Air temperature estimation with MSG-SEVIRI data: Calibration and validation of the TVX algorithm for the Iberian Peninsula

Héctor Nieto [a,b,*], Inge Sandholt [b], Inmaculada Aguado [a], Emilio Chuvieco [a], Simon Stisen [c]

[a] Department of Geography, University of Alcalá, Colegios 2, 28801, Alcalá de Henares, Spain
[b] Department of Geography and Geology, University of Copenhagen, Øster Volgade 10, 1350, Copenhagen K, Denmark
[c] GEUS, Geological Survey of Denmark and Greenland, Øster Volgade 10, 1350, Copenhagen K, Denmark

ARTICLE INFO

ABSTRACT

Air temperature can be estimated from remote sensing by combining information in thermal infrared and optical wavelengths. The empirical TVX algorithm is based on an estimated linear relationship between observed Land Surface Temperature (LST) and a Spectral Vegetation Index (NDVI). Air temperature is assumed to be equal to the LST corresponding to the effective full vegetation cover, and is found by extrapolating the line to a maximum value of $NDVI_{max}$. The algorithm has been tested and reported in the literature previously. However, the effect of vegetation types and climates and the potential variation in NDVI of the effective full cover has not been subject for investigation. The present study proposes a novel methodology to estimate $NDVI_{max}$ that uses observed air temperature to calibrate the $NDVI_{max}$ for each vegetation type. To assess the validity of this methodology, we have compared the accuracy of estimates using the new $NDVI_{max}$ and the previous $NDVI_{max}$ that have been proposed in literature with MSG-SEVIRI images in Spain during the year 2005. In addition, a spatio-temporal assessment of residuals has been performed to evaluate the accuracy of retrievals in terms of daily and seasonal variation, land cover, landscape heterogeneity and topography. Results showed that the new calibrated $NDVI_{max}$ perform well, with a Mean Absolute Error ranging between 2.8 °C and 4 °C. In addition, vegetation-specific $NDVI_{max}$ improve the accuracy compared with a unique $NDVI_{max}$.

1. Introduction

Air temperature is one of the most important variables in climate research and global change. Air temperature controls many biological and physical processes between the Earth surface and the atmosphere, including respiration, transpiration and photosynthesis (Prihodko & Goward, 1997). Many land surface process models require air temperature as a driving variable. These models include global production models (Prince et al., 1998), Soil-Vegetation-Atmosphere-Transfer (SVAT) models (Olioso et al., 1999), and distributed hydrological models (Stisen et al., 2008a). In wildfire danger assessment air temperature is a key variable in the calculation of most fire meteorological danger indices (Bradshaw & Deeming, 1983; Keetch & Byram, 1968; Van Wagner, 1987), since it regulates the heat and vapor exchange between fuels and the surrounding atmosphere.

Air temperature has traditionally been measured at meteorological stations. However these measurements are not usually placed with enough spatial density for accurate climate research (Vogt et al., 1997; Willmott et al., 1991) and for application of spatially distributed models (Stisen et al., 2007). In addition, meteorological stations tend to be in strategic places to be operational and thus they do not optimally represent all the environments (Prihodko & Goward, 1997). The lack of sufficient spatial coverage and density can result in inaccurate estimates that affect the knowledge of global climate and the assessment of biophysical processes such as water and carbon fluxes (Prihodko & Goward, 1997). Remote sensing techniques may be a sound alternative to provide spatially distributed information, since they enable the assessment of environmental conditions in ecosystems (Czajkowski et al., 1997; Goward et al., 1994). Several studies have empirically derived the near surface air temperature by means of thermal remote sensing (Chokmani & Viau, 2006; Jang et al., 2004). However, applicability of empirical relationships, are limited by the difficulty of spatial and temporal generalization and the need of a training dataset (Stisen et al., 2007). For instance, Cresswell et al. (1999) applied empirical hour-specific models to estimate the air temperature from the Land Surface Temperature and the sun zenith angle. In another study, Vogt et al. (1997) used linear regressions between surface temperature

* Corresponding author. Department of Geography and Geology, University of Copenhagen, Øster Volgade 10, 1350, Copenhagen K, Denmark.
E-mail addresses: hn@geo.ku.dk (H. Nieto), is@geo.ku.dk (I. Sandholt), inmaculada.aguado@uah.es (I. Aguado), emilio.chuvieco@uah.es (E. Chuvieco), sst@geus.dk (S. Stisen).

and air temperature for data from 31 meteorological stations in Andalusia (Spain). They suggested that accuracies would be improved by grouping the stations by environmental characteristics and then calibrating the subgroups separately.

An interesting approach to estimate air temperature with remote sensing data is based on the correlation between Land Surface Temperature (LST) and Spectral Vegetation Index measurements (SVI) (Goward et al., 1994; Prihodko & Goward, 1997). This approach to estimate near surface air temperature, called the Temperature-Vegetation Index (TVX), was proposed by Nemani and Running (1989) and Goward et al. (1994) and it includes as additional variable the Normalized Difference Vegetation Index (NDVI). This method assumes that the temperature of a full cover canopy approaches the temperature of the air within the canopy (Czajkowski et al., 1997; Prihodko & Goward, 1997), even without evapotranspiration taking place (Goward et al., 1994; Prince et al., 1998). On the contrary, soils have a much greater mass per unit volume than vegetation canopies (Czajkowski et al., 1997; Goward et al., 1994; Prihodko & Goward, 1997) and consequently they are significantly hotter than canopies at the same meteorological forcing (Czajkowski et al., 2000). Besides, the thermal conductivity of canopies is not much larger than air, whereas soil thermal conductivity is one order of magnitude larger than in vegetation (Prihodko & Goward, 1997). Therefore, LST tends to approach air temperature with increasing SVI (Czajkowski et al., 1997; Goward et al., 1994). If an estimate of a full cover SVI is found, its corresponding LST will be near to the shelter height air temperature (Czajkowski et al., 2000; Prince et al., 1998).

One of the main assumptions of this approach is that both uniform atmospheric forcing and soil moisture conditions must take place (Stisen et al., 2007). This assumption is presumably fulfilled by applying this approach in a relatively small pixel array. However, this contextual approach might induce too low variations in the retrieved air temperature over the moving window (Stisen et al., 2007) and therefore spatial autocorrelation of air temperature is expected (Prince et al., 1998; Stisen et al., 2007). This method is more suitable in windows with heterogeneous covers (Czajkowski et al., 2000; Czajkowski et al., 1997) since a high variance in SVI within the moving window is desired to avoid noisy linear relationships (Czajkowski et al., 1997; Prince et al., 1998; Stisen et al., 2007). On the other hand, the landscape elements should have a similar size to the pixel resolution, due to the non-linearity of surface temperature (Stisen et al., 2007). This approach also assumes that there are no differences in aerodynamic resistance within the moving window due to different surface types (Stisen et al., 2007). It seems that the air temperature is estimated more accurately in surfaces closely coupled with the atmosphere (Prince et al., 1998). Differences of sun illumination caused by topography reduce the LST/SVI relation (Czajkowski et al., 1997) and shadowing caused by the topography is assumed to be negligible (Stisen et al., 2007). In addition, significant variations in altitude within the moving window lead to spatial variations in air temperature due to the adiabatic cooling and thus it compromises the assumption of spatially correlated air temperatures. Intra-pixel variability in altitude makes the estimation of the pixel averaged air temperature difficult. Timing is an additional factor that affects the estimates, since the algorithm requires that surface and air temperature are not out of phase (Stisen et al., 2007). In addition, since the algorithm requires as input a Spectral Vegetation Index, the algorithm is only suitable for daylight hours. The accuracy of the method has been shown to be variable according to the season of the year, Goward et al. (1994) observed higher residuals during winter, which were attributed to presence of snow cover beneath the canopy.

The vegetation index that has typically been plotted against the LST is the NDVI. If all previous assumptions are met, a linear relationship between the NDVI and LST within the window can be expressed by Eq. (1)

$$LST = a_{t,i} + b_{t,i} \cdot NDVI \tag{1}$$

where the subscripts t and i indicate that the relationship is time dependent and variable in each moving window i.

A critical issue in the TVX approach is the definition of the value for the full cover NDVI (hereafter called $NDVI_{max}$). Once a value for $NDVI_{max}$ has been established the air temperature can be estimated by extrapolating Eq. (1) to $NDVI_{max}$ (Eq. (2)):

$$T_{t,i} = a_{t,i} + b_{t,i} \cdot NDVI_{max} \tag{2}$$

where $T_{t,i}$ is the estimated air temperature of the pixel centered at the moving window i in time t.

Several values for $NDVI_{max}$ have been proposed in the literature. Prihodko and Goward (1997) measured the leaf reflectance and transmittance of different herbaceous species and applied a radiative transfer model to derive a $NDVI_{max}$ of 0.86 for AVHRR images. Czajkowski et al. (1997) and Czajkowski et al. (2000) reduced the $NDVI_{max}$ of Prihodko and Goward (1997) to 0.7 and 0.65 respectively in order to account for atmospheric effects. More recently, Stisen et al. (2007) adapted the TVX algorithm to be used with Meteosat Second Generation SEVIRI images in the Senegal River Basin using an $NDVI_{max}$ of 0.65. Their purpose was to obtain the air temperature at a higher temporal resolution and thus enable the use of these estimates as inputs for specific models such as Soil-Vegetation-Atmosphere-Transfer (SVAT) models. Stisen et al. (2007) pointed out that differences in channels and pixel resolution between sensors should make the $NDVI_{max}$ sensor specific. Besides, Prihodko and Goward (1997) and Stisen et al. (2007) stated the need to address the value of $NDVI_{max}$ in different vegetation types.

This study continues the work of Stisen et al. (2007) and its main objective is to propose a novel methodology to estimate the $NDVI_{max}$ based on measured air temperatures. The $NDVI_{max}$ was computed and validated for different vegetation types along the Iberian Peninsula. To assess the performance of this methodology, the new calibrated $NDVI_{max}$ values have been compared with previously proposed $NDVI_{max}$ over the same validation set, covering data during the year 2005. In addition we have evaluated the effects that vegetation type, topography, landscape homogeneity, and timing have on the accuracy of estimations.

2. Study area and data

2.1. Study area

The Iberian Peninsula (Fig. 1) is located at the south-west limit of the European continent. The West and part of the North are bounded by the Atlantic Ocean whereas the East coast is washed by the Mediterranean. According to the Köppen classification, three climates can be distinguished. The north coast has an Oceanic climate (Cfb), with mild temperatures and abundant rainfall with no dry season. The natural limit of this climate is the Cantabric Ranges. The dominant vegetation types are deciduous broadleaved and mixed forests with evergreen needleleaved forest. South of the Cantabric Ranges, the dominant climate is the Mediterranean (Csa), with a hot and dry season during summer and mild winters. In this zone we find mainly evergreen broadleaved and needleleaved forests. A third climate can be found, the Mid-latitude steppe (BSk), with very dry and hot summers and very cold winters, located in the interior of the Peninsula in plateau zones. The vegetation in this climate is very adapted to these extreme conditions, with shrublands and, sometimes, sparse trees.

2.2. MSG-SEVIRI images

SEVIRI sensor provides images every 15 min in 12 spectral channels covering both optical and thermal spectrum at a spatial sampling of 3 km at nadir (Schmetz et al., 2002). Visible, thermal, and the cloud mask images were acquired and preprocessed at the Department of Geography and Geology at the University of Copenhagen (Denmark).
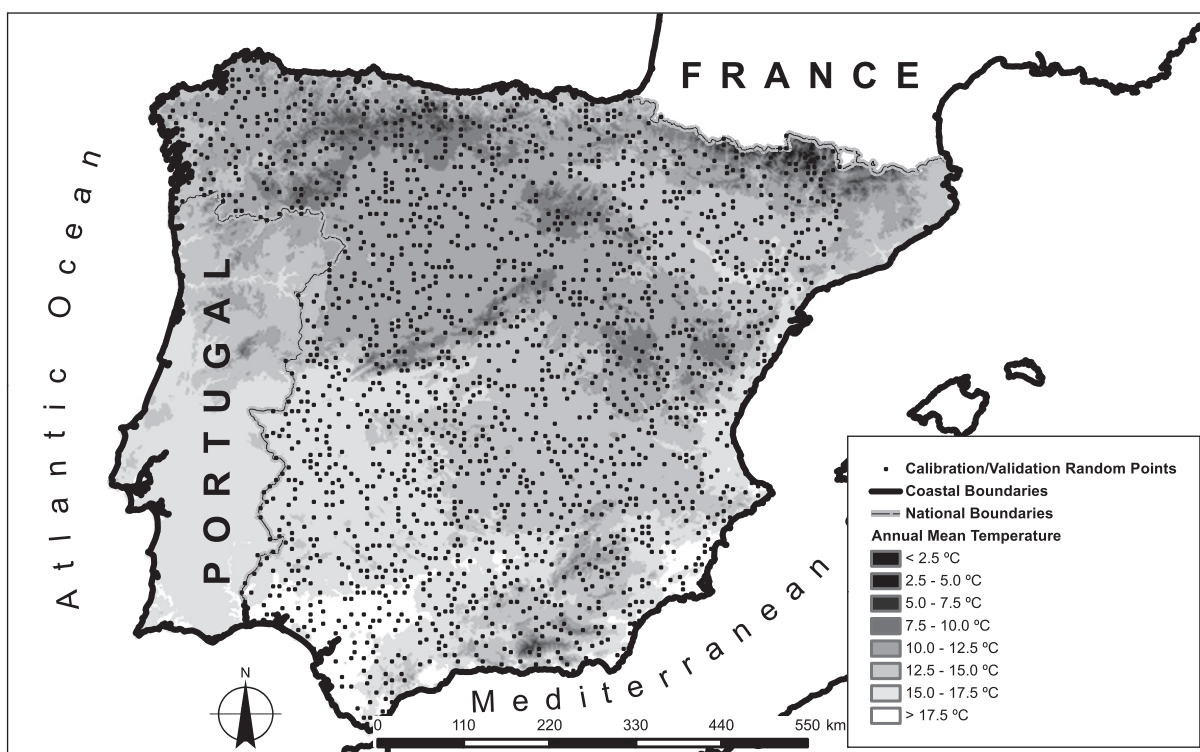
Fig. 1. The Iberian Peninsula and location of the 2000 random points used for calibration of validation of the algorithm. Source of annual mean temperatures: Digital Climatic Atlas of the Iberian Peninsula (Ninyerola et al., 2005).

The data consist of subsets of the Iberian Peninsula (upper left corner: 44.393968°N, 10.894700°W; lower right corner: 35.184872°N, 1.616629°E) spanning the year 2005 and resampled at a resolution of 4km. Atmospheric correction of visible channels was performed by using the Simplified Method for the Atmospheric Correction (SMAC) of satellite measurements in the solar spectrum (Rahman & Dedieu, 1994), more details of this correction process can be found in Fensholt et al. (2006) and in Stisen et al. (2007). Several additional products are provided by the EUMETCAST service, among these the cloud mask product was also acquired for the same period. The EUMETSAT cloud mask product is a simplification of the Cloud Analysis Image product which is based on a series of threshold tests to detect and characterize clouds (EUMETSAT, 2007).

2.3. Meteorological data

Two different datasets of air temperature were available for this study. 1) Hourly air temperature at an agrometeorological station, and 2) gridded interpolated daily maximum air temperature covering the Iberian Peninsula at 3×3 km resolution. The hourly dataset match accordingly to the MSG-SEVIRI temporal resolution enabling thus a temporal assessment of the proposed methodology, whereas the gridded daily data has been used to explore the spatial issues of the algorithm.

2.3.1. Cabañeros agrometeorological station

The agrometeorological station is located at the Cabañeros National Park (39.319758°N, 4.394824°W). Air temperature, as well as relative humidity, incoming solar radiation, wind speed and precipitation, were stored in an hourly and daily basis in a CR10X datalogger. The data are available since 1998, but a failure in sensors caused a loss of data during the year 2004. Only data from 2005 were thus used in this study. The station is located in flat terrain over a savanna with evergreen oak (*Quercus ilex* L.) and therefore shows the best conditions for the application of the TVX algorithm.

2.3.2. 3×3 km daily interpolated grid

This database is provided by the Spanish company Meteológica. It includes maximum and minimum daily temperatures from the reanalysis of the European Centre for Medium-range Weather Forecast, and interpolated at a spatial resolution of 3×3 km (Aguado et al., 2007). The database covers the years 2002 to 2005. To check the integrity of this dataset, the maximum interpolated daily air temperature was validated with daily data ranging the year 2005 from the Agro-climatic Information System for Irrigation (SIAR; Sistema de Información Agroclimática para el Regadío, http://www.mapa.es/siar/). This database comprises 361 automatic agrometeorological stations, which are mainly located over agricultural areas. Therefore, these stations do not cover the whole range of biomes found in the Iberian Peninsula. The results show that the interpolated data agree very closely with the observations during year 2005 (Fig. 2), with a Mean Absolute Error of 1.3 °C and Root Mean Square Error of 1.8 °C.

3. Methods

3.1. MSG-SEVIRI pre-processing

SMAC corrected radiances of channels 1 and 2, centered at 0.6 μm and 0.8 μm respectively, were used to compute NDVI values every 15 min (Fensholt et al., 2006). Land Surface Temperature was computed using the algorithm proposed by Sobrino and Romaguera (2004) for MSG-SEVIRI. This algorithm requires brightness temperatures and surface emissivity estimates for channels centered in 10.8 and 12 μm, as well as precipitable water vapor and view zenith angles. Precipitable water vapor was estimated using a split-window algorithm (Choudhury et al., 1995) originally developed for AVHRR data. The algorithm has been successfully applied for SEVIRI data (Stisen et al., 2007; Stisen, Sandholt, et al., 2008). Daily emissivity maps were produced combining an emissivity Look Up Table (Trigo et al., 2008) and daily fractional cover derived from NDVI values (Nishida et al., 2003). For this purpose, a daily NDVI was produced from the 15 min data by averaging all cloud free
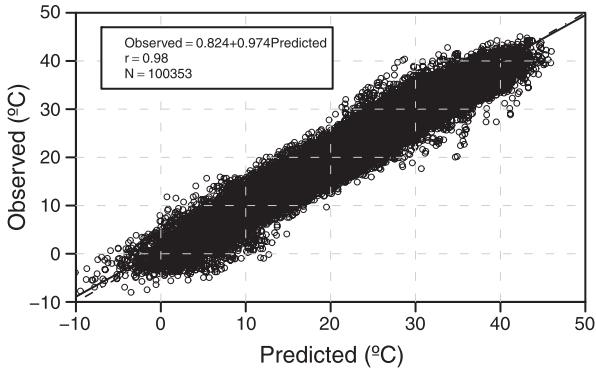
Observed = 0.824+0.974Predicted
r = 0.98
N = 100353

Fig. 2. Observed versus interpolated maximum daily values of air temperature in the SIAR network. Solid line represents the regression between observed versus predicted. Dashed line represents the 1:1 line.

observations between 8 and 17 h. This mean NDVI was smoothed using a Savitzky–Golay filter implemented in the TimeSat software (Jonsson & Eklundh, 2004) to reduce daily bidirectional effects since even daily

averages showed some noise due to the bidirectional reflectance. A complete description of this process can be found in Stisen et al. (2007). All the processing flow of the MSG processing is illustrated in Fig. 3.

3.2. Calibration of the TVX algorithm

The slope, intercept, and correlation coefficient between the NDVI and LST were computed in a 7×7 moving window. The size of this moving window responds to a trade-off between being big enough in order to obtain a reasonable number of valid observations (>30 cases), allowing thus a reliable fitting for Eq. (1); but small enough in order to fulfill the assumption of constant atmospheric forcing. This task was performed for every SEVIRI scene between 8 h and 17 h during the year 2005. This temporal window was chosen in order to assure daylight along the whole year. In order to improve the interpretation of the TVX relationship, factors of potential noise, such as clouds, water bodies and smoke, must be removed (Czajkowski et al., 1997; Prince et al., 1998). We have thus made use of the cloud mask, as well as several quality checks (Stisen et al., 2007), to flag contaminated pixels. The TVX algorithm was applied only to windows having more than 2/3 of non contaminated pixels.
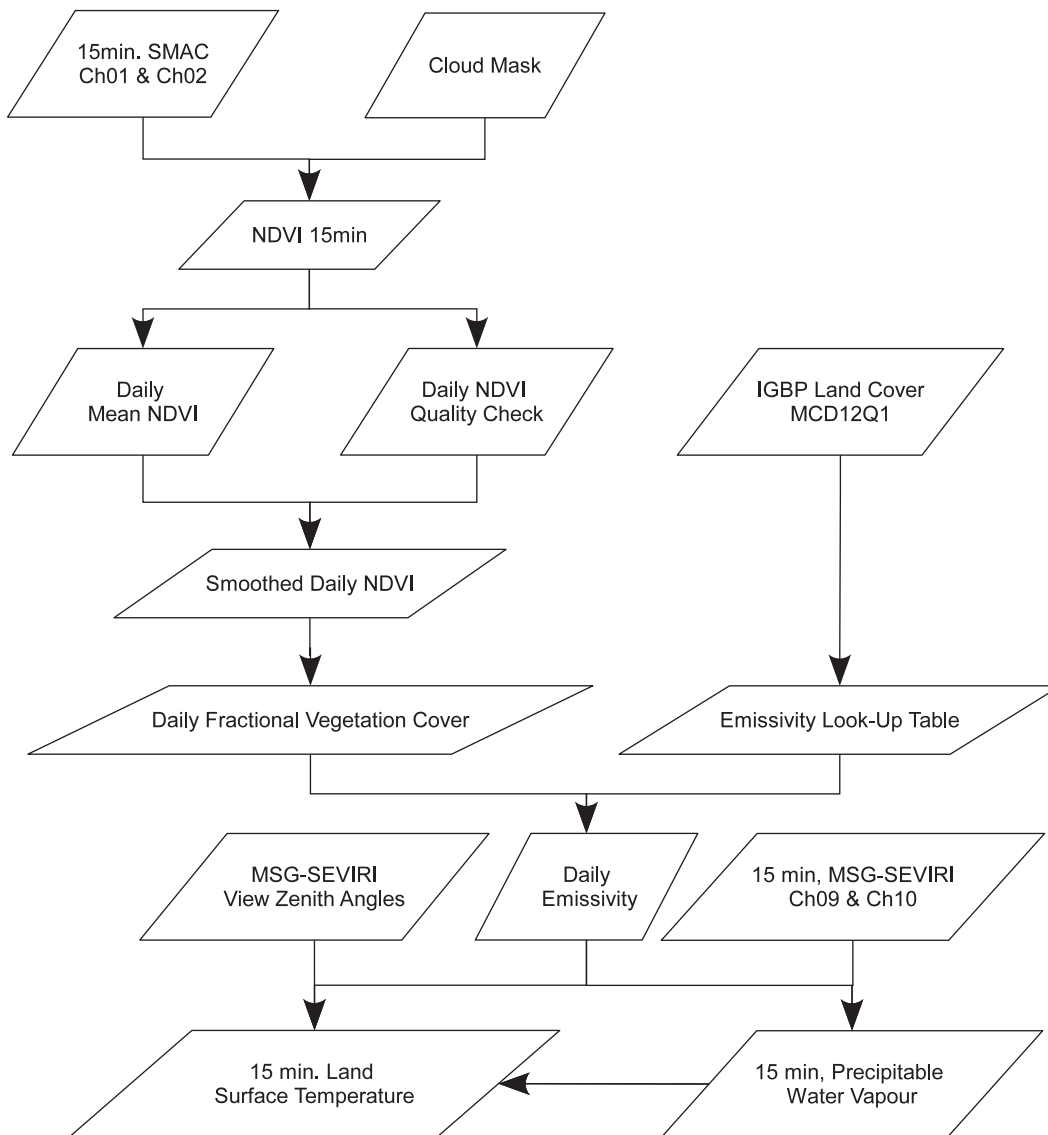
Fig. 3. MSG-SEVIRI processing flowchart.

In order to estimate $NDVI_{max}$ we can rewrite Eq. (2) as:

$$T_{t,i} - a_{t,i} = b_{t,i} \cdot NDVI_{max} \tag{3}$$

Knowing $T_{t,i}$ (the observed air temperature), $a_{t,i}$ and $b_{t,i}$, the only unknown variable is $NDVI_{max}$, which can be solved by the least squares method.

An a priori $NDVI_{max}$ must be assumed in order to compute a preliminary 15 min. SEVIRI air temperature and thus compare the observed hourly (daily) temperature with the retrieved 15 min. regression parameters. Every hour, the corresponding TVX parameters ($a_{t,i}$ and $b_{t,i}$) to the maximum retrieved air temperature are chosen to calibrate the new $NDVI_{max}$ with the observed hourly maximum temperature. Similarly, the corresponding TVX parameters to the maximum retrieved air temperature between 12 h and 17 h (the temporal window when it is expected to occur the maximum daily air temperature) were chosen to calibrate the $NDVI_{max}$ with the interpolated daily temperature. The a priori $NDVI_{max}$ was 0.65, which was initially proposed by Czajkowski et al. (2000) and gave good results in the Senegal River Basin study carried out by Stisen et al. (2007). In addition, only regression parameters corresponding to a good linear relationship between NDVI and LST were taken into account to perform the calibration. We have thus filtered all cases where the correlation coefficient between NDVI and LST is less than or equal to −0.95 (the relationship between the NDVI and the LST must be negative).

The $NDVI_{max}$ was calibrated for the hourly observations of the Cabañeros site, as well as for the daily interpolated temperature in the 3×3km UTM grid. The calibration and validation in this latter case was performed by sampling 2000 random points over all peninsular Spain along the whole year 2005. These 2000 locations will allow thus to cover the whole range of biomes in the Iberian Peninsula for the calibration/validation tasks. We have calibrated a unique $NDVI_{max}$ by using all points together, as well as an $NDVI_{max}$ for each biome by splitting the random points according to the MODIS IGBP Land Cover product, MCD12Q1 (Friedl et al., 2002).

3.3. Validation

Several statistics were computed in order to assess the accuracy of the retrieved air temperature. These error measurements have been calculated for an $NDVI_{max}$ of 0.65 (hereafter called NDVI$_{STI}$), the value proposed by Prihodko and Goward (1997) of 0.86 (hereafter called NDVI$_{PRI}$), and the new calibrated $NDVI_{max}$. Different measures were computed to assess the performance of models. According to Fox (1981) two types of measures can be computed to evaluate the performance of a model, measures based on residuals, and measures of correlation. The first set of measures represents a quantification of the difference between the observed and the predicted values. The second set represents the quantification of the agreement between predictions and observations. The following measures have been computed to perform the comparison of models: 1) mean of observed ($\bar{O}$) versus mean of predicted ($\bar{P}$), which shows us the mean bias between observed and predicted. 2) Standard deviation of observed versus standard deviation of predicted, which reflects the capacity of the model to explain the variability of the observed temperature. 3) The slope and intercept between the observed and predicted according to Piñeiro et al. (2008), under the hypothesis that a good model should follow the 1:1 line (slope unity and no intercept). 4) the index of agreement (d, Eq. (4)), which is more appropriate than the correlation index since the latter is not consistently related to the accuracy, in terms of magnitude, between the observed and predicted values (Willmott, 1982).

$$d = 1 - \frac{\sum (P_i - O_i)^2}{\sum (|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \tag{4}$$

5) The Root Mean Square Error (RMSE) and its systematic (RMSE$_s$) and unsystematic (RMSE$_u$) components. According to Willmott (1982), in a good model, the RMSE$_s$ should approach zero while the RMSE$_u$ approaches the RMSE. 6) The Mean Absolute Error (MAE), which is less sensitive to extreme values than RMSE. This latter tends to inflate when extreme values are present (Willmott, 1982). Willmott and Matsuura (2005) recommended to use this measure rather than RMSE since the latter is not a true measure of average error since it also measures the variability of the errors. And 7) the percentage of estimations that fall within $\pm 3\,°C$ and within $\pm 5\,°C$, since it is a measure that has been used in previous studies, and therefore it has been included for comparison purposes with other studies.

The final task of this study was to evaluate the spatio-temporal performance of the algorithm. We have compared the residuals in terms of daily and seasonal evolution, biome, topography, and landscape homogeneity. Topographic variations within the MSG-pixel and within the 7×7 moving window were evaluated. The standard deviation of altitudes, derived from a DTM at 250 m (source: National Cartographic Database), was computed at the extension of the SEVIRI pixels as well as in the 7×7 moving window. The landscape heterogeneity was evaluated with the Shannon's Entropy index h (Eq. (5) (Ricotta et al., 2006))

$$h = \sum_{i=1}^{n} p_i \cdot \ln p_i \tag{5}$$

where the relative abundance p_i of the class n_i is $p_i = n_i / \sum n_i$.

For each of the 2000 random points, this index was computed using a reclassification of the Corine Land Cover 2000 (European Environment Agency, 1999) into the IGBP land cover classes (Belward & Loveland, 1995). This source was chosen since it provides a higher spatial resolution than other global products such as MODIS Land Cover (with a resolution of 1000 m), and therefore heterogeneity is more accurately estimated at the SEVIRI pixel scale. Table 1 shows the total area covered by each class in Spain.

Analyses of Variance were performed to test whether time, season, and biome affects the residuals between observed and estimated air temperatures. The factor time was tested with the Cabañeros hourly data, whereas the other two factors, season and biome, were tested with the gridded interpolated daily data. In addition, Spearman correlation coefficient was computed between the residuals and the intra-pixel topographic variability, moving window topographic variability, and landscape heterogeneity.

Table 1
Surface occupied for each IGBP class in Spain.
Corine Land Cover 2000.

Code	Class	Area (ha)
1	Evergreen needleleaved forests	3,910,211
2	Evergreen broadleaved forests	1,437,917
4	Deciduous broadleaved forests	2,121,812
5	Mixed forests	1,498,125
6	Closed shrublands	10,380,236
7	Open shrublands	1425
8	Woody savannas	2,452,677
9	Savannas	2,616,773
10	Grasslands	620,147
11	Permanent wetlands	85,371
12	Croplands	16,955,950
13	Urban and built-up	1,020,885
14	Cropland/natural vegetation mosaic	5,060,491
15	Permanent snow and ice	302
16	Barren or sparsely vegetated	1,197,841

(a) $NDVI_{max}$ Calibration

$$T_{t,i} - a_{t,i} = 0.996 \times b_{t,i}$$
$$r = 0.898$$

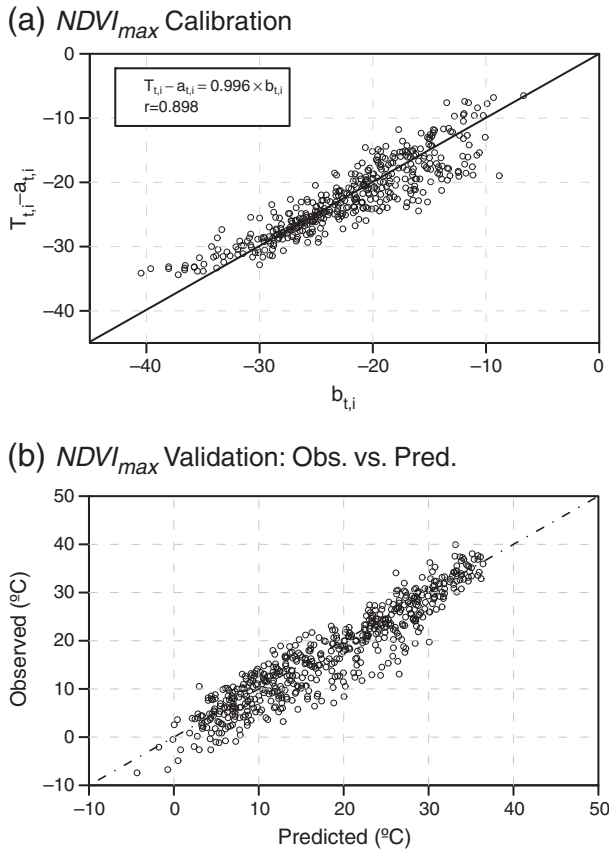## (b) $NDVI_{max}$ Validation: Obs. vs. Pred.

Fig. 4. (a) Calibration of maximum efficient NDVI corresponding to the hourly air temperature in Cabañeros. Points represent the observed values, the line corresponds to the calibrated $NDVI_{max} = 0.996$. (b) Validation of TVX air temperature. Observed versus predicted corresponding to the Cabañeros $NDVI_{max} = 0.996$. Dashed line represents the 1:1 line.

4. Results

4.1. Cabañeros data

A total of 982 valid observations centered in the meteorological station were obtained during the year 2005. For calibration purposes, 705 of these cases fulfilled the criteria of having a correlation coefficient less or equal than -0.95. Therefore 420 cases were finally selected for calibrating ($\approx 60\%$ of 705), leaving the remaining 562 observations for validation.

The calibrated $NDVI_{max}$ was 0.996 ($r = 0.898$, $p < 0.001$, Fig. 4). We performed then the validation of the new calibrated $NDVI_{max}$ (hereafter called $NDVI_{Cab}$) and compared the results with other previously proposed $NDVI_{max}$. The results in Table 2 show that the new $NDVI_{max}$ has a lower mean bias (1.0 °C), versus the ones obtained by using $NDVI_{PRI}$ and $NDVI_{STI}$ (3.2 °C and 6.5 °C, respectively). Similarly, MAE is significant lower with $NDVI_{Cab}$ (2.8 °C versus 3.8 °C and 6.7 °C).

Important as well is the decrease of the systematic $RMSE$ in relation to the total $RMSE$ that happens with $NDVI_{Cab}$ ($RMSE_s = 1.5$ °C, which represents the 41.50% of the total $RMSE$, compared to $NDVI_{PRI}$ and $NDVI_{STI}$, which represent the 70.44% and 86.99% of the total $RMSE$, respectively). On the other hand, the unsystematic $RMSE$ remains practically the same in all three $NDVI_{max}$ (3.4 °C, 3.3 °C and 3.7 °C). The index of agreement is significant lower in $NDVI_{STI}$ (0.88) compared to the $NDVI_{Cab}$ and $NDVI_{PRI}$ (0.96 and 0.95, respectively). Finally, a 62.5% of the estimates are within ± 3 °C with $NDVI_{Cab}$ whereas these proportion decreases in case of $NDVI_{PRI}$ and $NDVI_{STI}$ (44.7% and 17.8%, respectively).

4.2. Interpolated daily data

The total number of valid observations was 271,721 of which 54,763 had a valid correlation coefficient for calibration purposes. Taking the 60% of these latter observations resulted in a final dataset for calibration of 32,858, leaving the rest for validation (238,863).

The calibrated $NDVI_{max}$ for the whole dataset (hereafter called $NDVI_{All}$) was 0.818, $r = 0.756$ whereas the calibrated $NDVI_{max}$ for each of the IGBP classes (hereafter called $NDVI_{Var}$) are shown in Table 3. Validation results are shown in Table 4, all $NDVI_{max}$ tend to overestimate except $NDVI_{Cab}$, with a mean bias of -0.7 °C. From the $NDVI_{max}$ that overestimate, the lowest mean bias is achieved by $NDVI_{PRI}$ (0.9 °C), closely followed by $NDVI_{All}$ and $NDVI_{Var}$ (1.4 °C). Best MAE and $RMSE$ are obtained with $NDVI_{PRI}$ as well as with $NDVI_{All}$ and $NDVI_{Var}$ ($MAE = 3.9$ °C and $RMSE = 5.0$ °C). However, $NDVI_{PRI}$ has a lower systematic $RMSE$ (0.9 °C) compared to the other two (1.5 °C and 1.4 °C, respectively for $NDVI_{All}$ and $NDVI_{Var}$). Overall worst results are obtained with $NDVI_{STI}$ and $NDVI_{Cab}$. All $NDVI_{max}$ showed similar index of agreement (0.91–0.93) but the proportion of estimates within ± 3 °C is lower with $NDVI_{STI}$ (38.7%) compared to the other $NDVI_{max}$ (around 47%).

4.3. Spatio-temporal assessment

Fig. 5 shows the temporal dependence of residuals. We have compared the time dependence of residuals from $NDVI_{PRI}$ and the new calibrated $NDVI_{max}$ ($NDVI_{Cab}$ for hourly data, Fig. 5a; and $NDVI_{Var}$ for daily interpolated data, Fig. 5b). The results show that the greatest residuals are found in early hours in the morning and during spring and summer months, with both $NDVI_{max}$ showing a similar pattern.

Concerning the dependence of residuals to the vegetation type, different IGBP classes result in different residuals (Fig. 6a), where the highest bias are found in deciduous broadleaved and mixed forests, grasslands and mosaics of croplands and natural vegetation. These residuals decrease in case of using the calibrated NDVI for each IGBP class ($NDVI_{Var}$). Analysis of Variance (Table 5) showed significant interaction ($p < 0.001$) of these three factors in residuals.

Considering the topographic variability (within the pixel, Fig. 7a; and within the 7×7 moving window, Fig. 7b) as well as the landscape heterogeneity (Fig. 6b), the results show that increasing these factors overall increases the residuals. Both $NDVI_{PRI}$ and $NDVI_{Var}$ show the same behaviour, but $NDVI_{Var}$ presents a significant lower ($p < 0.01$) Spearman correlation coefficient (Table 6). This table also shows that

Table 2

Validation of hourly air temperature in the agrometeorological station of Cabañeros. Three different $NDVI_{max}$ have been compared: $NDVI_{STI}$ (0.65), $NDVI_{PRI}$ (0.86), and $NDVI_{Cab}$ (0.996). P_{mean}, mean of predicted values; O_{mean}, mean of observed values; s_p, standard deviation of predicted values; s_o, standard deviation of observed values; b, slope of the regression between observed and predicted values; a, intercept of the regression between observed and predicted values; MAE, Mean Absolute Error; RMSE, Root Mean Square Error; $RMSE_s$, systematic RMSE; $RMSE_u$, unsystematic RMSE; and d, index of agreement.

	N	P_{mean}	O_{mean}	s_p	s_o	(O vs. P)		MAE	RMSE	$RMSE_s$	$RMSE_u$	d	%	%
						b	a						≤ 3 °C	≤ 5 °C
$NDVI_{STI}$		23.6		11.2		0.85	-2.88	6.7	7.5	6.6	3.7	0.88	17.8	32.7
$NDVI_{PRI}$	562	20.3	17.1	10.1	10.0	0.94	-1.98	3.8	4.6	3.2	3.3	0.95	44.7	71.4
$NDVI_{Cab}$		18.1		10.0		0.98	-0.77	2.8	3.7	1.5	3.4	0.96	62.5	85.2

Table 3
Calibrated $NDVI_{max}$ for daily interpolated temperature corresponding to each of the IGBP biomes. N obs., number of cases for calibration/total number of cases; Obs. $NDVI_{99}$, 99th percentile of the observed NDVI; $NDVI_{max}$, calibrated maximum efficient NDVI; r, correlation coefficient.

Code	IGBP class	N obs	Obs. $NDVI_{99}$	$NDVI_{max}$	r
1	Evergreen needleleaved forest	1139/9160	0.896	0.849	0.78
2	Evergreen broadleave forest	111/3857	0.949	0.934	0.96
4	Deciduous broadleaved forest	70/262	0.658	1.162	0.80
5	Mixed forests	1032/7971	0.934	0.995	0.84
6	Closed shrublands	270/2559	0.752	0.915	0.72
7	Open shrublands	7373/69,757	0.821	0.803	0.77
8	Woody savannas	8239/59,344	0.852	0.835	0.79
9	Savannas	2039/18,216	0.851	0.777	0.76
10	Grasslands	116/1056	0.809	0.985	0.64
12	Croplands	11,633/94,425	0.849	0.800	0.71
13	Urban and built-up	678/3767	0.835	0.780	0.88
14	Cropland/natural vegetation mosaic	158/1359	0.946	0.937	0.93

the residuals are less related to spatial heterogeneity as well as to topographic heterogeneity with the use of NDVI$_{Var}$.

5. Discussion

The results show that overall the TVX algorithm performs adequately in the Mediterranean climate. We have achieved an *MAE* of 2.8 °C and *RMSE* of 3.7 °C with the hourly dataset, which is similar to the error of 2.96 °C found with the same sensor in Senegal (Stisen et al., 2007). However, an overestimation of 1.0 °C was found compared to the underestimation of 1.1 °C in Senegal. The daily maximum air temperature also showed a good accuracy with an *RMSE* = 5 °C and an overestimation of 1.0 °C to 1.5 °C, depending of the adopted $NDVI_{max}$. Similar results are found with AVHRR data by (Goward et al., 1994) in Oregon, *RMSE* = 5.4 °C; by (Czajkowski et al., 1997) in Canada, *RMSE* = 4.2 °C and bias of 3.2 °C; (Prihodko & Goward, 1997) in Kansas, *MAE* = 2.9 °C; and (Czajkowski et al., 2000) in Oklahoma, *RMSE* = 2.08 °C.

The new calibrated $NDVI_{max}$ showed a good performance, improving in general the results with other $NDVI_{max}$ proposed in the literature. In particular, NDVI$_{Cab}$, NDVI$_{All}$ and NDVI$_{Var}$ outperform the results obtained with the proposed $NDVI_{max}$ of Stisen et al. (2007) for MSG-SEVIRI. In both studies, SEVIRI scenes have been atmospherically corrected, but Stisen et al. (2007) adopted the value 0.65 assuming a residual atmospheric effect on the SEVIRI images. If that residual atmospheric effect is disregarded, the $NDVI_{max}$ should be obviously higher than 0.65 and closer to the value of 0.86 of Pridhodko and Goward (1997). On the other hand, Table 4 shows that NDVI$_{PRI}$ performs slightly better than the new $NDVI_{max}$. NDVI$_{All}$ and NDVI$_{Var}$ have similar error measurements and are close to NDVI$_{PRI}$. This is explained by the fact that most of the cases (85%) have an NDVI$_{Var}$ ranging between 0.80 and 0.85 (Table 3). Although NDVI$_{PRI}$ shows better error measurements than NDVI$_{Var}$, the distribution of residuals between IGBP classes is more homogeneous with NDVI$_{Var}$ (Fig. 6a),

considerably reducing the errors in classes such as evergreen and deciduous broadleaved forests, mixed forests, grasslands and mosaics. These categories have the highest $NDVI_{max}$ (above 0.93), but they only represent nearly the 5% of the total cases and thus they barely contribute to the global error measurements showed in Table 4.

At this point it is worth noting that the $NDVI_{max}$ for deciduous broadleaved forests present an unrealistic value (1.16). The purely empirical approach to derive the $NDVI_{max}$ together with the few valid cases available to calibrate this IGBP class (70 cases) can be responsible of this value. However, an analysis performed with the total number of cases for broadleaved forest (262 cases) lead to a similar result in the calibrated $NDVI_{max}$, with a value of 1.23. Table 3 also shows the 99th percentile of the measured NDVI as a measure of the maximum observed NDVI for each IGBP class. This statistic was used instead of the absolute maximum NDVI in order to avoid a possible influence of outliers due to noise effects. Overall, the observed maximum NDVI are in agreement with the estimated $NDVI_{max}$ with exception of broadleaved forests, closed shrublands and grasslands, with observed maximum NDVI sensible lower than the calibrated $NDVI_{max}$. These classes are precisely among the ones with fewer valid cases to perform the calibration of $NDVI_{max}$ and therefore these values should be taken with care. In addition, Prince et al. (1998) suggested a tendency to estimate *T* more accurately in areas where the vegetation is more closely coupled with the atmosphere. Therefore, another factor that could have influenced the anomalous $NDVI_{max}$ in broadleaved forest, is the leaf-in/leaf-out seasons that could contribute to noisy values. However, a calibration performed taking only into account spring and summer cases gave the same results. Finally, Stisen et al. (2007) pointed out that there is always a small canopy heat flux, even at maximum full cover, which induces the true air temperature to be below the canopy temperature. This minimum canopy heat flux is related to the minimum canopy resistance, which is directly proportional to the minimum leaf stomatal resistance and inversely to the Leaf Area Index (Monteith, 1973). In future research it would be interesting to explore the relationship between minimum canopy resistance and the retrieved $NDVI_{max}$.

Taking into account the unsystematic part of the *RMSE* it can be seen that independently of the adopted $NDVI_{max}$ this $RMSE_u$ is almost constant (3.5 °C and 5 °C for hourly and daily interpolated temperatures, respectively). According to Willmott (1982), these values of $RMSE_u$ are representative of the model noise level, and can be interpreted as a measure of potential accuracy. This noise can be caused by split-window estimates of LST — including satellite calibration, surface emissivity and precipitable water estimation — as well as radiometric errors in NDVI, residual cloud-contamination and the size of the contextual window (Czajkowski et al., 1997; Prihodko & Goward, 1997; Prince et al., 1998). The algorithm for LST retrieval proposed by Sobrino and Romaguera (2004) showed an accuracy of 3 °C in Senegal (Stisen et al., 2007). Surface emissivity retrieval has been pointed out to be a great source of error in LST estimation (Goward et al., 1994; Sobrino & Romaguera, 2004). On the other hand, Prihodko and Goward (1997) estimated that uncertainty in the measurements of NDVI with AVHRR images can result in air temperature errors of ± 4 °C. In addition errors in the observed air temperature can be added to this noise level (Czajkowski et al., 1997).

Table 4
Validation of daily interpolated air temperature. Five different $NDVI_{max}$ have been compared: NDVI$_{STI}$ (0.65), NDVI$_{PRI}$ (0.86), NDVI$_{Cab}$ (0.996), NDVI$_{All}$ (0.818), and NDVI$_{Var}$ (variable according to Table 3). P$_{mean}$, mean of predicted values; O$_{mean}$, mean of observed values; s$_p$, standard deviation of predicted values; s$_o$, standard deviation of observed values; b, slope of the regression between observed and predicted values; a, intercept of the regression between observed and predicted values; MAE, Mean Absolute Error; RMSE, Root Mean Square Error; RMSE$_s$, systematic RMSE; $RMSE_u$, unsystematic RMSE; and d, index of agreement.
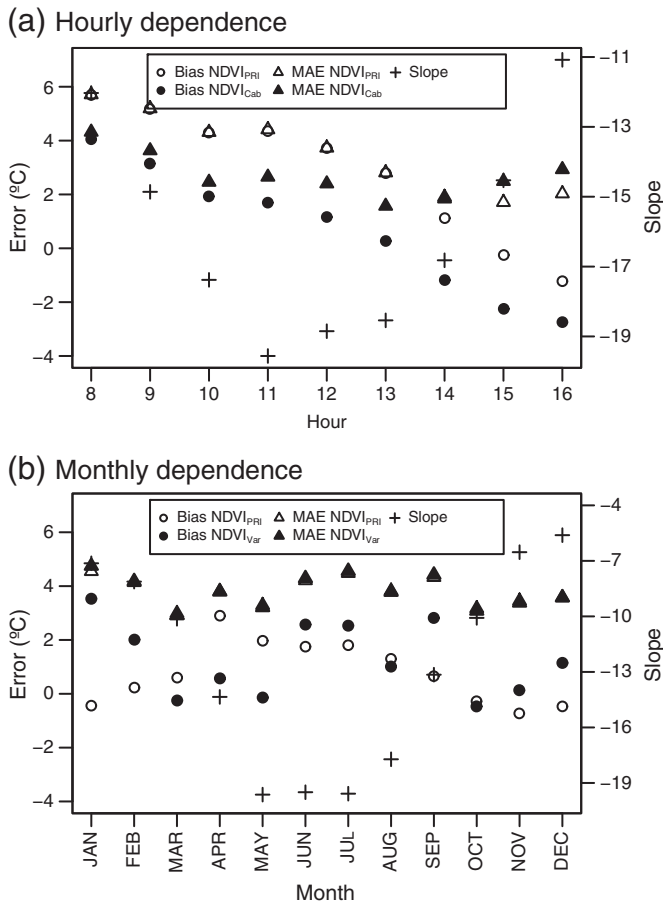
	N	P$_{mean}$	O$_{mean}$	s$_p$	s$_o$	(O vs. P)		MAE	RMSE	RMSE$_s$	RMSE$_u$	d	%	%
						b	a						≤3 °C	≤5 °C
NDVI$_{STI}$		26.0		11.0		0.74	3.39	4.8	6.0	3.5	4.8	0.91	38.7	59.9
NDVI$_{PRI}$		23.5		10.3		0.77	4.36	3.9	5.0	0.9	4.9	0.93	48.5	70.9
NDVI$_{Cab}$	238,863	21.9	22.6	10.0	9.1	0.78	5.53	4.0	5.2	0.9	5.1	0.92	46.4	69.0
NDVI$_{All}$		24.0		10.4		0.77	4.09	3.9	5.0	1.5	4.8	0.93	47.5	69.9
NDVI$_{All}$		24.0		10.4		0.77	4.08	3.9	5.0	1.4	4.8	0.93	48.1	70.4

(a) Hourly dependence

(b) Monthly dependence

Fig. 5. Temporal dependence of residuals. (a) Hourly dependence; (b) monthly dependence. For each factor two different $NDVI_{max}$ have been analysed in comparison with $NDVI_{PRI}$ (0.86): $NDVI_{Cab}$ (0.996), in a, and $NDVI_{Var}$ (IGBP variable NDVI), in b. The crosses represent the mean slope of the TVX relationship (secondary y-axis).

This is particularly important in our maximum daily air temperature, since they represent interpolated data and therefore are subject to model errors. Moreover, maximum air temperature estimated with MSG-SEVIRI is likely to be biased since the assessment of the daily temperature cycle will be subject to gaps due to the presence of clouds and failures to fulfill the quality criteria applied in the TVX algorithm.

In another study, Nieto et al. (2010) validated the algorithm using hourly ground data across different biomes using several agrometeorological stations in La Rioja (Spain). Their results showed a $MAE = 2.9\ °C$ and $RMSE = 3.5\ °C$, throwing some evidence on the actual accuracy of the algorithm. Further studies will thus be validated with ground data instead of modeled air temperature. Finally, as it was stated before, the size of the moving window was established according to a compromise between having enough valid observations within the window and assure the assumption of uniform atmospheric forcing following the work of Stisen et al. (2007). Future work will address the performance of the TVX algorithm for different moving window sizes.

An increasing bias in estimated air temperature was found with increasing pixel heterogeneity. This confirms the observations of Prince et al. (1998) and Czajkowski et al. (2000). However, accuracy, measured by the Mean Absolute Error showed no tendency related to landscape heterogeneity. On the other hand, TVX performance is clearly dependent on topographic variability within the pixel and within the moving window, as it was expected. The higher variation in altitudes within the pixel and within the moving window enhances the probability of different atmospheric forcing, caused either by differences in solar radiation and wind speed, as well as the adiabatic cooling of air with altitude. Both $NDVI_{max}$ that have been analyzed in this step showed similar patterns but $NDVI_{Var}$ tends to overestimate compared to $NDVI_{PRI}$.

Finally, a temporal dependence of residuals has been observed. In terms of the diurnal cycle, highest accuracies are found during the hours around midday (12 h–13 h), when the Land Surface Temperature reaches its maximum, and coinciding with the steepest slopes of the TVX relationship. Angular effects on observed NDVI and LST can be responsible of this trend (Prince et al., 1998; Stisen et al., 2007). Besides, Czajkowski et al. (1997) showed that the TVX slope decreases with high solar zenith angles and, therefore, the slope should be smaller during the morning and the afternoon. After noon, accuracy begins to decrease, probably due to the decoupling between the Land Surface Temperature and the air temperature (Stisen et al., 2007). It has been seen as well that bias decreases with time, being negative after noon. The underestimation has its maximum at 16–17 h, in coincidence with the air temperature peak of our study site.

On the other hand, annual variability showed that the lowest accuracy and the highest bias are found during spring and summer. The bias can be caused by the influence of physiological activity of vegetation during these months, since vegetation canopies always cause a small sensible heat flux, even at maximum canopy cover and moisture content (Stisen et al., 2007). Assuming that bulk surface resistance is inversely proportional to Leaf Area Index (Allen et al., 2006; Noilhan & Planton, 1989), the sensible heat flux will be negligible when LAI tends to infinite. In addition, meteorology

Table 5
Analysis of variance of residuals for $NDVI_{PRI}$ (0.86), $NDVI_{Cab}$ (0.996) and $NDVI_{Var}$ ($NDVI_{max}$ variable according to Table 3).

Factor	$NDVI_{max}$		Sum of squares	d.f.	Mean square	F	Sig.
Hour	$NDVI_{PRI}$	Between groups	4147.7	8	518.5	106.1	0.00
		Within groups	4756.2	973	4.9		
		Total	8903.9	981			
	$NDVI_{Cab}$	Between groups	4150.4	8	518.8	80.2	0.00
		Within groups	6296.7	973	6.5		
		Total	10,447.1	981			
Month	$NDVI_{PRI}$	Between groups	299,732.7	11	27,248.4	1213.8	0.00
		Within groups	6,099,498.0	271,709	22.4		
		Total	6,399,230.7	271,720			
	$NDVI_{Var}$	Between groups	420,770.9	11	38,251.9	1826.9	0.00
		Within groups	5,689,175.8	271,709	20.9		
		Total	6,109,946.7	271,720			
IGBP	$NDVI_{PRI}$	Between groups	116,314.7	11	10,574.1	457.3	0.00
		Within groups	6,282,916.0	271,709	23.1		
		Total	6,399,230.7	271,720			
	$NDVI_{Var}$	Between groups	12,421.2	11	1129.2	50.3	0.00
		Within groups	6,097,525.5	271,709	22.4		
		Total	6,109,946.7	271,720			

(a) Land cover dependence

(b) Landscape heterogeneity dependence
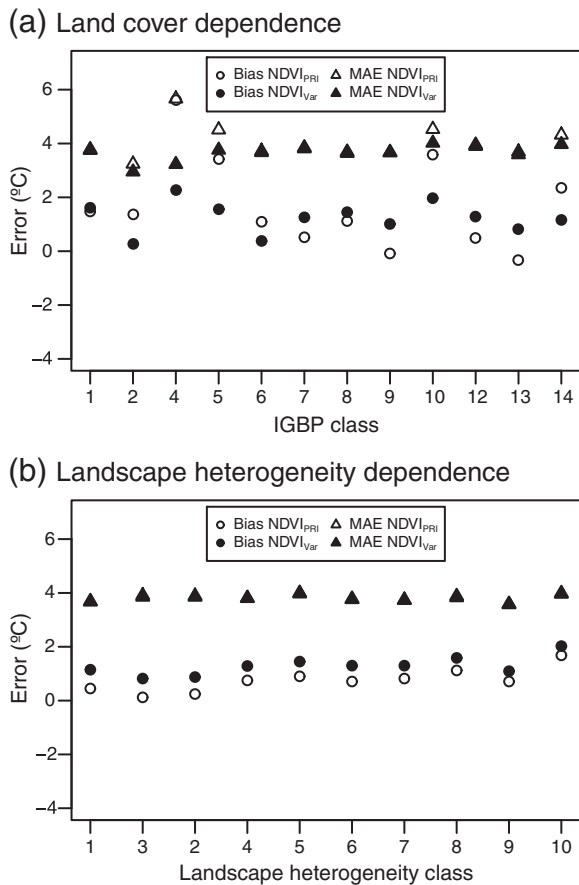
Fig. 6. Spatial dependence corresponding to variations in land cover and landscape heterogeneity. (a) IGBP classes; (b) equal size interval Shannon's *h* bins.

(a) Intra-pixel topographic dependence

(b) Moving window topographic dependence

Fig. 7. Spatial dependence corresponding to variations topography. (a) Equal size interval intra-pixel variability bins; (b) equal size interval moving window variability bins.

during the year 2005 in Spain was significantly anomalous, with warmer temperatures than usual during the spring months, April, May, and especially June (Instituto Nacional de Meteorología, 2005). Prihodko and Goward (1997) found that lower errors occur with decreasing temperatures and thus it is expected that better accuracies had been found in winter and autumn. Finally, that year was extremely dry in the Iberian Peninsula (491.4 mm accumulated rainfall in average), with historical minimums beginning in November 2004 until September 2005 (Instituto Nacional de Meteorología, 2005). Fig. 5.b shows that the highest monthly mean of the TVX slope are found from April until September. Since the slope of the TVX relationship can be related to surface wetness (Czajkowski et al., 1997; Goward et al., 1994; Prihodko & Goward, 1997), the temporal evolution of the slope is in agreement with the drought suffered in 2005.
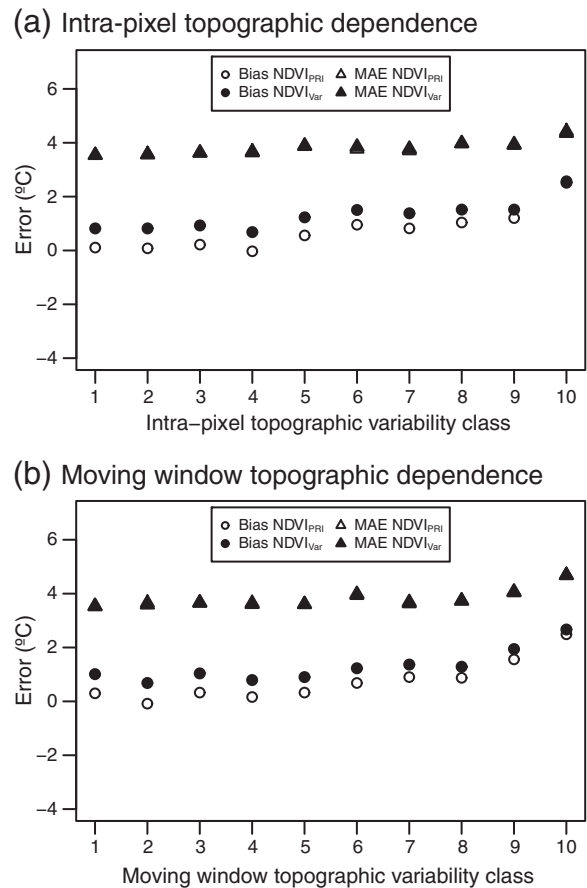
6. Conclusion

A methodology to estimate the full cover NDVI has been developed to better compute air temperature using the TVX algorithm. This empirical

Table 6
Spearman correlation coefficient between residuals and binned topographic variability (intra-pixel and moving window) and Shannon's *h* index of entropy. Significant correlations ($p < 0.01$) are marked with an asterisk (*).

	$NDVI_{PRI}$		$NDVI_{Var}$	
	Residual	Abs. residual	Residual	Abs. residual
STD Pixel	0.14*	0.06*	0.09*	0.06*
STD Window	0.13*	0.07*	0.10*	0.07*
h Shannon	0.07*	0.00	0.05*	0.00

approach only needs ground measurements of air temperature and the retrieved parameters (slope and intercept) of the TVX relationship. This allows us to easily obtain a different $NDVI_{max}$ depending on the vegetation type. All retrieved $NDVI_{max}$ values were coherent with exception of the deciduous broadleaved forest value that had an unrealistic value of 1.2, probably due to insufficient data for calibration.

The retrieved air temperature can be used as an input for a large range of models such as energy and water flux models, ecosystem functioning, etc. Since the accuracy of the algorithm ranged between 3 °C and 5 °C, the error transmission of the retrieved temperature into a model depends largely on the model sensitivity to the air temperature. Therefore, in very sensitive models to air temperature, the TVX algorithm could not be accurate enough to be used instead of ground measurements. For instance, Nieto et al. (2010) estimated the Equilibrium Moisture Content (Simard, 1968; Van Wagner & Pickett, 1985) for wildfire danger assessment with MSG-SEVIRI data. The authors found biased estimates of the Equilibrium Moisture Content caused by an overestimation of the TVX air temperature.

The results showed similar performances compared to previous studies and previous $NDVI_{max}$ but the application of a $NDVI_{max}$ dependent with vegetation type improved the estimations in several land covers, homogenising thus the accuracy between vegetation types. Although we have used exactly the same methodology as Stisen et al. (2007) for image processing, great differences have been found in terms of the relationship between $NDVI_{max}$ and the residuals. Further research would address the viability of this approach at global scales, testing the variability of the $NDVI_{max}$ and the accuracy of the TVX algorithm between different climates.

Using the appropriate $NDVI_{max}$ most of the error to estimate the air temperature can be attributed to noise ($RMSE_u$ near 4 °C). Therefore,

future efforts should be focused on improving the input variables. Firstly, the performed calibration/validation with interpolated air temperature is subject to uncertainties because of inaccuracies in the weather model. Secondly, additional noise is caused by the satellite input data. Several factors affect the accuracy of Land Surface Temperature and NDVI and their correlation. A cloud screening is crucial for this method since it affects LST, NDVI and the regression between both (Prihodko & Goward, 1997; Prince et al., 1998). Atmospheric correction for aerosols and water vapor are needed for addressing the effects of the atmosphere on the radiometric signal. Emissivity estimation is a key factor for accurate retrievals of LST (Sobrino & Romaguera, 2004). Finally, if hourly variations of air temperature are required, it would be necessary to take into account the bidirectional effects on the NDVI as well as the diurnal cycle of temperature.

Acknowledgements

References

Aguado, I., Chuvieco, E., Borén, R., & Nieto, H. (2007). Estimation of dead fuel moisture content from meteorological data in Mediterranean areas. Applications in fire danger assessment. *International Journal of Wildland Fire*, 16, 390−397.

Allen, R. G., Pruitt, W. O., Wright, J. L., Howell, T. A., Ventura, F., Snyder, R., Itenfisu, D., Steduto, P., Berengena, J., Yrisarry, J. B., Smith, M., Pereira, L. S., Raes, D., Perrier, A., Alves, I., Walter, I., & Elliott, R. (2006). A recommendation on standardized surface resistance for hourly calculation of reference ETo by the FAO56 Penman-Monteith method. *Agricultural Water Management*, 81, 1−22.

Belward, A., & Loveland, T. R. (1995). The IGBP-DIS 1 km land cover project: Remote sensing in action. *21st Annual Conference of the Remote Sensing Society* (pp. 1099−1106). Southampton, UK..

Bradshaw, B. S., & Deeming, J. E. (1983). *The 1978 national fire danger rating system*. Technical Report USDA Forest Service Ogden, Utah: Technical documentation.

Chokmani, K., & Viau, A. A. (2006). Estimation of the air temperature and the vapour quantity in atmospheric water with the help of the AVHRR data of the NOAA. *Canadian Journal of Remote Sensing*, 32, 1−14.

Choudhury, B. J., Dorman, T. J., & Hsu, A. Y. (1995). Modeled and observed relations between the AVHRR split window temperature difference and atmospheric precipitable water over land surfaces. *Remote Sensing of Environment*, 51, 281−290.

Cresswell, M. P., Morse, A. P., Thomson, M. C., & Connor, S. J. (1999). Estimating surface air temperatures, from Meteosat land surface temperatures, using an empirical solar zenith angle model. *International Journal of Remote Sensing*, 20, 1125−1132.

Czajkowski, K. P., Goward, S. N., Stadler, S. J., & Walz, A. (2000). Thermal remote sensing of near surface environmental variables: Application over the Oklahoma Mesonet. *Professional Geographer*, 52, 345−357.

Czajkowski, K. P., Mulhern, T., Goward, S. N., Cihlar, J., Dubayah, R. O., & Prince, S. D. (1997). Biospheric environmental monitoring at BOREAS with AVHRR observations. *Journal of Geophysical Research, [Atmospheres]*, 102, 29,651−29,662.

EUMETSAT (2007). Cloud detection for MSG — Algorithm theoretical basis document. *Technical Report EUM/MET/REP/07/0132 EUMETSAT Darmstadt, Germany*.

European Environment Agency (1999). CORINE land cover 2000. Technical guide. *Technical Report European Commission, Joint Research Centre*.

Fensholt, R., Sandholt, I., Stisen, S., & Tucker, C. (2006). Analysing NDVI for the African continent using the geostationary Meteosat Second Generation SEVIRI sensor. *Remote Sensing of Environment*, 101, 212−229.

Fox, D. G. (1981). Judging air quality model performance. *Bulletin of the American Meteorological Society*, 62, 599−609.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., & Schaaf, C. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, 83, 287−302.

Goward, S. N., Waring, R. H., Dye, D. G., & Yang, J. L. (1994). Ecological remote-sensing at OTTER: Satellite macroscale observations. *Ecological Applications*, 4, 322−343.

Instituto Nacional de Meteorología (2005). Resumen Anual Climatológico del Año 2005. *Technical Report Ministerio de Medio Ambiente Madrid*.

Jang, J. D., Viau, A. A., & Anctil, F. (2004). Neural network estimation of air temperatures from AVHRR data. *International Journal of Remote Sensing*, 25, 4541−4554.

Jonsson, P., & Eklundh, L. (2004). TIMESAT — A program for analyzing timeseries of satellite sensor data. *Computers and Geosciences*, 30, 833−845.

Keetch, J., & Byram, G. (1968). A drought index for forest fire control. *Technical Report USDA Forest Service*. Asheville, NC: Southeastern Forest Experiment Station.

Monteith, J. L. (1973). *Principles of environmental physics*. London: Edward Arnold.

Nemani, R. R., & Running, J. W. (1989). Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data. *Journal of Applied Meteorology*, 28, 276−284.

Nieto, H., Aguado, I., Chuvieco, E., & Sandholt, I. (2010). Dead fuel moisture estimation with MSG-SEVIRI data. Retrieval of meteorological data for the calculation of the equilibrium moisture content. *Agricultural and Forest Meteorology*, 150, 861−870.

Ninyerola, M., Pons, X., & Roure, J. M. (2005). *Atlas climático digital de la Península Ibérica. Metodología y aplicaciones en bioclimatología y geobotánica. Bellaterra*. Spain: Universidad Autónoma de Barcelona (http://www.opengis.uab.es/wms/iberia/english/en cartografia.htm).

Nishida, K., Nemani, R, Running, S., & Glassy, J. (2003). An operational remote sensing algorithm of land surface evaporation. *Journal of Geophysical Research, [Atmospheres]*, 108, D4270.

Noilhan, J., & Planton, S. (1989). A simple parameterization of land surface processes for meteorological models. *Monthly Weather Review*, 117, 536−549.

Olioso, A., Chauki, H., Courault, D., & Wigneron, J. P. (1999). Estimation of evapotranspiration and photosynthesis by assimilation of remote sensing data into SVAT models. *Remote Sensing of Environment*, 68, 341−356.

Piñeiro, G., Perelman, S., Guerschman, J. P., & Paruelo, J. M. (2008). How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecological Modelling*, 216, 316−322.

Prihodko, L., & Goward, S. N. (1997). Estimation of air temperature from remotely sensed surface observations. *Remote Sensing of Environment*, 60, 335−346.

Prince, S. D., Goetz, S. J., Dubayah, R. O., Czajkowski, K. P., & Thawley, M. (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*, 213, 230−249.

Rahman, H., & Dedieu, G. (1994). SMAC — A simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing*, 15, 123−143.

Ricotta, C., Corona, P., Marchetti, M., & Chirici, G. (2006). On parametric fragmentation measures. *European Journal of Forest Research*, 125, 441−444.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., & Ratier, A. (2002). An introduction to Meteosat Second Generation (MSG). *Bulletin of the American Meteorological Society*, 83, 977−992.

Simard, A. (1968). *The moisture content of forest fuels — A review of the basic concepts*. Forest Service Ottawa, Ontario: Technical Report USDA.

Sobrino, J. A., & Romaguera, M. (2004). Land surface temperature retrieval from MSG1-SEVIRI data. *Remote Sensing of Environment*, 92, 247−254.

Stisen, S., Jensen, K. H., Sandholt, I., & Grimes, D. I. F. (2008). A remote sensing driven distributed hydrological model of the Senegal River basin. *Journal of Hydrology*, 354, 131−148.

Stisen, S., Sandholt, I., Norgaard, A., Fensholt, R., & Eklundh, L. (2007). Estimation of diurnal air temperature using MSG SEVIRI data in West Africa. *Remote Sensing of Environment*, 110, 262−274.

Stisen, S., Sandholt, I., Norgaard, A., Fensholt, R., & Jensen, K. H. (2008). Combining the triangle method with thermal inertia to estimate regional evapotranspiration — Applied to MSG-SEVIRI data in the Senegal River basin. *Remote Sensing of Environment*, 112, 1242−1255.

Trigo, I. F., Peres, L. F., DaCamara, C. C., & Freitas, S. C. (2008). Thermal land surface emissivity retrieved from SEVIRI/Meteosat. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 307−315.

Van Wagner, C. E. (1987). Development and structure of the Canadian Forest Fire Weather Index System. *Technical Report Canadian Forest Service Otawa.*.

Van Wagner, C. E., & Pickett, T. L. (1985). Equations and FORTRAN program for the Canadian Forest Fire Weather Index System. *Technical Report Canadian Forest Service Ottawa*.

Vogt, J. V., Viau, A. A., & Paquet, F. (1997). Mapping regional air temperature fields using satellite-derived surface skin temperatures. *International Journal of Climatology*, 17, 1559−1579.

Willmott, C. J. (1982). Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society*, 63, 1309−1313.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, 79−82.

Willmott, C. J., Robeson, S. M., & Feddema, J. J. (1991). Influence of spatially-variable instrument networks on climatic averages. *Geophysical Research Letters*, 18, 2249−2251.