New Developments and Challenges in Remote Sensing, Z. Bochenek (ed.)
 ISBN 978-90-5966-053-3

Estimation of daily mean air temperature from MODIS LST in Alpine areas

A. Colombi, M. Pepe & A. Rampini
IREA–CNR, Via Bassini 15, 20133 Milano, Italy–email: colombi.a@irea.cnr.it

C. De Michele
DIIAR, Politecnico di Milano, P.zza Leonardo da Vinci 32, 20133 Milano, Italy

Keywords: LST, air temperature, MODIS, weather station, spatialization

ABSTRACT: This study is aimed to demonstrate the feasibility of MODIS LST products as a source for calculating the spatially distributed daily mean air temperature to be used as input for hydrological or climate models. The test area is located in the Alpine area. The proposed procedure solves, by empirical approaches, the problems of relating LST to T_{air} and instantaneous T values to daily mean ones, exploiting ground data weather station measurements as a reference. The LST to T_{air} relationship is determined by correlation analysis and equation generalisation for spatial distribution. The extrapolation of daily mean values from instantaneous values is addressed again by correlation analyses taking into account the altitude variability and exploiting historical series. The validation was accomplished by accuracy assessment procedures both punctual and spatially distributed, the last performed by comparison with the Inverse Distance Weighting interpolation method.

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

1 INTRODUCTION

Air temperature near 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 hydrological and climate models. The measurement of this variable by meteorological stations, however, provides only punctual values, whereas most models require spatial-distributed variables and parameters to evaluate physical processes with an adequate spatial scale. In most cases, because of the strong relationship between air temperature and elevation, interpolation of point site data is carried out only by using a vertical temperature lapse rate. But, when considering a regional scale, it's necessary also to refer to spatial interpolation methods, allowing for land cover and topographic effects factors. Moreover, often, high-elevated areas are 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 ungauged areas.

In this context, thermal infrared satellite imagery represents an effective solution to obtain spatially distributed information. For that purpose, a semi-empirical method to estimate air temperature by using thermal infrared information is presented. Coarse/medium spatial resolution satellite sensors, observing all Earth surface twice a day, offer insights about thermal infrared radiation, from which it is possible to retrieve Land Surface Temperature (LST). Optical sensors provide surface averaged value of LST at the pixel resolution. The strong correlation between LST and air temperature, demonstrated also in previous works (Jones *et al.* 2004) and depending on local conditions, represents the basis of our analysis.

There is a major constraint in the use of optical thermal infrared remotely sensed observations which concerns cloud coverage: under cloud sky conditions it is impossible to retrieve any information. As compared to a weather station network that acts everyday, information can result incomplete, considering that most models require data on a daily basis for a long period.

In previous researches, similar analyses were carried out using a Temperature Vegetation indeX (TVX) technique (Prihodko & Goward 1994; Prince *et al.* 1996). This method is based on the assumption that the surface temperature of a closed canopy is equal to air temperature and so air temperature can be derived by solving the regression equation (LST = a + b * 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 air temperature in Alpine area, where high altitudes make unusable NDVI because of the presence of snow and bare ground. Moreover, by using 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 (Garand *et al.* 2003); in this work we deal with a regional scale for which a higher accuracy and resolution are needed.

2 MATERIALS AND 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 at a spatial resolution of 1 km. Its suitability for the purpose is based on the MODIS sensor spatial, temporal and spectral characteristics. These last allow the solution of the thermal infrared radiometric correction by means of the *MODIS Generalized Split-Window LST Algorithm* (Wan, Z. 1999). 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 adjustments in order to properly map daily mean air temperature, as required by most models. Since the satellite sensor gives information only about LST, the first step (hereafter STEP1) of analysis is the evaluation of relationships between LST and air temperature measured by ground weather stations. The second step of analysis (hereafter STEP2) concerns the evaluation of daily mean value of air temperature from instant values. In fact, unlike weather stations, EO data are acquired instantaneously twice a day, while often hydrological and climate models require a daily mean value. Nevertheless, polar orbit satellites overpass everyday the same area at approximately the

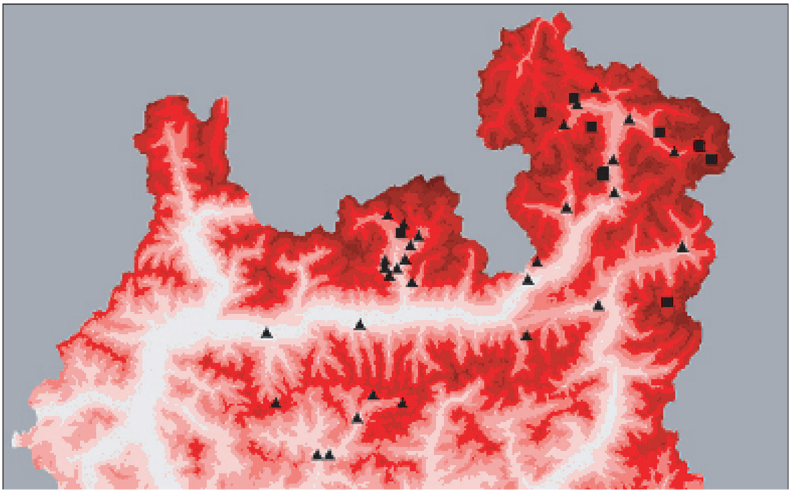

Figure 1. Study area. In evidence the weather stations (triangles) and the stations at high altitudes (squares).

same time, so data sets belonging to different dates are comparable. The study of both these relationships – STEP1 and 2 - permits to obtain input data for different hydrological and climate 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 Italian Alps, Lombardia district (Figure 1), where TERRA occurs daily at 10-11 a.m. and 9-10 p.m.

STEP1 consisted of a statistical method to retrieve a relationship between LST and air temperature during the period January–June 2003. The data set is composed by all passages of TERRA satellites in this period: 364 MODIS LST products (MOD11_L2 swath) corresponding to 182 days, together with contemporary ground data collected by 42 weather stations within the study area. From this data set, the clear sky control resulted in 124 usable dates (70% of whole); in the remaining dates only few pixel were sufficiently clear to be processed.

STEP2 of the procedure, dealing with the extrapolation of daily mean temperature from instant values, was obtained by statistical analyses of a data set composed by historical recordings of air temperature between 1996 and 2003, collected from 15 of the weather stations within the study area. The number of ground stations is a sub-set with respect to the ones of STEP1, as a consequence of the historical data availability.

3 STATISTICAL ANALYSES

3.1 *Air Temperature and LST*

STEP1 correlation analyses was applied: (1) at first separately to each station, to assess the feasibility of such a statistical approach, then (2) to the entire data set for identifying a typical equation representative of the area at hand, or for some sub-areas.

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

This analysis points out a strong linear correlation for all the selected stations (determination coefficient always higher than 0.75 and often higher than 0.9) and there are no evidences of dynamic depending on different months. Obtained relationships were compared with the 1:1 line, which represents $T_{air} = LST$. In 69 of the 84 cases the regression line has an angular coefficient less than 1 and intersects the 1:1 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 for low temperatures it is the opposite. Figure 2 shows this trend for one meteorological station (*Alla Braccia*).

Starting from these encouraging results in the second analysis some procedures of generalization were tested on the basis of the statistical sample and considering the literature, focusing the attention on the land cover, the altitude and the proximity of the available stations. These 3 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 the evaluation of a different behaviour (31 of 42 stations belong to the *Forest and seminatural area* CORINE Land Cover macrocategory).

The best results were achieved by considering altitude as the main important variable that affects this kind of relationship. In fact, on the assumption that stations at close altitude behave similarly, an estimation for high altitudes was found, using the 11 stations located at an altitude higher than 2,000 meters (6 stations over 182 days for calibration procedure and other 5 for validation on the same period). The outcomes clearly demonstrate the actual possibility to obtain two general relationships (day-time/night-time) as the altitude is considered the sub-setting criterion; and the more satisfactory results are obtained considering the night-time equation. In particular the determination coefficients 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:

$T_{air} = 0.649 * LST + 1.4036$ day-time

$T_{air} = 0.791 * LST + 2.7691$ night-time

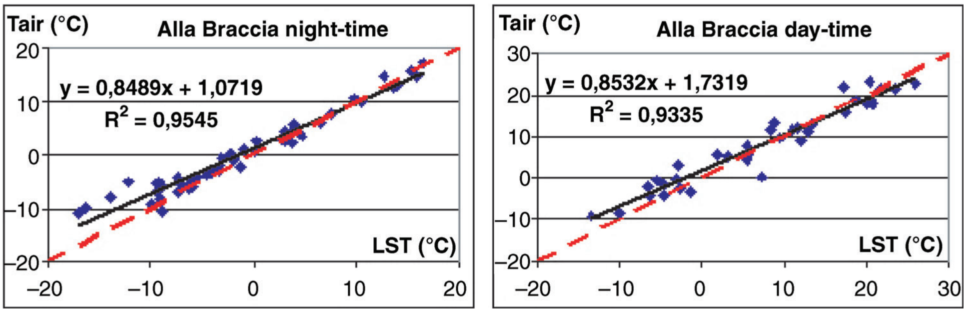

Figure 2. Relationships estimated for Alla Braccia station. For low temperatures LST results less than T_{air}, while for high temperatures LST is major.

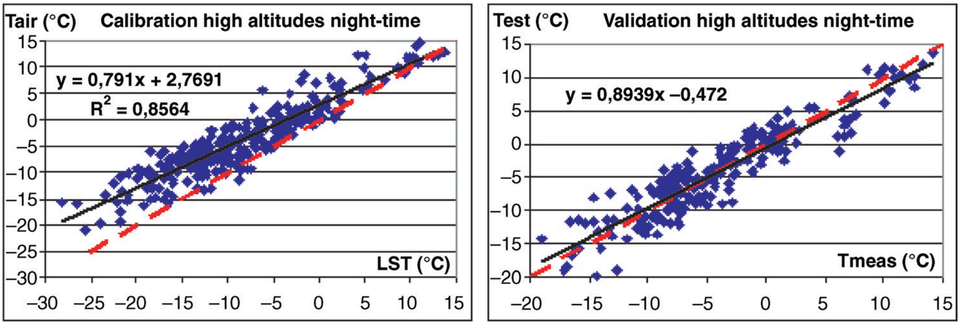

Figure 3. Relations estimated for high altitudes and night-time. The calibration (left figure) shows a R^2 of 0.86; from the validation (left figure) a regression line very close to the 1:1 line is obtained.

The last generalization procedure, based on the closeness of the selected stations, led to worse results, in terms of RMSE (between 3.14 and 4.64 °C).

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

3.2 *Instantaneous to daily mean temperature*

STEP2 of analysis regarded the extrapolation of daily mean air temperature from an instant value. The estimation of this kind of relationship is based on the existence of a standard air temperature diurnal cycle, characterized by a typical shape, dependent on several conditions, such as altitude, latitude and season. First of all some empirical

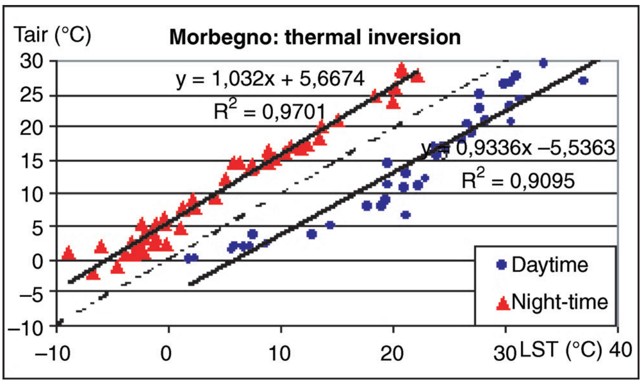

Figure 4. Thermal inversion observed at Morbegno weather station.

relationships that consider all these conditions were estimated. The main feature that clearly emerges is the strong linear correlation between instant and daily mean air temperature observed in all stations, hours and months. Then, 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, higher than 2,000 m) and the month (from January to June), with reference only to the satellite overpass hours, whereas latitude influence is negligible, due to the regional scale of the study. Then a number of 60 linear correlations was obtained, each characterized by high values of R^2 (from 0.74 to 0.93), each one resulting from the processing of several hundreds of ground data.

4 VALIDATION OF RESULTS

For evaluating together the STEP1 and STEP2 analyses, 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 one retrieved in STEP1 for each station) to retrieve the instant air temperature from LST. Then, according to the month and the altitude, the correlations found with STEP2 were applied, obtaining the final estimations of daily mean air temperature from the instant values. Therefore, by a comparison with the actual values of mean temperature measured by the ground stations, it was possible to verify the accuracy of the results.

The test data set is composed by 120 data, representing 12 days (two for each month, one day-time and one night-time) as acquired by 10 weather stations (lying at different altitudes in order to represent all the altitude ranges). Of course, the test data were excluded from the correlation analysis of STEP1 and STEP2.

The validation shows RMSE values ranging between 1.44 and 3.69 °C, according to the different stations considered, and a global RMSE of 2.14 °C, for 100 ground reference data[1] (see Table 1). 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 over or under estimation. By these results we can infer that in most cases errors are consistent with their use as inputs to climate or hydrological models. Modest 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, in order to qualitatively estimate the efficiency of the empirical method proposed as for generalisation capability (i.e. spatial distribution), a comparison to an alternative estimation method was made. A spatial interpolation methodology, the Inverse Distance Weighting method (IDW), was used to alternatively estimate the mean air temperature on the basis of ground data alone. Then, by a comparison with the mean temperature measures collected by test stations it was possible to verify the result accuracy.

[1]This subset of the 120 test data is due to cloud covered areas.

Table 1. Results of validation of satellite procedure for deriving daily mean air temperature, on the basis of 100 data collected by 10 different stations.

Station	N	RMSE (°C)
Alla Braccia	11	2.42
Aprica	9	2.09
Carona Arpa	11	2.02
Funivia Bernina	11	1.82
Lanzada 2/L	11	2.22
Lanzada 5/L	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**

The test was conducted on one test station (*Val Pola* station, 2,330 m) considering for the Inverse Distance Weighting method a number of 10 near weather stations. This choice was forced by the fact that test stations should be located in a well-monitored area and should belong to a high altitude (more than 2,000 m, see analysis of STEP1). Also in this case the statistical sample is composed by two values for each month, referring to the different hours (half day-time and half night-time), and the test data were excluded from the correlation analysis of STEP1 and STEP2.

The daily mean air temperature at *Val Pola* station was estimated, by the IDW method, as a particular function of the values measured by the near selected stations and their distance from it (Figure 5), taking into account the station altitude, by means of a vertical lapse rate of 0.6 °C / 100 m.

For the remote sensing procedure, in STEP1 the generalized correlation of high altitude was used to evaluate T_{air}, from LST, and, as a consequence, the relationship

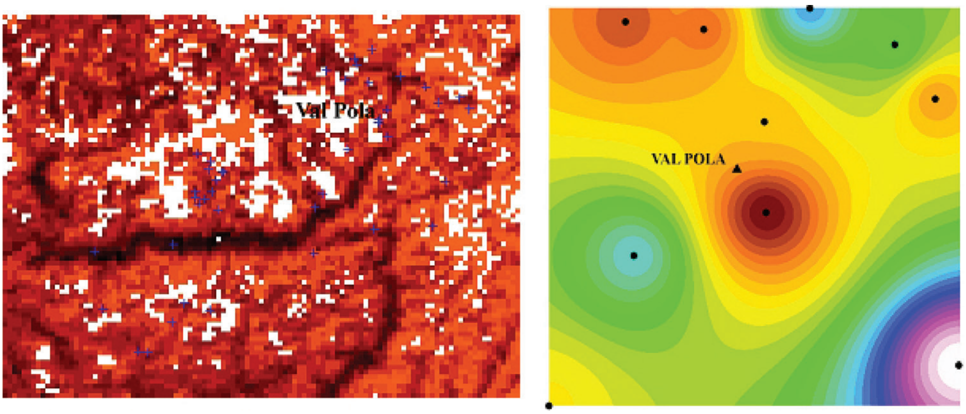

Figure 5. Map of instant day-time LST temperature from MOD11 (on the left) and map of daily mean air temperature from IDW method (on the right) around Val Pola weather station (29th of June 2003).

referring to the highest altitude range was used in STEP2 to derive the T_{air} mean value from the instant one. This allows verifying the behaviour of the two general relationships together, which is more related to our purpose.

The analysis shows that the two contender methods lead to similar results (see Table 2): in 4 cases remote sensing method is outperforming, in 3 cases Inverse Distance is more accurate and in the remaining 5 cases the errors are similar. This result is very encouraging since it demonstrates that this remote sensing method is comparable with a standard spatialization method even in a very well monitored area. On the other hand, this method resolves as the only one usable in an ungauged area. 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 during day-time it is less performing. This is a direct consequence of the previous results of STEP1: all the estimated relations are better in night-time than in day-time.

Finally a statistical analysis on 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 method 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

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 h 21:30	−9.7	−12.5	−9.7	0.0	−2.8
18-Jan 2003 h 10:25	−6.9	−8.9	−7.9	1.0	−1.0
21-Feb 2003 h 10:15	−4.0	−7.8	−5	1.0	−2.8
28-Feb 2003 h 21:30	−4.0	−6.0	−3	−1.0	−3.0
09-Mar 2003 h 21:20	−1.4	−3.0	−2.1	0.7	−0.9
14-Mar 2003 h 10:30	−3.2	−4.7	−6.7	3.5	2.0
24-Apr 2003 h 21:35	3.4	3.6	2	1.4	1.6
24-Apr 2003 h 10:25	−0.3	3.6	2	−2.3	1.6
03-May 2003 h 21:30	3.4	3.2	2.4	1.0	0.8
15-May 2003 h 10:45	1.0	−0.1	−1.9	2.9	1.8
12-Jun 2003 h 20:40	12.5	12.6	14.3	−1.8	−1.7
29-Jun 2003 h 10:15	13.1	10.6	13.5	−0.4	−2.9
RMSE				1.89	2.23

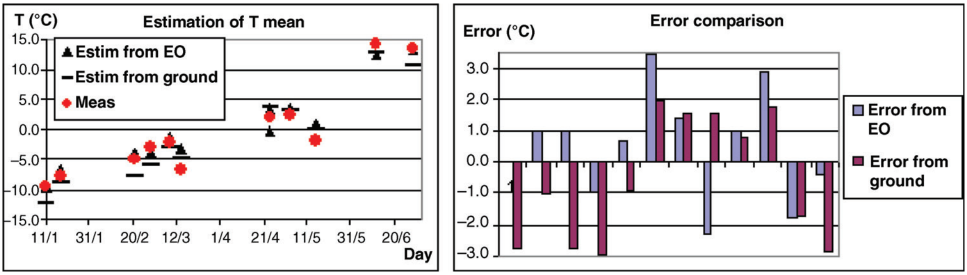

Figure 6. Comparison of the two estimation methods (on the left) and their errors (on the right).

and estimated temperature from both methods. At present, this is just a qualitative evaluation and a quantitative cannot be made, above all because of the fairly small number of data analysed. In fact, the mean of estimation errors and their standard deviation calculated for the remote sensing method are equal to 1.4 °C and 1 °C respectively, while for the IDW are equal to 1.9 °C and 0.8 °C. The closeness of the mean values, besides the high values of the standard deviations, does not provide for a significant prevalence of one method and does not allow a ranking of the two methods.

5 CONCLUSIONS

This study demonstrates the validity of remotely sensed LST (MODIS 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 instantaneous T values to daily mean ones, exploiting ground data weather station measurements as a reference. The LST to T_{air} relationship is determined by correlation analysis and equation generalisation (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 processing historical series (STEP2). The validation was accomplished by accuracy assessment procedures both punctual and spatially distributed, the last 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 temperature input values for distributed and semi-distributed models. The suitability for the purpose relies not only on the accuracy but mainly on the possibility to derive a map of daily mean temperature 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 general correlations between T_{air} and LST, as assessed and validated for high altitudes, demonstrate a good generalisation capability (considering also stations far from each other) and acquire 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 permits to estimate mean air temperature also in an usually ungauged altitude range, solving one of the most common problem of hydrology and climatology of the Alps.

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, 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 climate 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, allowing the acquisition of two more values of LST for each day. This will likely improve the procedure both in terms of accuracy and frequency of daily mean air temperature maps.

ACKNOWLEDGEMENTS

This study was co-funded by EC in the framework of AWARE (A tool for monitoring and forecasting Available WAter REsource in mountain environment), a STREP Project in the VI F.P., GMES Initiative.

The authors wish to thank ARPA Lombardia for all the available data from its network weather stations.

REFERENCES

De Wit, A.J.W., Boogaard, H.L. and Van Diepen, C.A. 2004. Using NOAA–AVHRR estimates of land surface temperature for regional agrometeorogical modelling. *International Journal of Applied Earth Observationand Geoinformation* 5: 187–204.

Garand, L., Buehner, M. and Wagneur, N. 2003. Dynamic Inference of Background Error Correlation between Surface Skin and Air Temperature. *XIII International TOVS Study Conference*.

Ignatov, A. and Gutman, G. 1998. Diurnal cycles of land surface temperature. *Adv. Space Res.* 22: 641–644.

Jin, M. 2004. Analysis of land skin temperature using AVHRR observations. *American Meteorological Society* 587–600.

Jin, M. and Dickinson, R.E. 1999. Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle 1.Without clouds. *Journal of Geophysical Research* 104: 2105–2116.

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, No. D3: 4061–4076.

Jones, P., Jedlovec, G., Suggs, R. and Haines, S. 2004. Using MODIS LST to estimate minimum air temperature at night. *13^{th} Conference on Satellite Meteorology and Oceanography.*

Petitcolin, F. and Vermote, E. 2002. Land surface reflectance, emissivity and temperature from MODIS middle and thermal infrared data. *Remote Sens. of Environ.* 83: 112–134.

Prihodko, L. and Goward, N. 1997. Estimation of Air Temperature from Remotely Sensed Surface Observations. *Remote Sens. Environ.* 60: 335–346.

Prince, S. D., Goetz, S. J., Dubayah, R.O., Czajkowski, K.P. and Thawley, M. 1996. 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. *Journ. Of Hydrol.* 212–213: 230–249.

Wan, Z. 1999 MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD), *NASA contract NAS5–31370*.

Wan, Z. and Li, Z.L. 1997. A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. *IEEE Trans. Geosci. Remote Sens.* 35: 980–996.