ESTIMATION OF DAILY MEAN AIR TEMPERATURE FROM MODIS LST IN ALPINE AREAS

Angelo Colombi[1], Carlo De Michele[2], Monica Pepe[1] and Anna Rampini[1]

1. IREA-CNR, Institute for Electromagnetic Sensing of the Environment, National Research Council, Milano, Italy; {colombi.a / pepe.m / rampini.a}(at)irea.cnr.it
2. Polytechnic of Milan, Department of Hydraulics, Milan, Italy; carlo.demichele(at)polimi.it

ABSTRACT

This study is aimed at demonstrating the feasibility of the MODIS Land Surface Temperature (*LST*) product as a source for calculating spatially distributed daily mean air temperature to be used as input for hydrological or environmental models. The test area is located in the Italian Alpine area. The proposed procedure solves, by empirical approaches, the problem of relating *LST* to the Air Temperature (T_{air}) and instantaneous T_{air} values to daily mean values, exploiting ground data weather station measurements as a reference. The relationship between *LST* and T_{air} is determined by correlation analysis and equation generalisation for spatial distribution. The extrapolation of daily mean values of T_{air} from instantaneous values is addressed again by correlation analyses taking into account the altitude variability and exploiting historical series. Validation was accomplished by accuracy assessment procedures both punctual and spatially distributed, the latter performed by comparison with the Inverse Distance Weighting (IDW) interpolation method.

The proposed methodology produced satisfactory results as related to the objective: The daily mean air temperatures derived by *LST* showing an overall *RMSE* of 1.89°C, and slightly outperforms the interpolation method used as comparison.

INTRODUCTION

The air temperature near the Earth's surface is a key variable to describe energy and water cycles of the Earth-atmosphere system and its measurement is necessary to run several snowmelt hydrological (HBV (1); SRM (2)) and environmental models. The measurement of this variable by meteorological stations, however, provides only punctual values, whereas most models require spatially distributed variables and parameters to evaluate physical processes with an adequate spatial scale. In most cases, an interpolation of point site data is carried out only by using a vertical temperature lapse rate because of the strong relationship between the air temperature and the elevation. But, when considering a regional scale, it's necessary also to refer to spatial interpolation methods, allowing for land cover and topographic effect factors. Moreover, high-elevated areas are often not properly monitored by weather station networks. In several cases, weather stations are installed only after natural disasters (such as avalanches or landslides) leading to irregularly equipped regions: very dense station networks and un-gauged areas.

In this context, remote sensing from satellites can offer a contribution since it regularly provides spatially distributed information also about isolated areas.

However, radiometers cannot directly measure air temperature, but estimations of the land surface temperature (*LST*) can be obtained by means of specific algorithms from Thermal Infra-Red (TIR) channels satellite data (3). Coarse/medium spatial resolution sensors like the Advanced Very High Resolution Radiometer (AVHRR) and, more recently, the Moderate-resolution Imaging Spectroradiometer (MODIS) have already been extensively used and tested for the LST retrieval (4,5,6). In particular, MODIS furnishes a *LST* product at 1 km spatial resolution covering all the Earth's surface twice a day.

A strong correlation between *LST* and air temperature, already observed and analysed in previous works (7), represents the basis of our analysis.

However, there is a major constraint in the use of optical TIR remotely sensed observations which concerns the cloud coverage: Under cloud sky conditions it is impossible to retrieve any information. As compared to a weather station network that operates every day, information can therefore result incomplete, considering that most models require data on a daily basis for a long period.

In previous research, similar analyses were carried out using a Temperature Vegetation indeX (TVX) technique (8,9). This methodology is based on the assumption that the surface temperature of a closed canopy is equal to the air temperature and so the air temperature can be derived by solving the regression equation $LST = a + b \cdot SVI$ for the saturated value of a Spectral Vegetation Index (*SVI*), usually the Normalized Difference Vegetation Index (*NDVI*). However, the aim of our study is to obtain an estimation of T_{air} in the Alpine area, where high altitudes make the *NDVI* unusable because of snow and bare ground. Moreover, by using the *NDVI*, only day-time satellite overpass can be considered, losing night-time observations. Other previous experiences about the derivation of air temperature from TIR channels analysed the problem at a continental scale in order to verify climate global change (10); in this work, we deal with a regional scale for which a higher accuracy and resolution are required.

METHODS

The selected sensor for the study is the Moderate Resolution Imaging Spectroradiometer (MODIS), onboard EOS NASA's TERRA platform, that provides thermal infrared data in ten spectral bands (6.5-15 µm bandwidth) at a spatial resolution of 1 km. Its suitability for the purpose is based on the MODIS sensor's spatial, temporal and spectral characteristics. The latter allow of the solution of the thermal infrared radiometric correction by means of the *MODIS Generalized Split-Window LST Algorithm* (11,12,13). This algorithm is based on average emissivity in bands 31 (10,780-11,280 µm) and 32 (11,770-12,270 µm) to retrieve LST from the MODIS sensor.

Although satellites provide global coverage, this kind of measure requires two important operations in order to properly map daily mean air temperature as required by most models; therefore our study can be conceptually divided into two different steps. Since the satellite sensor gives information only about LST, the first step (hereafter STEP1) of the analysis is the evaluation of relationships between the satellite-derived LST and the *in situ* measurements of T_{air}, taken as close as possible to the moment of the satellite overpass. The second step of the analysis (hereafter STEP2) concern the evaluation of daily mean values of air temperature from instantaneous values. In fact, unlike weather stations, Earth Observation data are acquired instantaneously twice a day, while hydrological and climate models often require a daily mean value (14). Nevertheless, polar orbit satellites overpass the same area every day at approximately the same time, so that data sets belonging to different dates are comparable.

The study of both of these relationships – STEP1 and STEP2 - enable input data to be obtained for different hydrological and environmental models. Naturally, the accuracy of this analysis affects that of the models, and so their sensitivity must be considered too.

The selected study area (3,500 km^2) is located in the Italian Alps, Lombardia district (Figure 1), where Terra overpasses every day between 10 and 11 AM and between 9 and 10 PM.

STEP1 consisted in a statistical procedure to retrieve a relationship between *LST* and T_{air} during the period between January and June 2003.

The EO data set is composed of all passages of Terra satellite in this 182-day period: 364 MODIS LST products (MOD11_L2 swath, 182 day-time and 182 night-time). From this data set, the clear sky control resulted in 124 usable dates (70%); whereas only few pixels were sufficiently clear to be analysed in the remaining dates.

In situ measurements of the air temperature, at 2-meter height, were collected at half-hour time steps in the January-June period from 1996 to 2003 by a network of 42 meteorological stations (Figure 1), managed by the Regional Environmental Protection Agency of Lombardia (ARPA Lombardia).

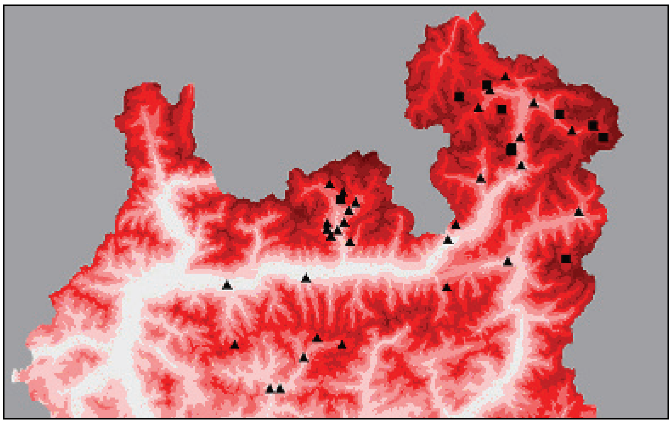

Figure 1: Study area. In evidence all the weather stations used in the analysis. Station sited at the highest altitudes (more than 2000 m) are shown as squares instead of triangles.

Statistical analyses – STEP 1: Air temperature and land surface temperature relationship

The STEP1 correlation analyses were applied at first separately to each of the 42 selected stations to assess the feasibility of such a statistical approach; then they were applied to the entire data set for identifying a typical equation representative of the area at hand, or of some sub-areas.

The first correlations were obtained for each station, distinguishing between night-time and day-time satellite overpass, as a direct consequence of the different trend observed for each station during the two different moments.

This analysis pointed out a strong linear correlation for all selected stations (determination coefficient R^2 always higher than 0.75 and often higher than 0.9) and there were no evidences of different dynamic depending on different months. Obtained relationships were then compared with the T_{air} - LST equation. In 69 of the 84 cases (42 day-time and 42 night-time) the regression line showed an angular coefficient lower than 1 and intersects the T_{air} - LST line. This can be considered as a consequence of a physical characterisation of the area: For high temperatures LST is greater than T_{air}, while it is the opposite for low temperatures. Figure 2 shows this feature for one meteorological station (*Alla Braccia*).

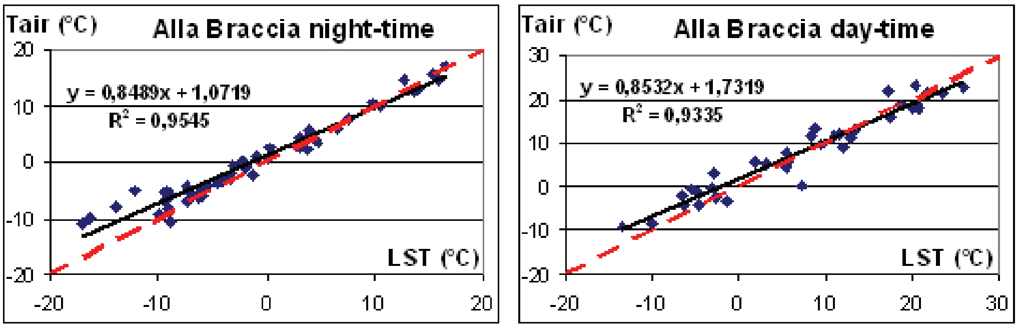

Figure 2: Night-time and day-time relationships estimated for Alla Braccia station. For low temperatures LST is lower than T_{air}, while for high temperatures LST is higher than T_{air}.

Starting from these encouraging results, some procedures of generalization were tested in order to obtain, for a selected group of stations, a unique equation representative of the relationship between T_{air} and LST. On the basis of the statistical sample and considering the literature, we focused our attention on the land cover, on the altitude and on the proximity of available stations. These three criteria were used to introduce a sub-setting of the data set in different sub-areas, and their consistency and performances were tested during the analysis.

In particular, as regard to the land cover, the stations showed a similarity that did not allow of the evaluation of a different behaviour (31 of 42 stations belong to the Forest and seminatural area CORINE Land Cover macrocategory).

The best generalization results were achieved by considering altitude as the main variable that affects this kind of relationship. In fact, on the assumption that stations at close altitude behave similarly, two relationships (one day-time and one night-time) for high altitude stations were found, using the 11 stations, among the 42 available (Figure 1), located at an altitude higher than 2,000 metres. In particular, from these selected 11 stations, we used six stations for the calibration of these two relationships over the 182-day period, while the other five were used to validate the relationships during the same period.

Outcomes clearly demonstrated the actual possibility to obtain two general relationships (one day-time and one night-time) as the altitude is considered the sub-setting criterion. The most satisfactory results are obtained considering the night-time equation: In particular, the determination coefficients (R^2) are equal to 0.86 for night-time (see Figure 3) and to 0.80 for day-time, with validation Root Mean Square Error (*RMSE*) of 2.47 and 3.36°C, respectively. The resulting equations are:

$$T_{air} = 0.649 \cdot LST + 1.4036 \qquad \text{day - time}$$

$$T_{air} = 0.791 \cdot LST + 2.7691 \qquad \text{night - time}$$

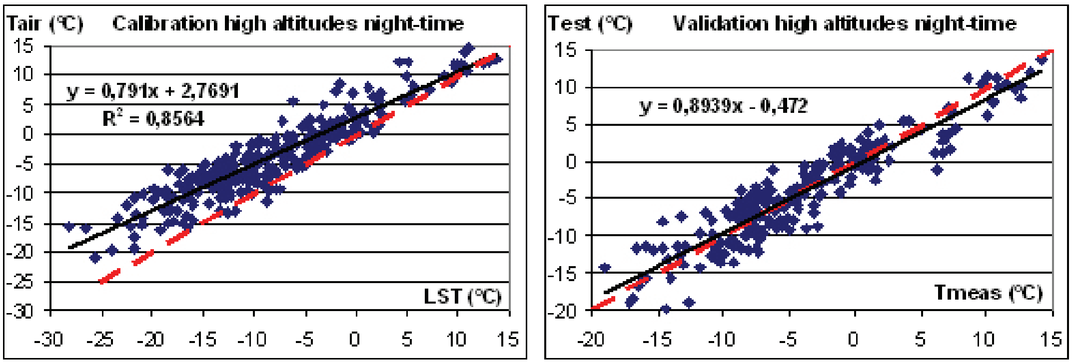

Figure 3: Relations estimated for high altitude stations during night-time. The calibration (left figure) shows a R^2 of 0.86; from the validation (right figure) a regression line very close to the 1:1 line was obtained.

The last generalization procedure, based on the closeness of the selected stations, was performed by considering stations belonging to different hydrological watershed. However, this last analysis led to worse results in terms of *RMSE* (between 3.14 and 4.64°C).

These results confirm that the relationship between *LST* and T_{air} depends on the altitude range rather than on other parameters, which may be explained by the importance of altitude for both air and surface temperatures. The analysis based on the altitude criterion showed another important result: While night-time and day-time relationships are very similar for stations located at high altitudes, the trend is very different for the lowest stations and, in some cases, seems to demonstrate the thermal inversion phenomenon. In fact, during night-time T_{air} is greater than *LST*, while LST is higher during day-time. Figure 4 shows this behaviour for one station (*Morbegno* sited at 242 m a.s.l).

Statistical analyses – STEP2: Instantaneous and daily mean air temperature relationship

STEP2 of the analysis regarded the extrapolation of daily mean value of air temperature (T_{air} mean) from the instantaneous values (T_{air} inst). The estimation of this kind of relationship is based on the existence of a standard air temperature diurnal cycle, characterised by a typical shape, dependent on several conditions such as the altitude, the latitude and the season, as a consequence of solar irradiation.

First of all, some empirical relationships considering all of these conditions were estimated. The main feature that clearly emerged was the strong linear correlation between instantaneous and daily mean air temperature observed in every station, every hour and every month. Therefore, the similarity of these relationships suggested, also in this case, to subdivide the analysis according to the altitude range (0-500 m, 500-1,000 m, 1,000-1,500 m, 1,500-2,000 m and higher than 2,000 m) and according to the month (from January to June), with reference only to the closest satellite overpass hours, whereas the latitude influence was considered negligible due to the regional scale of the study.

From the processing of the data, two linear correlations between T_{air} mean and T_{air} inst were obtained for each month in each elevation range, one referring to the day-time and one to the nigh-time MODIS overpasses, respectively. All obtained relationships were characterised by high values of the determination coefficient (R^2 from 0.83 to 0.92).

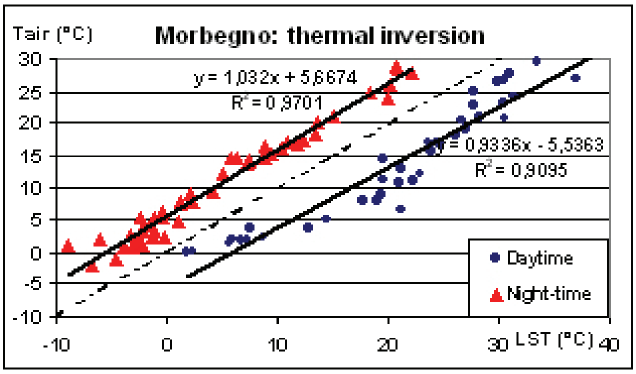

Figure 4: Thermal inversion observed at Morbegno weather station.

RESULTS

For evaluating STEP1 and STEP2 analyses together, a validation procedure was designed taking into account the twofold character of the analyses: punctual and spatially distributed.

The first validation procedure is based on the application of single punctual relationships (i.e. the ones retrieved in STEP1 for each station) to retrieve the instantaneous air temperature from *LST*. Then, according to the month and the altitude, the correlations found within STEP2 were applied, obtaining the final estimations of daily mean air temperature from the instantaneous values. Therefore, by a comparison with the actual values of mean air temperature measured by the ground stations, it was possible to verify the accuracy of results.

In order to evaluate the results under different conditions, we selected ten weather stations from the previous 42, located at different altitudes and not used in the calibration of the STEP2 relationships. Then, we selected, within the analysed six-month period, two days for each month, not used in the calibration phase of STEP 1, during which satellite derived LST were taken, one day-time and one night-time. From a total number of 120 data - twelve days over ten stations - the validation showed *RMSE* values ranging between 1.44 and 3.69°C, according to the different stations considered, and a global *RMSE* of 2.14°C. It was also observed that 64% of estimation errors are less than 2°C and 39% are less than 1°C; and there is no systematic error of overestimation or underestimation. From these results we can infer that in most cases errors are consistent with their usability as inputs to environmental models. Low errors resulting from this validation procedure also suggest that the error propagation induced by STEP1 in STEP2 does not affect the overall results. In fact, errors due to the whole procedure are comparable to those of the STEP1 alone.

After that, a comparison with an alternative estimation standard method, commonly used in environmental modelling, was made in order to qualitatively estimate the efficiency of the empirical proposed method as for generalisation capability (i.e. spatial distribution). A spatial interpolation methodology, the Inverse Distance Weighting method (IDW), integrated with a vertical lapse rate (0.6°C / 100 m), was used to alternatively estimate daily mean air temperature on the basis of only ground data. In particular, the IDW procedure estimates the daily mean air temperature in a certain site as a particular function of the daily mean values measured by the nearest selected stations and of their distance from the site.

Table 1: Results of the validation of the satellite procedure for deriving daily mean air temperature, on the basis of 120 data collected by 10 different stations. Actually, only 100 of 120 data were suitable for the validation analysis because of the cloud coverage.

Station	**# data**	***RMSE***
Alla Braccia	11	2.42
Aprica	9	2.09
Carona ARPA	11	2.02
Funivia Bernina	11	1.82
Lanzada 2	11	2.22
Lanzada 5	11	1.67
Lenna	10	2.99
Monte Masuccio	9	3.69
Semogo	9	1.44
Val Torreggio	8	2.83
Total	**100**	**2.14**

Then, by a comparison of the results with the actual mean air temperature measured by the test station, it was possible to verify the result accuracy. The comparison was conducted on one test station (*Val Pola* station, 2,330 m), by considering for the Inverse Distance Weighting method a number of ten near weather stations. This choice was forced by the fact that we decided to compare the two methodologies in a well-monitored area; thus, under the best conditions for IDW methodology.

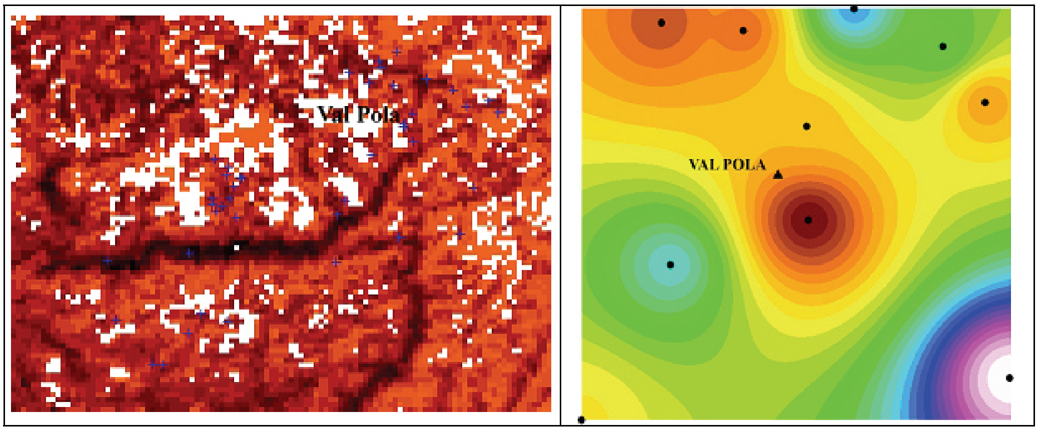

Figure 5: Map of instantaneous day-time LST temperature from MOD11 (on the left) and map of daily mean air temperature from IDW method (on the right) around Val Pola weather station (29 June 2003).

Also in this case, the statistical sample was composed of two values for each month, referring to the different moment of the satellite overpass (half day-time and half night-time), and the test data were excluded from the correlation analysis of STEP1 and STEP2.

For the remote sensing procedure, in STEP1 the generalised correlation of high altitude was used to evaluate T_{air}, from *LST*, and, as a consequence, the relationship referring to the highest altitude range was used in STEP2 to derive the T_{air} mean value from the T_{air} inst. This allows the behaviour of the two general relationships to be verified together, which is more appropriate to our purpose.

The analysis showed that the two contending methodologies lead to similar results (Table 2): in four cases, the satellite-derived *LST* procedure is outperforming, in three cases Inverse Distance is more accurate and in the remaining five cases the errors are similar. This result is very encouraging, since it demonstrates that this remote sensing method is comparable with a standard spatialisation method even in a very well-monitored area. On the other hand, this method is the only one usable in un-gauged areas. Moreover, by a deeper analysis considering also the acquisition moment, we can notice that in all night-time simulations, the remote sensing methodology is better or similar than Inverse Distance, while it is less performing during day-time. This is a direct consequence of the previous results of STEP1: the punctual relationships of all the stations are better in night-time than in daytime.

Finally, a statistical analysis of the results was made in order to identify the best method for this case study. While Inverse Distance method shows a *RMSE* of 2.23°C, the remote sensing procedure shows a *RMSE* of only 1.89°C: In this case study better results are therefore achieved by processing remotely sensed data. Figure 6 shows the analysis results in terms of errors and differences between actual measured temperatures and estimated temperatures from both methods. At present, this is just a qualitative evaluation and a quantitative evaluation cannot be made, mainly because of the fairly small number of data analysed. In fact, the mean of estimation errors and their standard deviations calculated for the remote sensing method are equal to 1.4 and 1.0°C, respectively, while for the IDW, they are equal to 1.9 and 0.8°C. The closeness of the mean values, besides the high values of the standard deviations, does not point to a significant prevalence of one method and does not allow us to rank the two methods.

Table 2: Results of comparison between satellite and spatial interpolation methods on the basis of 12 dates for Val Pola station.

	T mean from satellite (°C)	*T* mean from spatial interpolation (°C)	*T* measured (°C)	Error from satellite (°C)	Error from spatial interpolation (°C)
11 Jan 2003 21:30h	-9.7	-12.5	-9.7	0.0	-2.8
18 Jan 2003 10:25h	-6.9	-8.9	-7.9	1.0	-1.0
21 Feb 2003 10:15h	-4.0	-7.8	-5.0	1.0	-2.8
28 Feb 2003 21:30h	-4.0	-6.0	-3.0	-1.0	-3.0
09 Mar 2003 21:20h	-1.4	-3.0	-2.1	0.7	-0.9
14 Mar 2003 10:30h	-3.2	-4.7	-6.7	3.5	2.0
24 Apr 2003 21:35h	3.4	3.6	2.0	1.4	1.6
24 Apr 2003 10:25h	-0.3	3.6	2.0	-2.3	1.6
03 May 2003 21:30h	3.4	3.2	2.4	1.0	0.8
15 May 2003 10:45h	1.0	-0.1	-1.9	2.9	1.8
12 Jun 2003 20:40h	12.5	12.6	14.3	-1.8	-1.7
29 Jun 2003 10:15h	13.1	10.6	13.5	-0.4	-2.9
RMSE				**1.89**	**2.23**

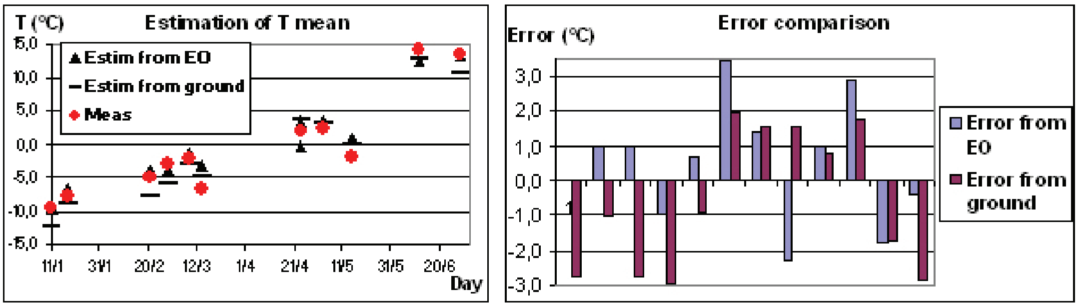

Figure 6: Comparison of the two estimation methods (on the left) and their errors (on the right).

CONCLUSIONS

This study demonstrates the validity of remotely sensed *LST* (MOD11_L2 product) as a source for calculating daily mean air temperature. The proposed procedure solves, by empirical approaches, the problems of relating *LST* to T_{air} and T_{air} inst to T_{air} mean, exploiting ground data weather station measurements as a reference. The relationship between *LST* and T_{air} is determined by a correlation analysis and a generalised procedure (STEP1). The extrapolation of daily mean *T* values from instantaneous *T* values is addressed again by correlation analyses taking into account the altitude variability and by processing historical series (STEP2). Validation was accomplished by accuracy assessment procedures both punctual and spatially distributed, the latter by comparison with the Inverse Distance Weighting interpolation method.

Results are satisfactory since the obtained *RMSE*, approximately 2°C, can be considered acceptable for the purpose of providing daily mean air temperature values as input for distributed and semi-distributed models. The suitability for the purpose relies not only on the accuracy but mainly on the possibility to directly derive a map of daily mean air temperatures at 1 km resolution.

Nevertheless, this error can hardly be reduced, since the punctual relationships between *LST* and T_{air} estimated for each station in STEP1 on the basis of hundreds data showed an error of about 1-2°C.

The generalized correlations between T_{air} and *LST*, as assessed and validated for high altitudes, demonstrates a good generalisation capability (considering also stations far from each other) and acquires much importance by considering that, generally, very few stations are located above 2,000 meters and there is a gap of information regarding high elevated areas. Therefore, availability of such a relation enables mean air temperatures to be estimated also in a usually un-gauged altitude range, solving one of the most common problems of hydrology and climatology of the Alps. Moreover, the calibration of this kind of relationship between T_{air} and LST for high altitude zones does not require a dense meteorological network, but only a high number of data that could be collected by several stations – not necessarily close - or by few stations with historical series.

Further developments will consider mainly three topics. At first, the quantitative evaluation of the difference between the proposed methodology and traditional spatial interpolation methods, as well as of the influence of the weather station network density on this difference, in particular to understand, if there is a network density value below which the use of satellite data is advisable. The second development will regard the sensitivity analysis of environmental and hydrological models to such inputs, in order to evaluate errors induced by the assimilation of remotely sensed data. Another development will consider the introduction in the analyses of the MODIS imagery onboard the EOS-Aqua satellite, allowing of the acquisition of two more values of LST for each day. This will likely improve the procedure in terms of both accuracy and frequency of daily mean air temperature maps.

ACKNOWLEDGEMENTS

This study was co-funded by the EC in the framework of AWARE (A tool for monitoring and forecasting Available WAter REsource in mountain environment), a STREP Project in the FP6 GMES Initiative. The authors wish to thank ARPA Lombardia for all the data available from its network weather stations.

REFERENCES

1 Bergstrom S, 1976. Development and applications of a conceptual runoff model for Scandinavian catchments. Swedish Meteorological and Hydrological Institute, Report RHO, No. 7, Ph.D. Thesis (Norrköping, Sweden)

2 Martinec J, A Rango & R Roberts, 1994. The Snowmelt Runoff Model (SRM) User's Manual. Edited by M F Baumgartner, Geographica Bernensia, P29, Department of Geography, University of Berne (Berne, Switzerland) 65 pp.

3 Garand L, M Buehner & N Wagneur, 2003. Dynamic inference of background error correlation between surface skin and air temperature. In: XIII International TOVS Study Conference (Sainte Adèle, Canada) 5 pp.

4 Jin M, 2004. Analysis of land skin temperature using AVHRR observations. American Meteorological Society, 85(4): 587-600

5 Jin M & R E Dickinson, 1999. Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle 1. Without clouds. Journal of Geophysical Research, 104(D2): 2105-2116

6 Jin M, 2000. Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle 2. Cloudy-pixel treatment. Journal of Geophysical Research, 105(D3): 4061–4076

7 Jones P, G Jedlovec, R Suggs & S Haines, 2004. Using MODIS LST to estimate minimum air temperature at night. In: 13th Conference on Satellite Meteorology and Oceanography, (American Meteorological Society, Norfolk, Virginia) 6 pp.

8 Prihodko L & N Goward, 1997. Estimation of air temperature from remotely sensed surface observations. Remote Sensing of Environment, 60: 335-346

9 Prince S D, S J Goetz, R O Dubayah, K P Czajkowski & M Thawley, 1998. Inference of surface and air temperature, atmospheric precipitable water and vapor pressure deficit using Advanced Very High-Resolution Radiometer satellite observations: comparison with field observations. Journal of Hydrology, 212-213: 230-249

10 De Wit A J W, H L Boogaard & C A Van Diepen, 2004. Using NOAA–AVHRR estimates of land surface temperature for regional agrometeorological modelling. International Journal of Applied Earth Observation and Geoinformation, 5(3): 187–204

11 Wan Z, 1999. MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD), Version 3.3, NASA contract NAS5-31370 (Institute for Computational Earth System Science)

12 Wan Z & Z L Li, 1997. A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. IEEE Transactions on Geoscience and Remote Sensing, 35: 980-996

13 Petitcolin F & E Vermote, 2002. Land surface reflectance, emissivity and temperature from MODIS middle and thermal infrared data. Remote Sensing of Environment, 83: 112-134

14 Ignatov A & G Gutman, 1998. Diurnal cycles of land surface temperature. Advances in Space Research, 22: 641-644