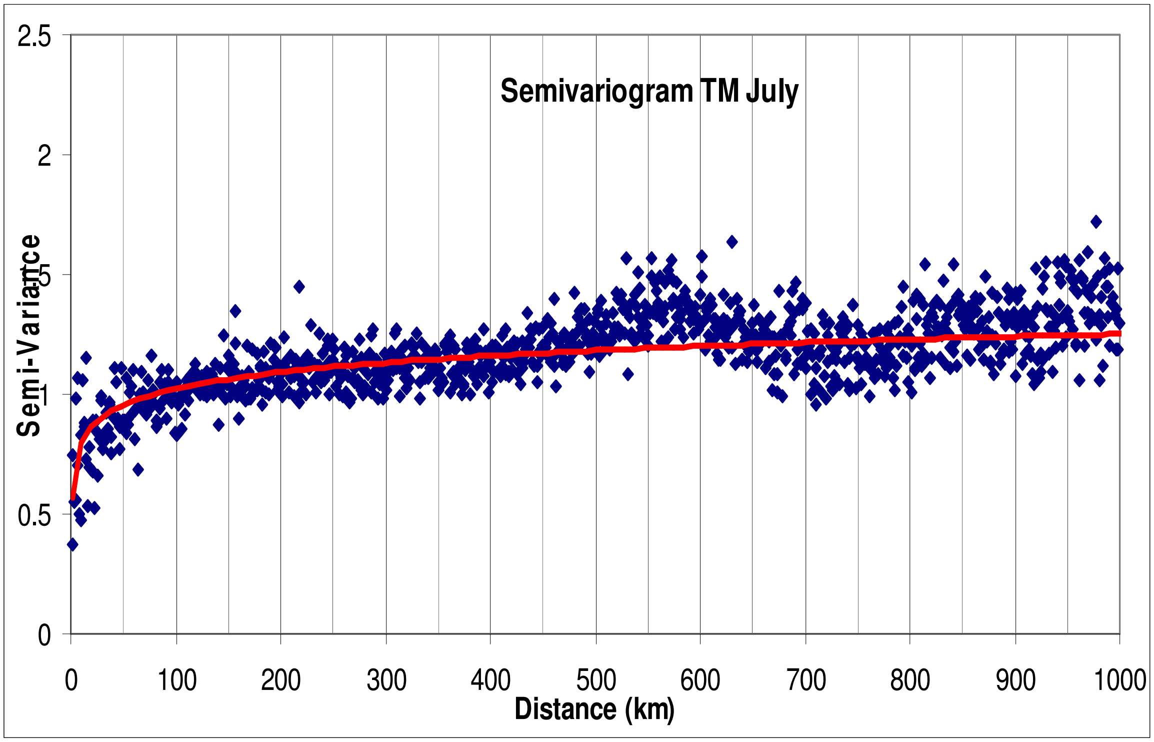


Fig.63 Logarithmic semi-variogram for January T_N residuals

Using a logarithmic semi-variogram, we deduced from monthly plots that the nugget value C_0 can be approximately found around a semi-variance of 0.4, whilst the sill value varies monthly from 150 to 300 km, for a semi-variance of about 1.2. Thus the search radius for stations in the radial weight was set up to 250 km, i.e. every station within a radius of 250 km from a grid cell was used in the weighting process.

Furthermore, we deeply analyzed the semi-variogram and we found that the mean value of semi-variance between the semi-variance corresponding to nugget and sill values is approximately at 33 km, thus we set up 33 km as the threshold distance where the Gaussian weight equals to *0.5*.

In the next page, we show the T_M residual correction for January and July (as a raster we used *USGS GTOPO30 DEM*).

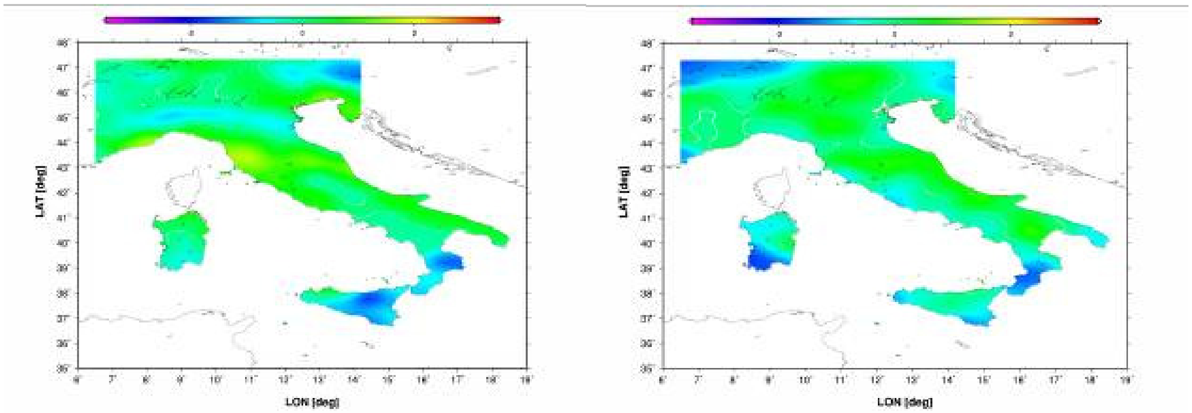


Fig.64-65 Interpolation of the residuals with a GIDW for January (left) and July (right) T_M (°C)

As we can see, in January the interpolation of the residuals mainly depends on the radial weight and such a correction strengthens the Po-Plain effect, marks some regions that were included as improvements such as Apennines (especially in Calabria), the regions north of the Alps. It also shows that the Sicily and Sardinia latitudinal gradients were not yet satisfactorily modelled. Similar considerations can be made for July.

In the end, we added back the residual *GIDSW* effects to the T_M modelled with *MLR* plus local and global improvements, plus the residual interpolation as:

$$TM9_M = TM8_M + IDS_M \qquad \textbf{(103)}$$

• Minimum Temperature

We used the same methodology used for T_M because semi-variograms for T_N (and for T_X) yielded very similar results.

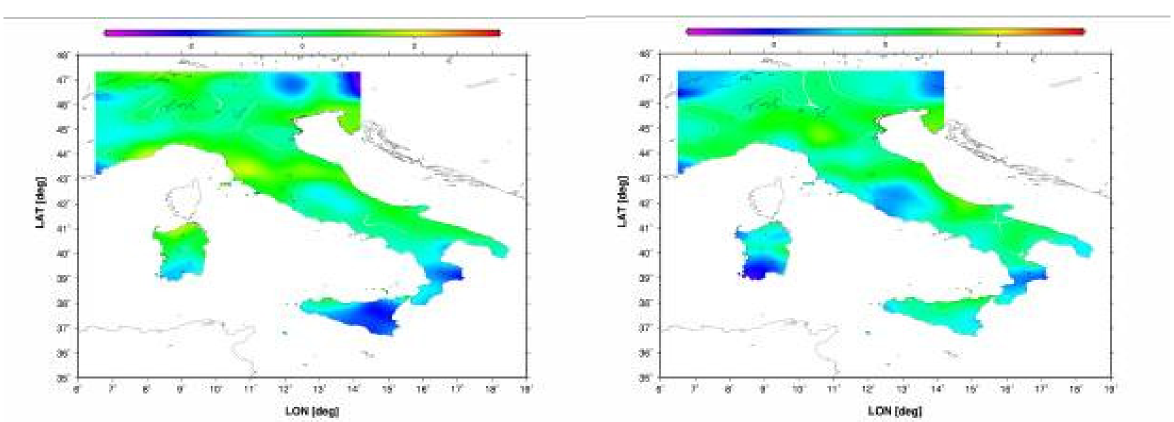


Fig.66-67 Interpolation of the residuals with a GIDW for January and July T_N (°C)

For T_N the residual correction is more important than for T_M because the T_N residuals are higher, in absolute values, especially in Sicily and Trentino in January and in Sardinia and Lazio in July.

• Maximum Temperature

The same considerations made for T_N are valid for T_X.

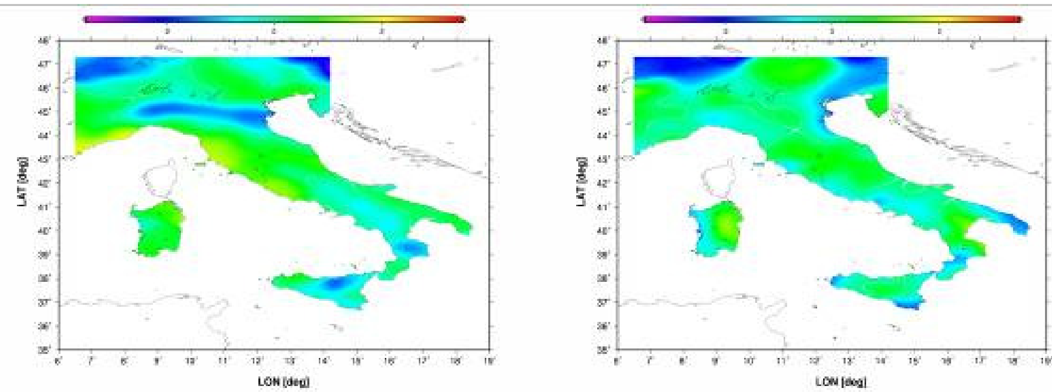


Fig.68-69 Residual interpolation with a GISDW for January and July T_X (°C)

Even for T_X the residuals correction is more important than for T_M. In absolute values the T_X residuals are higher, especially in Po Plain, French Riviera in January and over the Alps and in the north-east in July.

• Improvements after the interpolation of the Residuals and final errors

Once again, residuals were calculated as:

$$RESTM_{MLR+SEA+LAKE+PO+SUN+ASP+TV+UHI+IDS} = RESTM_{MLR+SEA+LAKE+PO+SUN+ASP+TV+UHI} - IDS_M \qquad \textbf{(104)}$$

The same holds for T_N and T_X.

In tab. 36, in the next page, we show the final residuals of temperature models.

		ME	MAE	RMSE		ME	MAE	RMSE		ME	MAE	RMSE
	JAN	0.00	0.81	1.04	JAN	0.01	0.98	1.26	JAN	-0.01	0.87	1.17
	FEB	0.00	0.69	0.90	FEB	0.01	0.93	1.20	FEB	-0.01	0.79	1.08
	MAR	0.00	0.59	0.79	MAR	0.01	0.85	1.11	MAR	-0.01	0.74	1.02
	APR	-0.01	0.57	0.76	APR	0.01	0.83	1.07	APR	-0.01	0.75	1.01
	MAY	0.00	0.58	0.75	MAY	0.01	0.86	1.10	MAY	0.00	0.77	1.02
	JUN	0.00	0.61	0.80	JUN	0.01	0.92	1.17	JUN	-0.01	0.83	1.09
	JUL	0.00	0.65	0.86	JUL	0.02	1.00	1.28	JUL	0.00	0.89	1.16
	AUG	0.00	0.64	0.84	AUG	0.02	0.98	1.25	AUG	0.00	0.88	1.15
	SEP	0.00	0.59	0.78	SEP	0.02	0.90	1.15	SEP	0.00	0.81	1.07
	OCT	0.01	0.61	0.80	OCT	0.02	0.86	1.09	OCT	0.00	0.77	1.04
	NOV	0.01	0.66	0.85	NOV	0.01	0.85	1.08	NOV	-0.01	0.76	1.03
	DEC	0.01	0.81	1.04	DEC	0.01	0.94	1.20	DEC	-0.01	0.88	1.17
After GISDW	YEAR	0.00	0.65	0.85	YEAR	0.01	0.91	1.16	YEAR	-0.01	0.81	1.08
Improvement	YEAR	0.04	0.17	0.20	YEAR	0.05	0.17	0.20	YEAR	0.04	0.22	0.25

Tab.36 Monthly statistical error values after GIDW corrections and comparisons (°C)

After the *GIDW* of the residuals, the *MAE* were considerably improved: 12 out of 12 months for T_M, 11 out of 12 for T_N and 12 out of 12 are lower than 1.*0* °C, and even the *RMSE* were improved and the average year value is only 0.85 °C for T_M, 1.16 °C for T_N and 1.08 °C. *GIDW* led to an improvement of about 0.2 °C for *MAE* and *RMSE* of every variable and it erased the *ME*.

A *GIDW* with a shorter search distance and a shorter distance where the radial weight is halved would lead to lower *MAE* and *RMSE* but it would introduce bulls' eyes and too local effects that are not realistic.

In fig. 70-71 we show the final distributions of the residuals for T_M for January and July.

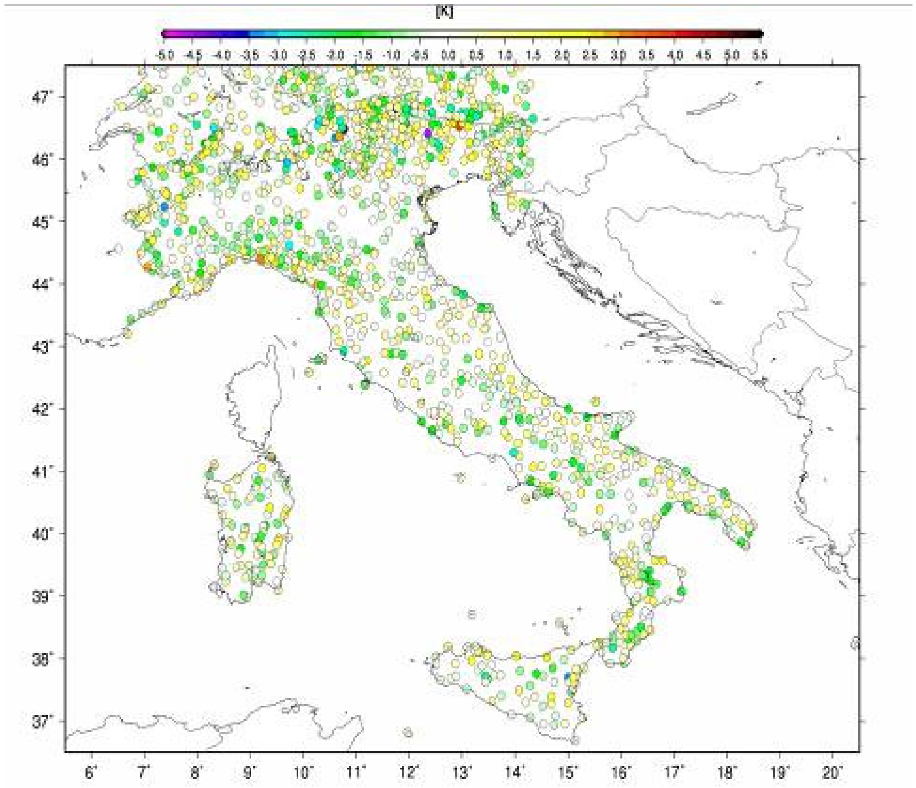


Fig.70 January map of final T_M residuals

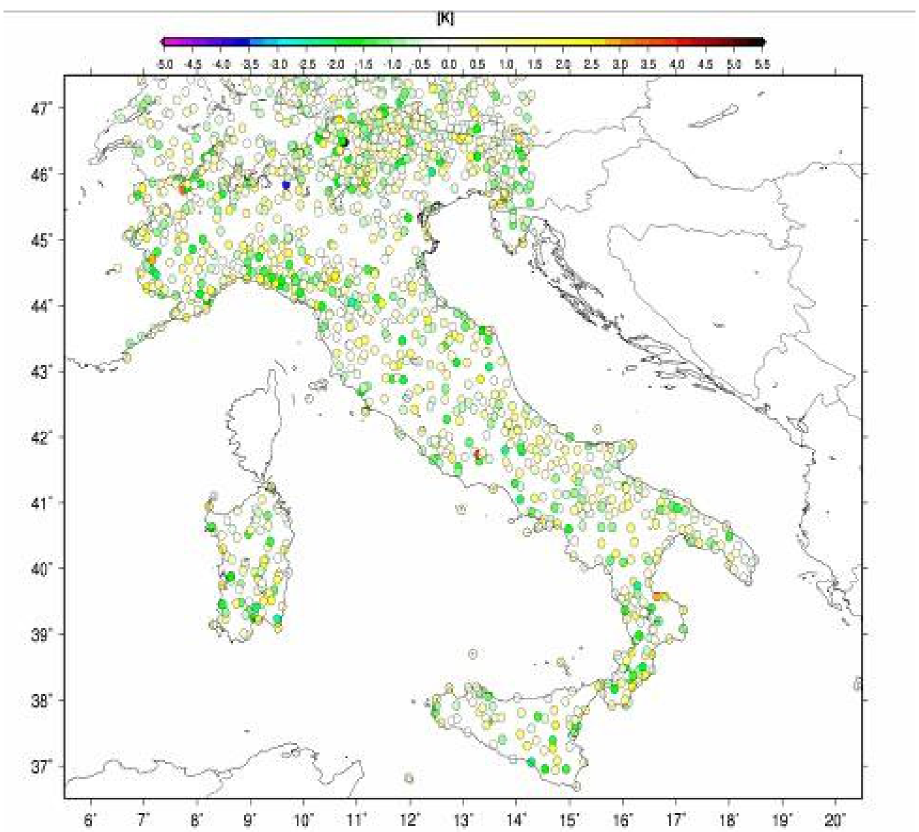


Fig.71 July map of final T_M residuals

As we can see in fig. 70 and fig. 71, the final distributions of the residuals show heterogeneous distributions. No biased regions still exist after residual interpolation.

4.3.11 Step 10: final T_N-T_M-T_X-DTR grids and maps

In Italy, mean temperature data are measured as:

$$TM = \frac{TN + TX}{2} \tag{105}$$

However, we implemented three different models for T_N, T_X, T_M, without considering equation (104). This could lead to unrealistic single effects modelling (as for the *UHI*).

Thus, we first created gridded data for 12 months for T_M, T_N, T_X and a year average grid for T_N, T_X, T_M, with a Fortran dedicated code (used only for these models). Then we considered T_M grids only as the final grids, because T_M data density and quality are higher than that of T_N and T_X. We also calculated and gridded:

$$DTR = (TX - TN) \tag{106}$$

Where *DTR* is the Daily Temperature Range.

Then we obtained the final grids as *New et al. (1999, 2000)*:

$$TN = TM - \frac{DTR}{2} \tag{107}$$

$$TX = TM + \frac{DTR}{2} \tag{108}$$

In the end, we have 52 gridded raster data: 12 + 1 monthly grids for 4 variables (T_N, T_M, T_X, *DTR*). The high-resolution grids were stored in matrix or vector data files and are freely available upon request. The data analyses were performed by means of Windows

Office™ and Ubuntu OpenOffice™ tools, free Fortran compilers (as gfortran™ or Force™) were used to grid single rasters and final grids, *GMT™*, ArcGis™ and R™ tools were used to convert data, to draw maps and to evaluate statistical errors (see Chapter 9.3 for references and websites). In the next pages we will show the 52 maps.

The average elevation of the area modelled is 727 m; the mean temperature is 10.8 °C. Italy is colder than the globe of approximately 4 °C due to the two long and high mountain ridges.

• 1961-1990 Mean Temperature Maps

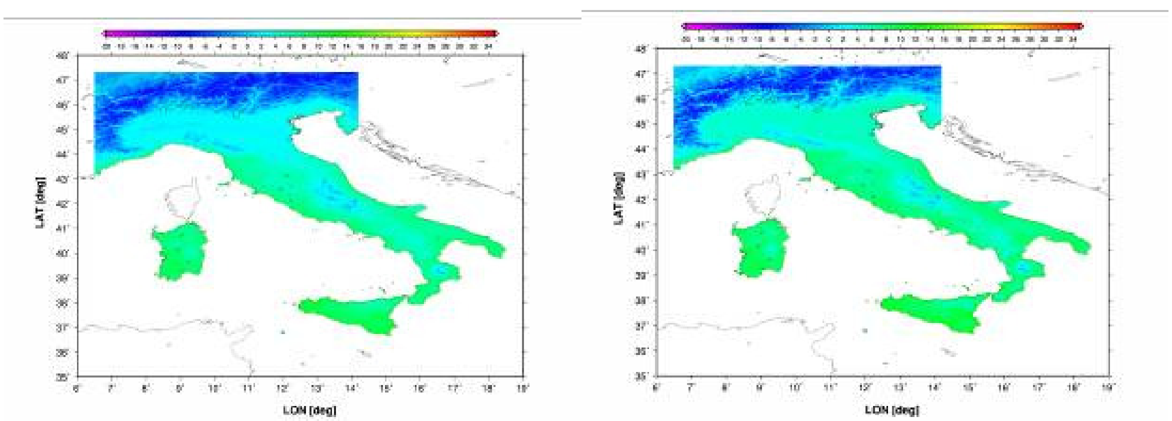


Fig.72-73 1961-90 January and February T_M maps (°C)

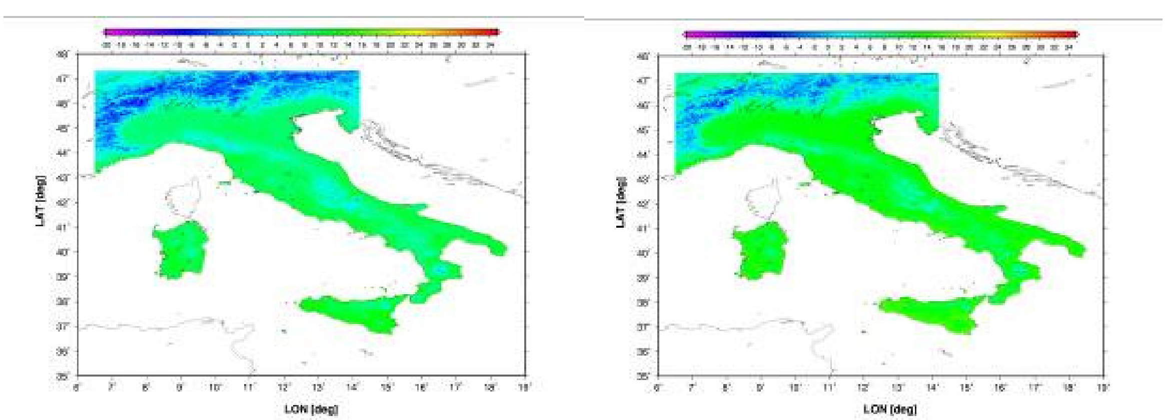


Fig.74-75 1961-90 March and April T_M maps (°C)

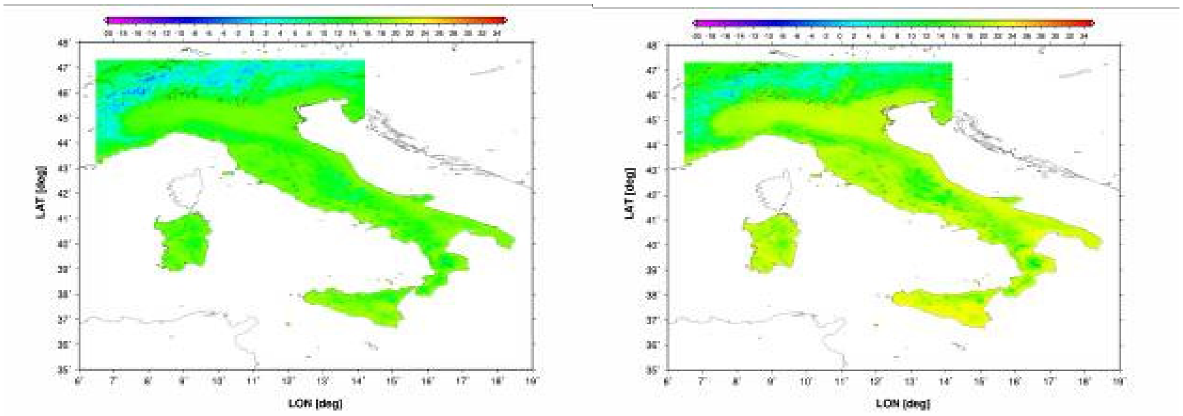


Fig.76-77 1961-90 May and June T_M maps (°C)

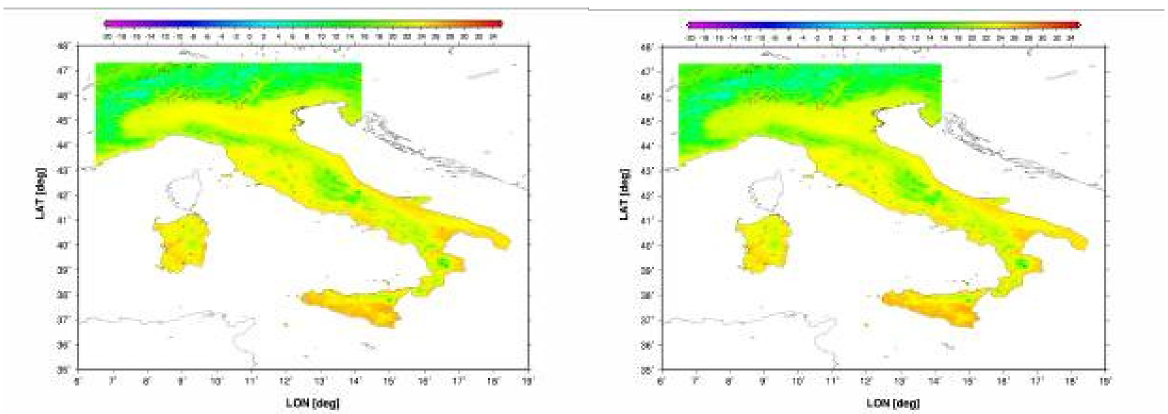


Fig.78-79 1961-90 July and August T_M maps (°C)

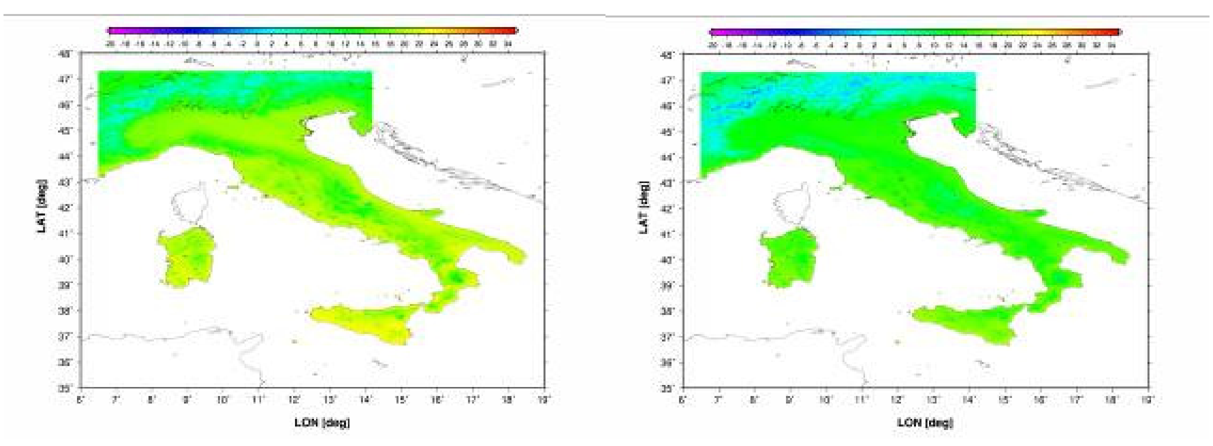


Fig.80-81 1961-90 September and October T_M maps (°C)

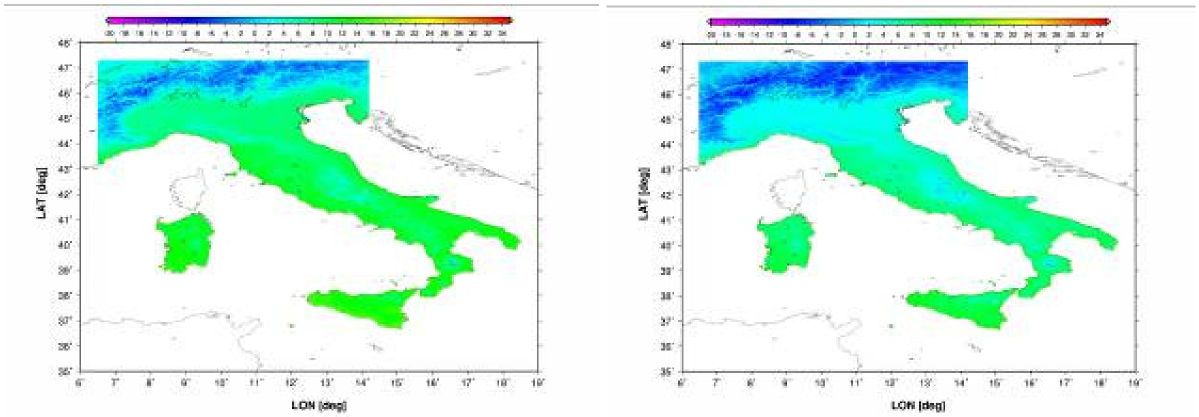


Fig.82-83 1961-90 November and December T_M maps (°C)

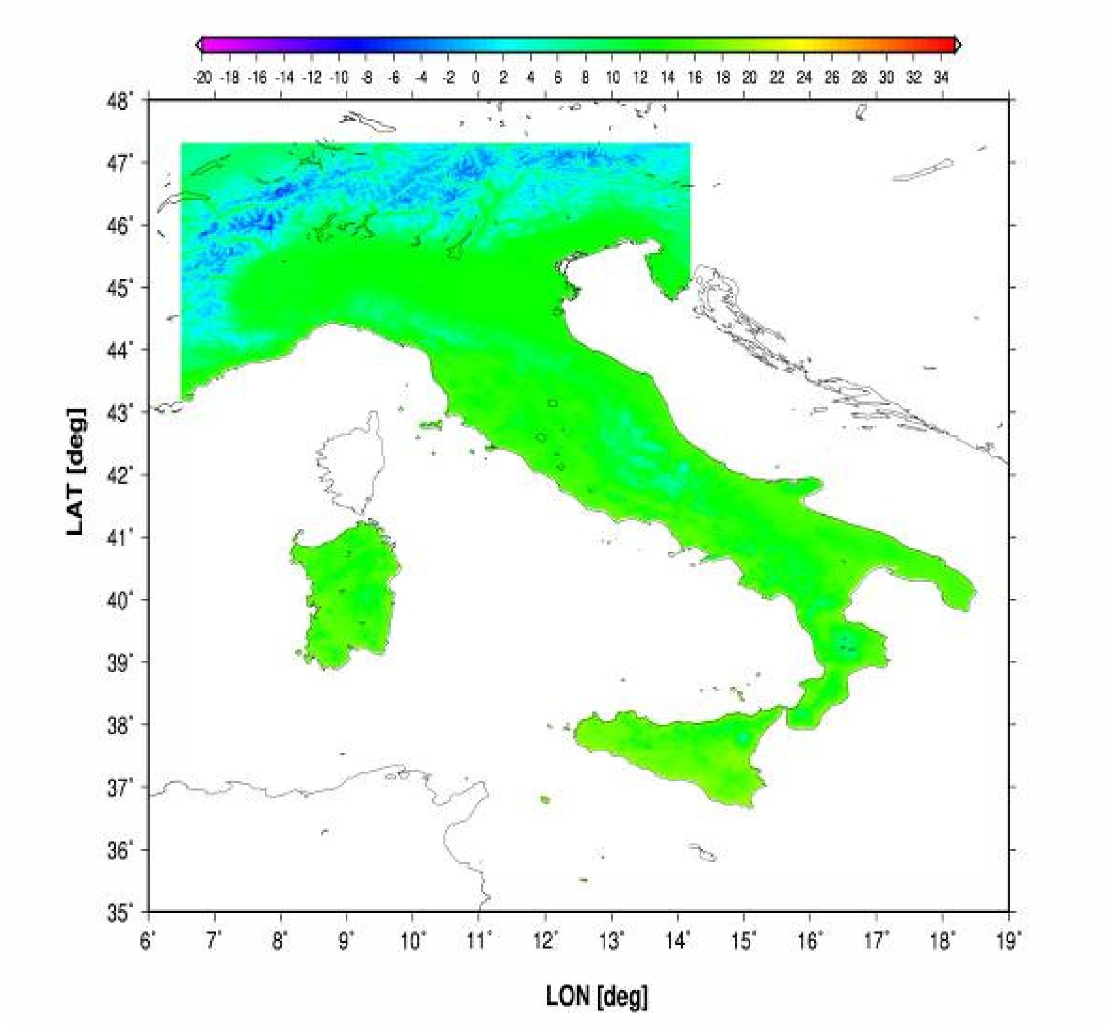


Fig.84 1961-90 Year T_M map (°C)

• 1961-1990 Minimum Temperature Maps

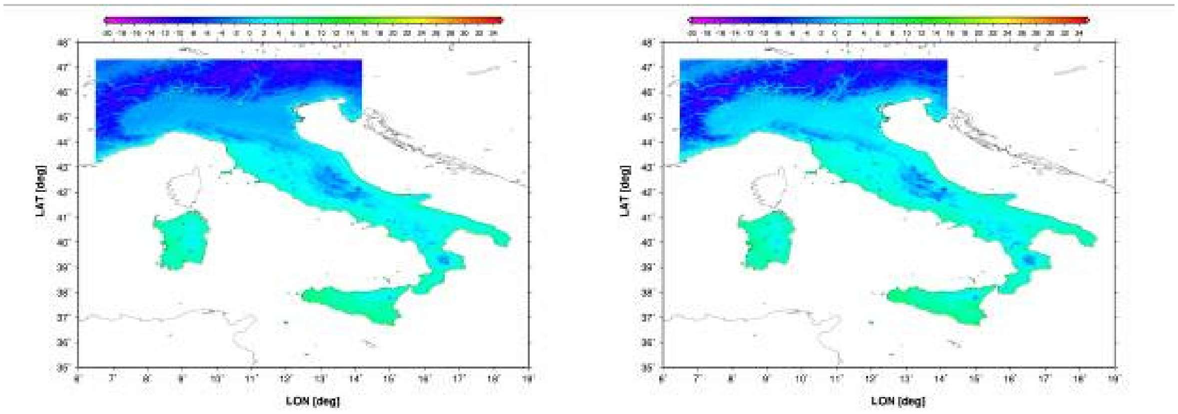


Fig.85-86 1961-90 January and February T_N maps (°C)

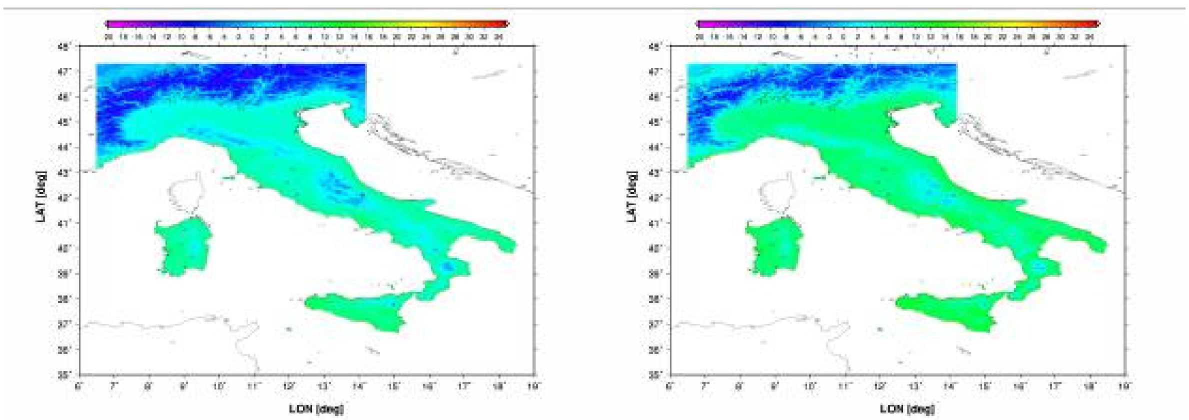


Fig.87-88 1961-90 March and April T_N maps (°C)

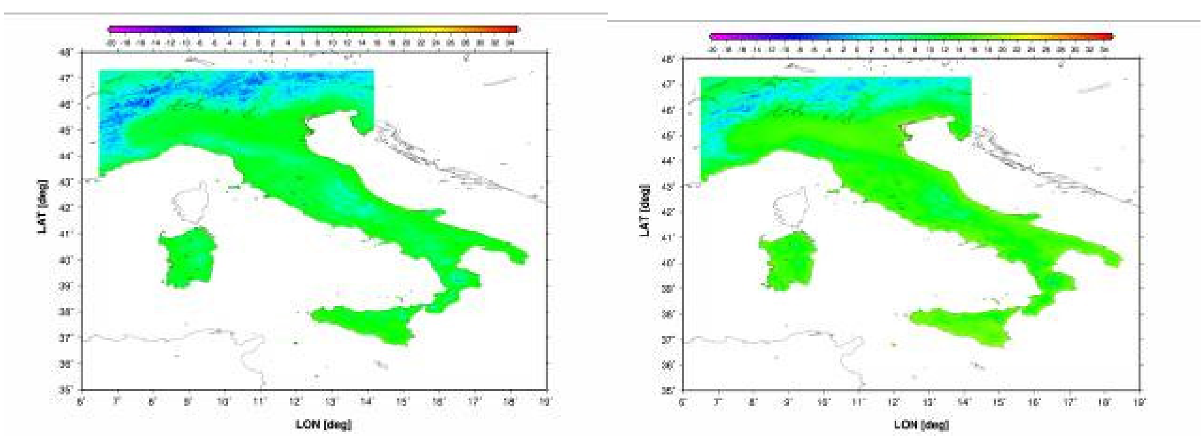
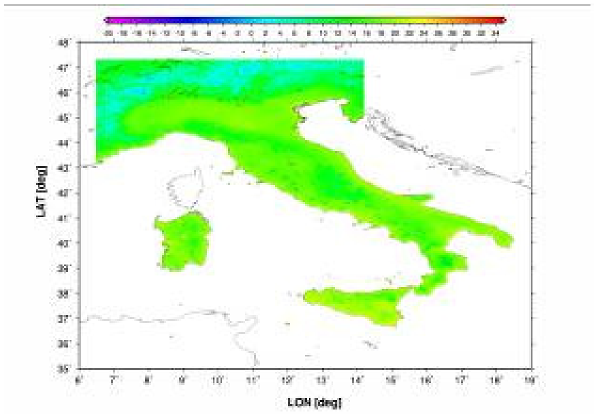


Fig.89-90 1961-90 May and June T_N maps (°C)



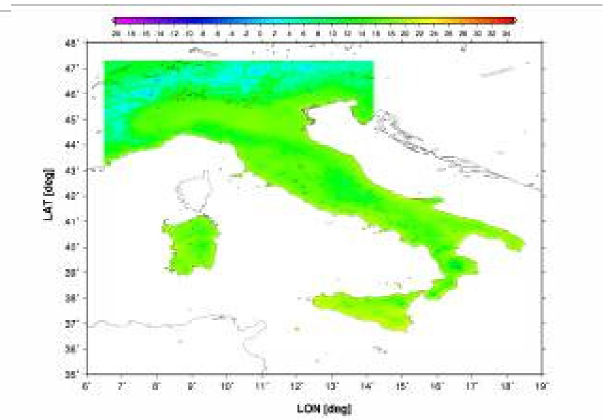
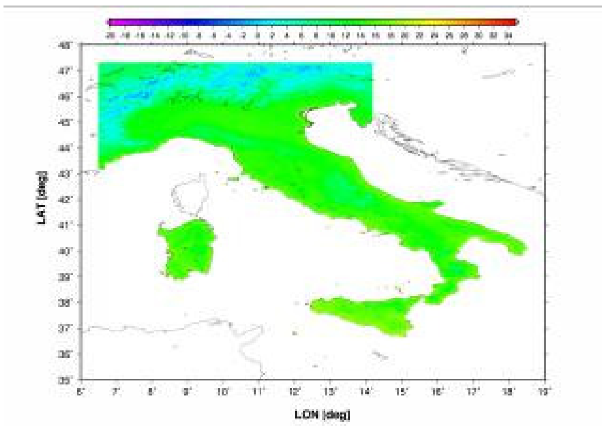


Fig.91-92 1961-90 July and August T_N maps (°C)



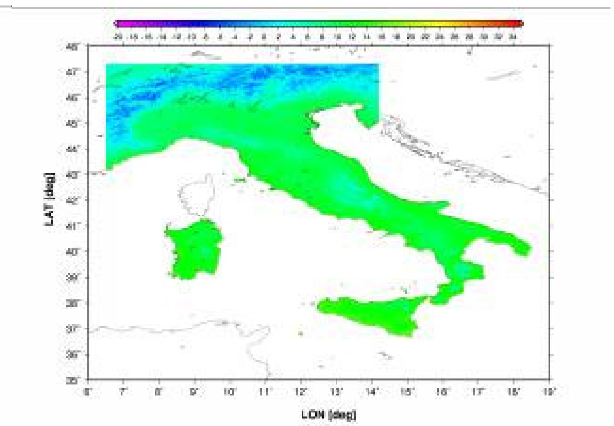
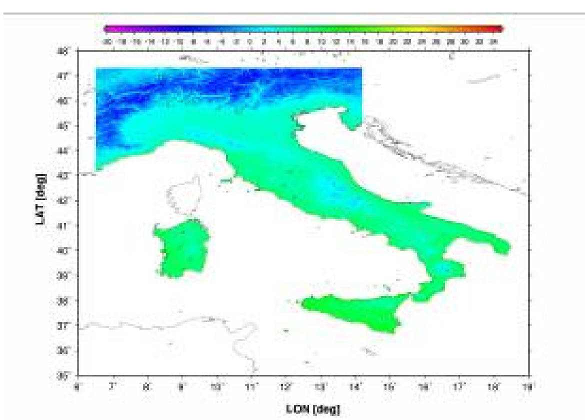


Fig.93-94 1961-90 September and October T_N maps (°C)



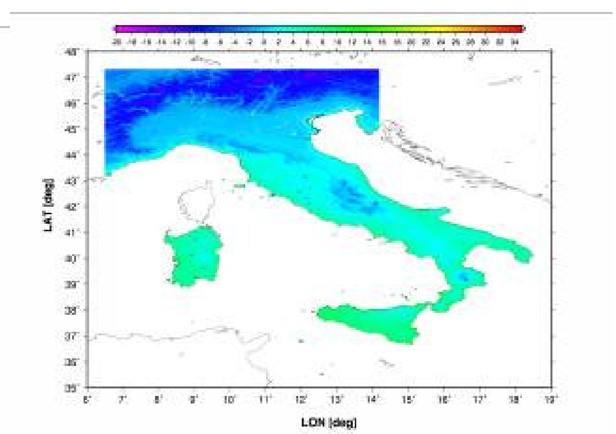


Fig.95-96 1961-90 November and December T_N maps (°C)

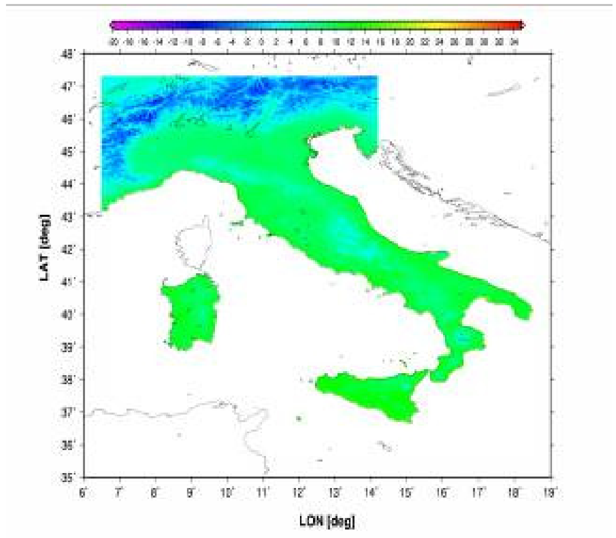


Fig.97 1961-1990 Year T_N maps (°C)

• 1961-1990 Maximum Temperature Maps

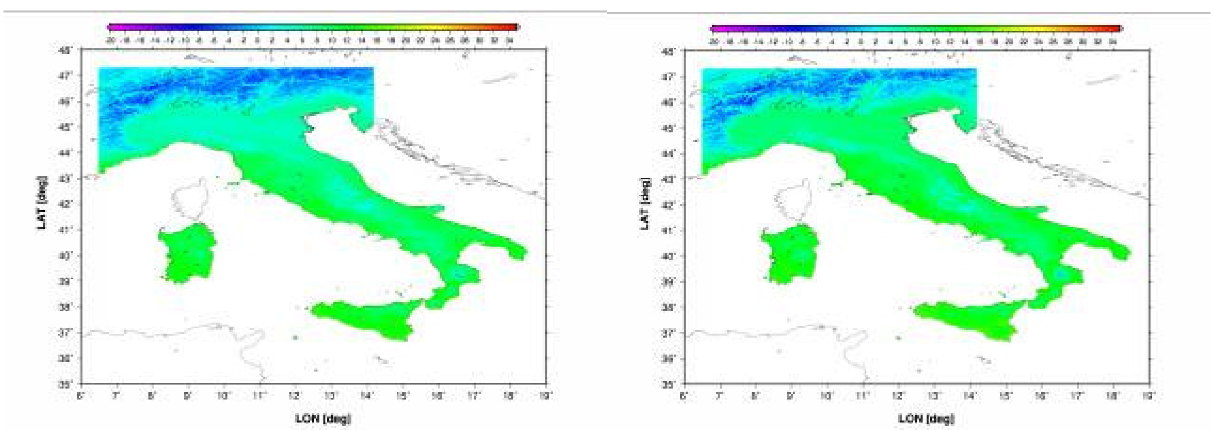


Fig.98-99 1961-90 January and February T_X maps (°C)

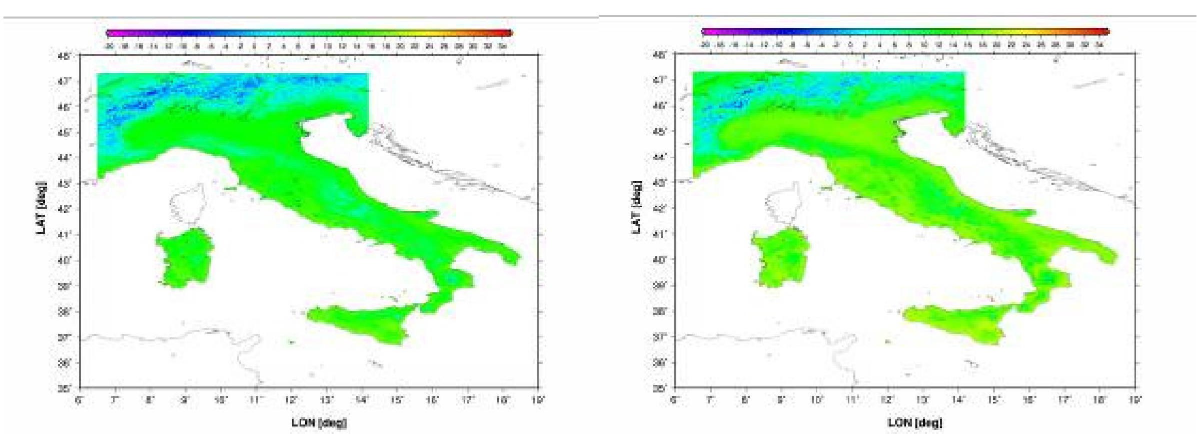


Fig.100-101 1961-90 March and April T_X maps (°C)

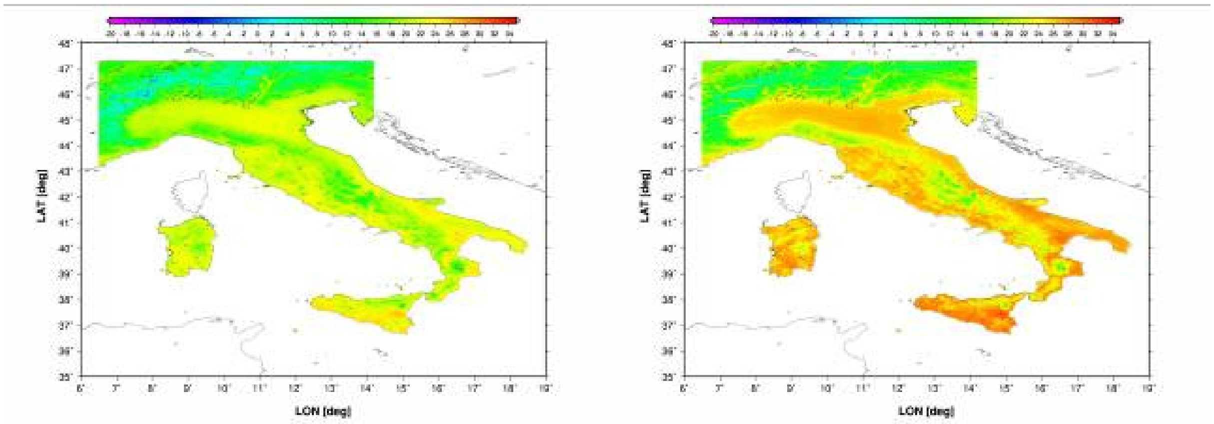


Fig.102-103 1961-90 May and June T_X maps (°C)

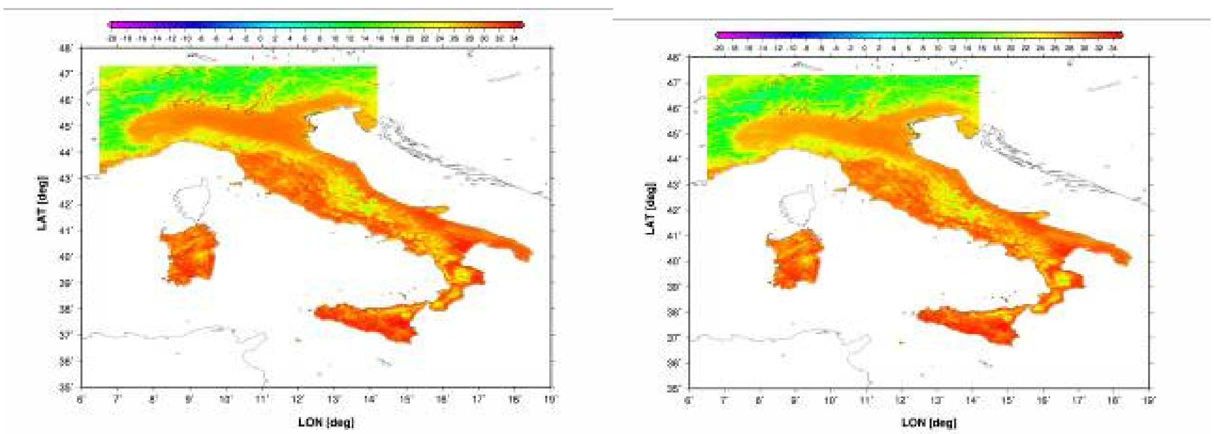


Fig.104-1051961-90 July and August T_X maps (°C)

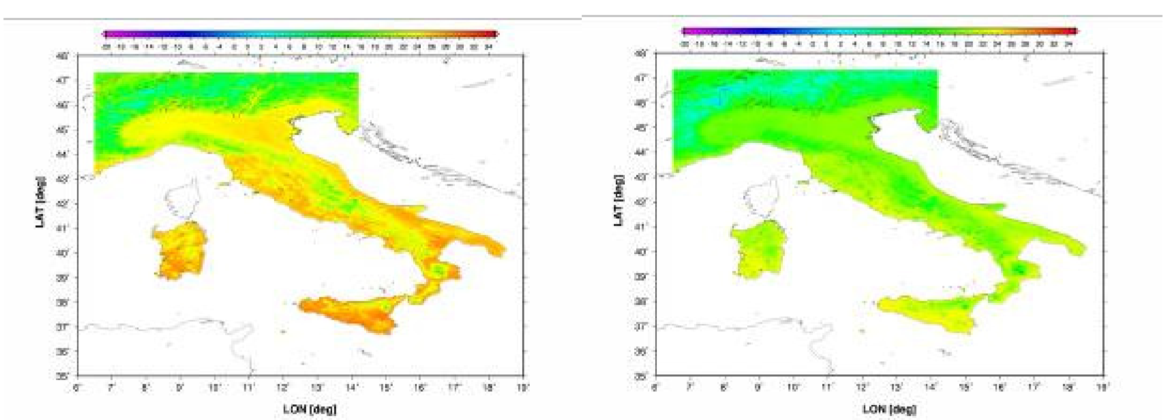


Fig.106-107 1961-90 September and October T_X maps (°C)

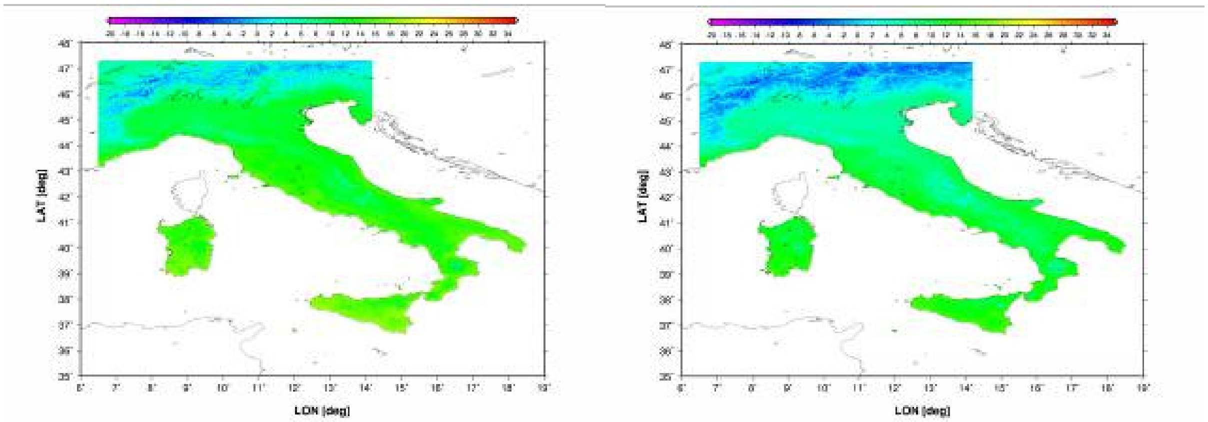


Fig.108-109 1961-90 November and December T_X maps (°C)

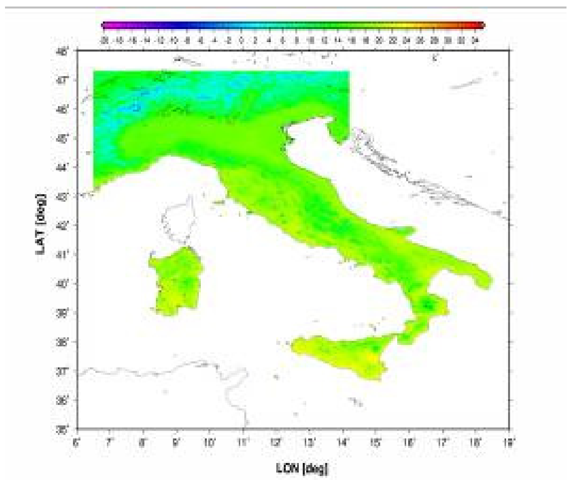


Fig.1101961-90 Year T_X map (°C)

• 1961-1990 Daily Temperature Range Maps

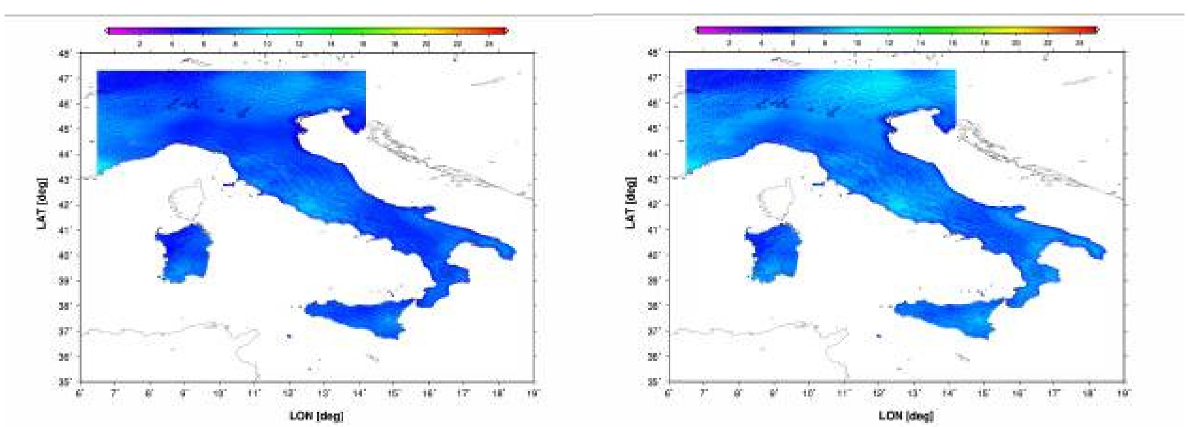


Fig.111-112 1961-90 January and February DTR maps (°C)

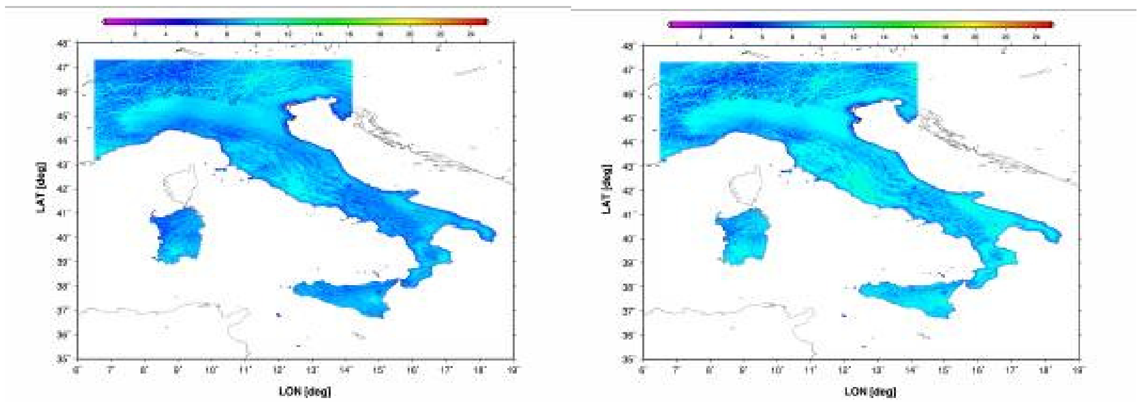


Fig.113-114 1961-90 March and April DTR maps (°C)

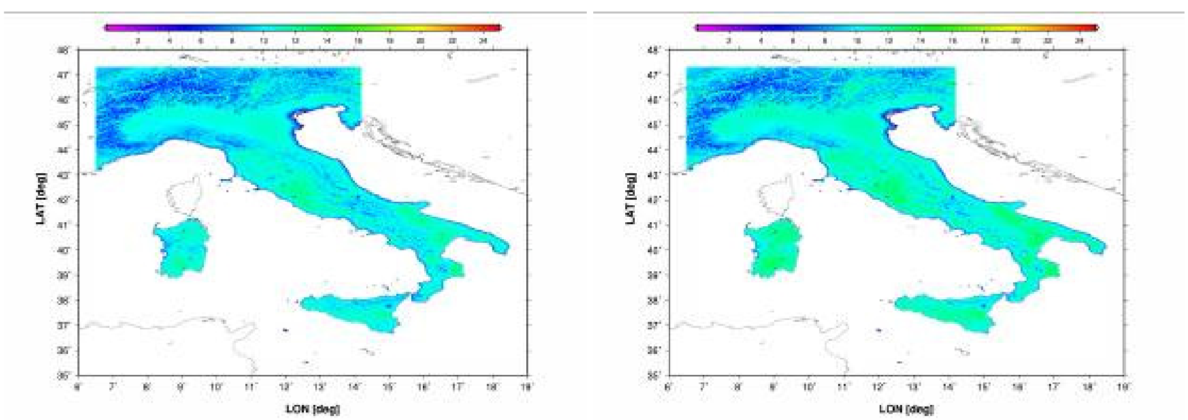


Fig.115-116 1961-90 May and June DTR maps (°C)

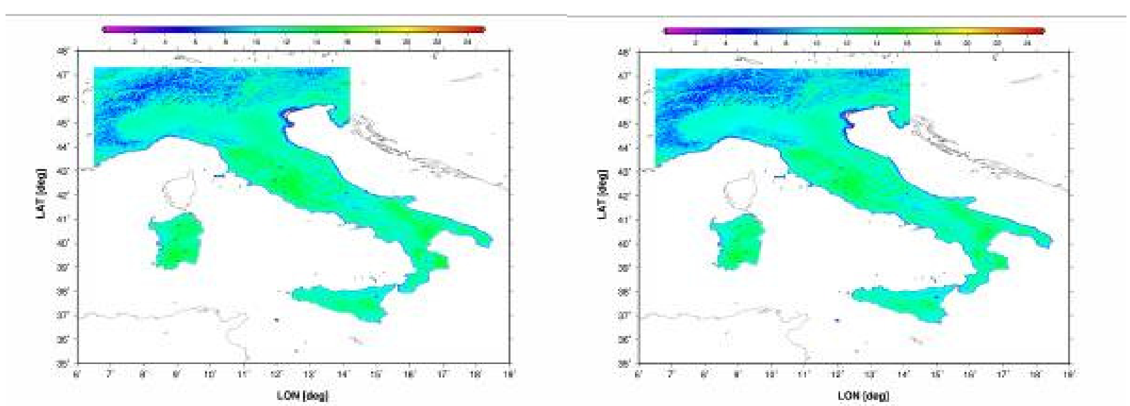


Fig.117-118 1961-90 July and August DTR maps (°C)

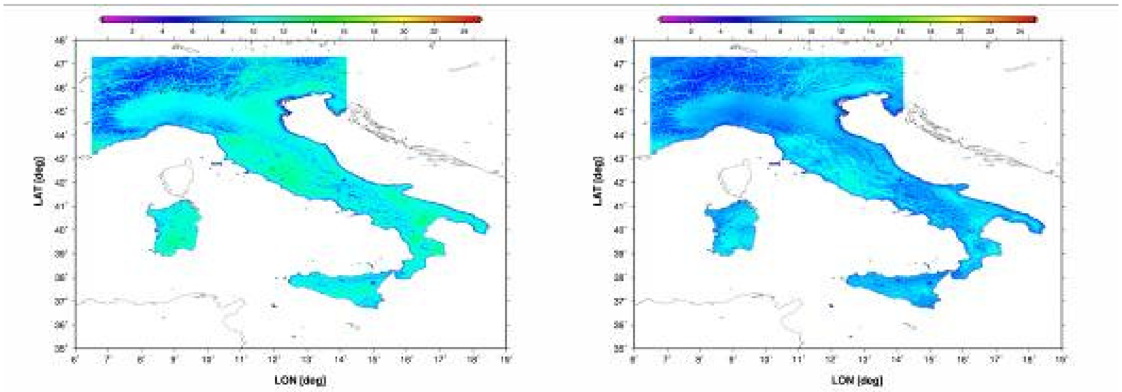


Fig.119-120 1961-90 September and October DTR maps (°C)

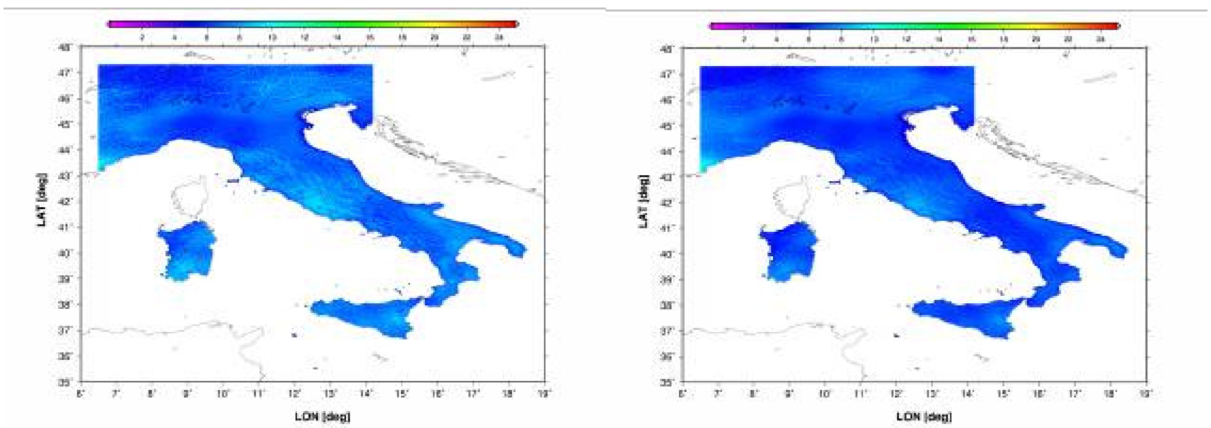


Fig.121-122 1961-90 November and December DTR maps (°C)

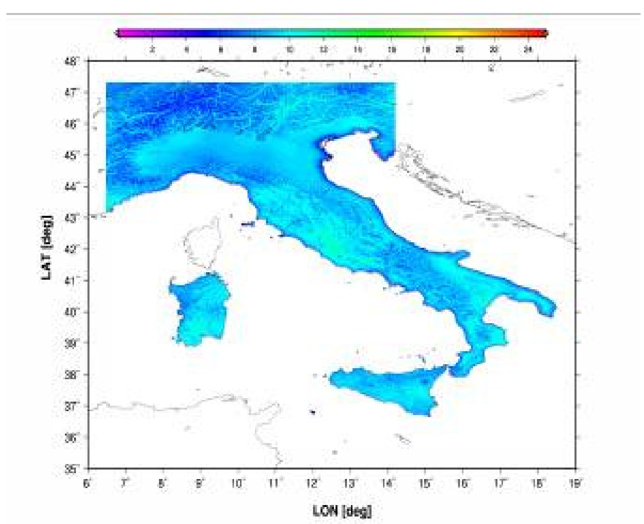


Fig.123 1961-90 Year DTR maps (°C)

4.3.12 Cross Validation

In order to validate our models better, besides the evaluation of the statistical error parameters shown in tab. 36, we withheld a subset of 150 T_M stations and 120 T_N and T_X stations (not the same stations). They were chosen randomly from a Fortran subroutine and recalculated the coefficients using the same methodology, without the chosen stations. Then we applied the coefficients to the withheld subset of stations, and we found the following errors:

1 CV subset	ME	MAE	RMSE			ME	MAE	RMSE			ME	MAE	RMSE
JAN	0.20	0.90	1.34		JAN	0.07	0.76	0.94		JAN	-0.08	0.62	0.79
FEB	0.18	0.87	1.29		FEB	0.07	0.63	0.79		FEB	-0.03	0.59	0.75
MAR	0.17	0.82	1.18		MAR	0.04	0.54	0.69		MAR	0.02	0.56	0.72
APR	0.15	0.81	1.12		APR	0.04	0.57	0.71		APR	0.03	0.56	0.72
MAY	0.16	0.84	1.16		MAY	0.05	0.58	0.73		MAY	0.05	0.67	0.84
JUN	0.15	0.90	1.27		JUN	0.05	0.61	0.77		JUN	0.04	0.78	0.99
JUL	0.15	0.98	1.37		JUL	0.04	0.64	0.82		JUL	0.02	0.83	1.04
AUG	0.15	0.95	1.33		AUG	0.02	0.63	0.78		AUG	0.03	0.76	0.98
SEP	0.13	0.88	1.25		SEP	0.02	0.59	0.74		SEP	-0.01	0.66	0.86
OCT	0.12	0.83	1.19		OCT	0.03	0.59	0.74		OCT	-0.08	0.63	0.81
NOV	0.16	0.78	1.15		NOV	0.01	0.63	0.78		NOV	-0.05	0.57	0.74
DEC	0.18	0.86	1.24		DEC	0.04	0.76	0.94		DEC	-0.09	0.63	0.82
YEAR	0.16	0.87	1.24		YEAR	0.04	0.63	0.78		YEAR	-0.01	0.65	0.84

Tab.37 Monthly statistical error related to a Cross-Validation subset (°C)

Then we performed CV withholding 10 random subsets (10% stations) for T_M and we averaged results, thus we found:

10 CV subsets	ME	MAE	RMSE
JAN	0.02	0.87	1.07
FEB	-0.01	0.69	0.88
MAR	-0.04	0.60	0.77
APR	-0.08	0.56	0.73
MAY	-0.07	0.55	0.70
JUN	-0.07	0.56	0.75
JUL	-0.06	0.61	0.83
AUG	-0.02	0.60	0.81
SEP	-0.03	0.57	0.72
OCT	-0.01	0.62	0.76
NOV	0.03	0.66	0.84
DEC	0.04	0.86	1.07
YEAR	-0.02	0.65	0.83

Tab.38 Monthly statistical error related to 10 Cross-Validation subsets (°C)

If we compare the errors in tab. 38 with the statistical parameters calculated over all the grid cells corresponding to the stations, we find that the first method yields a null *ME* versus a -0.02 °C *CV ME*, a 0.65 °C *MAE* that is equal to *CV MAE*, a 0.85 °C *RMSE* that is higher than the 0.83 °C *CV RMSE*.

The 1961-1990 High-Resolution Temperature models for Italy were satisfactorily validated with two independent methods.

4.4 Other temperature models realized during the PhD project

In the framework of this PhD project, we realised other temperature models that were also used as tests to develop the methodology that yields to the 1961-1990 temperature models for Italy.

• Northern Italy 1961-90 T_M models for *EUMETNET/ECSN GAR* grids

In the frame of the *EUMETNET/ECSN* GAR project (*Auer et al., 2007)*, we implemented monthly 1961-1990 T_M climatologies for Italy north to 42.5 °N. We used a stepwise *LR* model plus local improvements without residual interpolations; the database was part of the *HISTALP* (*Auer et al., 2007*) database. These *GAR* climatologies can be found in *Hiebl et al. (2009)*, the northern Italy climatologies used for *GAR* climatologies can be found in *Spinoni et al., (2008)*.

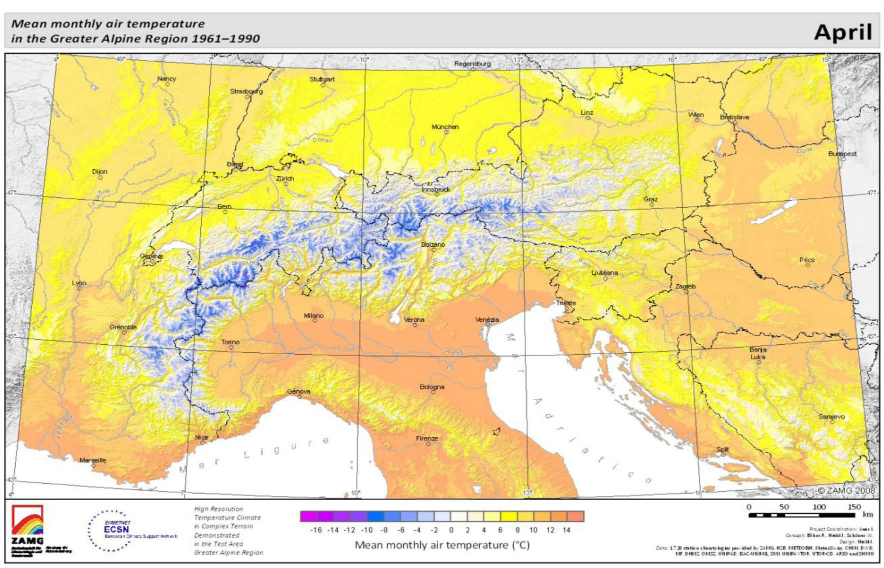


Fig.124 1961-1990 April T_M map for GAR region (Hiebl et al., 2009) (°C)

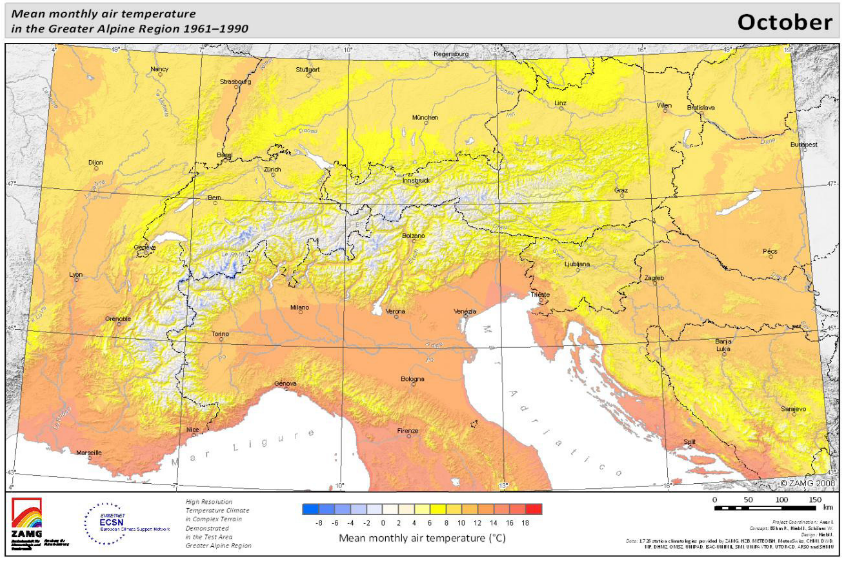


Fig.125 1961-1990 October T_M map for GAR region (Hiebl et al., 2009) (°C)

• Northern and Central Italy 1961-1990 T_M models

We tested our updated database in the northern part and central part of Italy, by means of a stepwise *LR*, plus local improvements. The T_M climatologies can be found in *Brunetti et al., (2009)*.

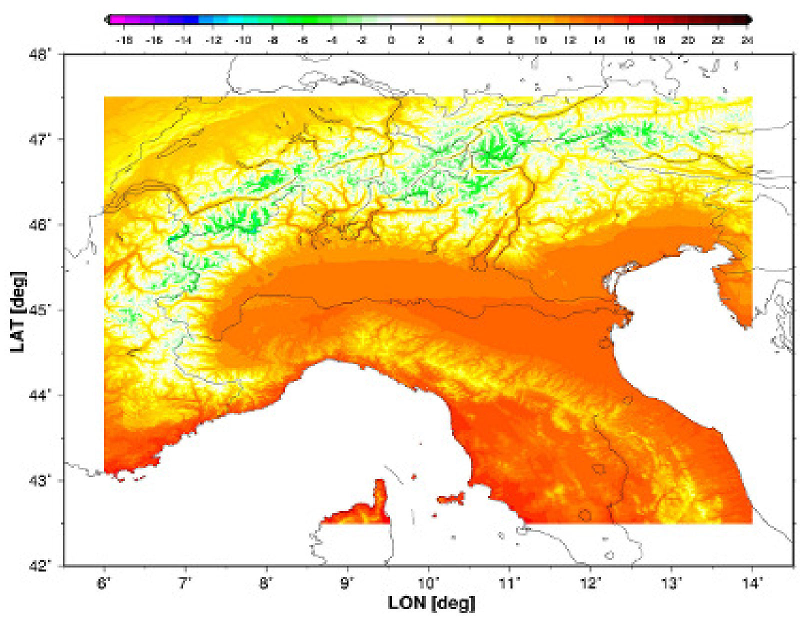


Fig.126 1961-1990 Year T_M map for northern and central Italy (Brunetti et al., 2009) (°C)

• North-Eastern Italy1961-1990 T_N-T_M-T_X models

We tested some new parameters as the "*NCEL*" parameter (see equation (91)) in the north-eastern part of Italy, by means of a stepwise *LR* plus local improvements for T_N, T_M and T_X. The monthly climatologies can be found in *Brunetti et al., (2010)*.

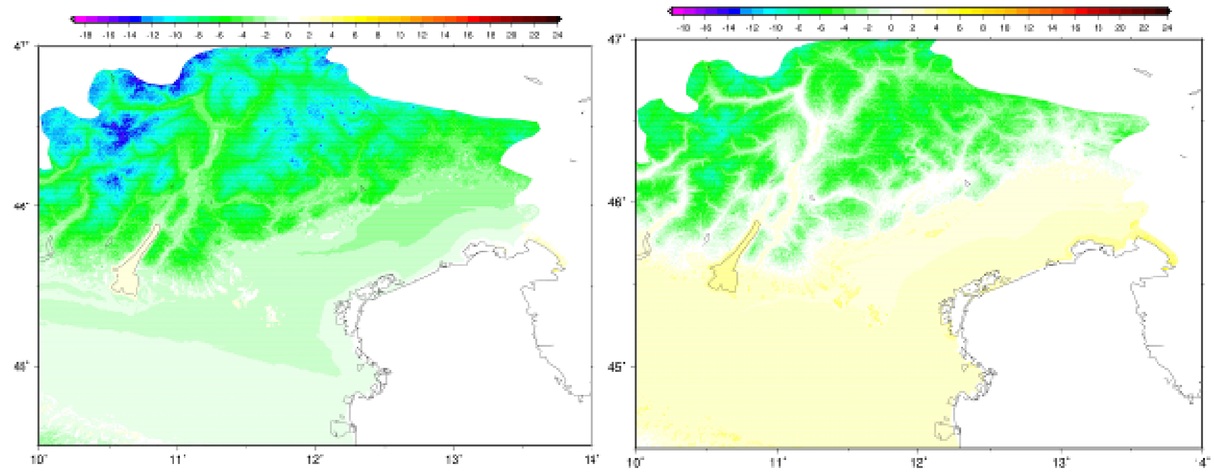


Fig.127-128 1961-1990 January and March T_M maps for north-eastern Italy (Brunetti et al., 2010) (°C)