Recibido: 04/12/2022
Aceptado: 25/01/2023

REVISTA DE TELEDETECCIÓN
Asociación Española de Teledetección
(2023) 61, 59-71
ISSN 1133-0953
EISSN 1988-8740
https://doi.org/10.4995/raet.2023.18909

A simple method for the estimation of minimum and maximum air temperature monthly mean maps using MODIS images in the region of Murcia, Spain

Alfonso Galdón-Ruíz[1], Guillermo Fuentes-Jaque[2,3], Jesús Soto[4], Luis Morales-Salinas*[2]

[1] *Computer and Environmental Engineering Technology Doctorate Program, Universidad Católica de Murcia (UCAM), Murcia, Spain.*

[2] *Laboratory for Research in Environmental Sciences (LARES), Faculty of Agricultural Sciences, University of Chile, Santiago, Chile.*

[3] *Master in Territorial Management of Natural Resources, Post-graduate School, Faculty of Agricultural Sciences, University of Chile, Santiago, Chile.*

[4] *IT Department, Universidad Católica de Murcia (UCAM), Murcia, Spain.*

Abstract: Air temperature records are acquired by networks of weather stations which may be several kilometres apart. In complex topographies the representativeness of a meteorological station may be diminished in relation to a flatter valley, and the nearest station may have no relation to a place located near it. The present study shows a simple method to estimate the spatial distribution of minimum and maximum air temperatures from MODIS land surface temperature (LST) and normalized difference vegetation index (NDVI) images. Indeed, there is a strong correlation between MODIS day and night LST products and air temperature records from meteorological stations, which is obtained by using geographically weighted regression equations, and reliable results are found. Then, the results allow to spatially interpolate the coefficients of the local regressions using altitude and NDVI as descriptor variables, to obtain maps of the whole region for minimum and maximum air temperature. Most of the meteorological stations show air temperature estimates that do not have significant differences compared to the measured values. The results showed that the regression coefficients for the selected locations are strong for the correlations between minimum temperature with LST_{night} (R^2 = 0.69–0.82) and maximum temperature with LST_{day} (R^2 = 0.70–0.87) at the 47 stations. The root mean square errors (RMSE) of the statistical models are 1.0 °C and 0.8 °C for night and daytime temperatures, respectively. Furthermore, the association between each pair of data is significant at the 95% level (p<0.01).

Key words: MODIS, land surface temperature, topoclimate, spatial regression models, geographically weighted regression, geostatistical interpolations.

Un método simple para la estimación de los mapas medios mensuales de temperaturas mínimas y máximas del aire utilizando imágenes MODIS en la región de Murcia, España

Resumen: Los registros de temperatura del aire son adquiridos por redes de estaciones meteorológicas las cuales podrían estar alejadas varios kilómetros entre sí. En topografías complejas la representatividad de una estación meteorológica podría verse disminuida en relación con un valle más plano, y la estación más cercana no

To cite this article: Galdón-Ruíz, A., Fuentes-Jarque, G., Soto, J., Morales-Salinas, L. 2023. A simple method for the estimation of minimum and maximum air temperature monthly mean maps using MODIS images in the region of Murcia, Spain. *Revista de Teledetección*, *61*, 59-71. https://doi.org/10.4995/raet.2023.18909

* *Corresponding author: lmorales@uchile.cl*

tener relación con un lugar ubicado cerca de ella. El presente estudio, muestra un método simple para estimar la distribución espacial de las temperaturas mínimas y máximas del aire a partir de imágenes MODIS de temperatura de la superficie terrestre (LST) y el índice de vegetación de diferencia normalizada (NDVI). En efecto, existe una fuerte correlación entre los productos LST día y noche MODIS y los registros de temperatura del aire de las estaciones meteorológicas, lo que se obtiene al usar ecuaciones de regresión ponderadas geográficamente, encontrándose resultados confiables. Luego, los resultados permiten interpolar espacialmente los coeficientes de las regresiones locales usando como variable descriptora la altitud y el NDVI, para obtener mapas de la región completa para la temperatura del aire mínima y máxima. La mayoría de las estaciones meteorológicas muestran estimaciones de temperatura del aire que no tienen diferencias significativas en comparación con los valores medidos. Los resultados mostraron que los coeficientes de regresión para las ubicaciones seleccionadas son fuertes para las correlaciones entre temperatura mínima con LST_{noche} (R^2 = 0,69–0,82) y temperatura máxima con $LST_{día}$ (R^2 = 0,70–0,87) en las 47 estaciones. Los errores cuadráticos medios (RMSE) de los modelos estadísticos son 1,0 °C y 0,80 °C para las temperaturas nocturna y diurna, respectivamente. Además, la asociación entre cada par de datos es significativa al nivel del 95% (p<0,01).

Palabras clave: MODIS, temperatura de la superficie terrestre, topoclimatología, modelos de regresión espacial, regresiones ponderadas geográficamente, interpolación geoestadística.

1. Introduction

Air temperature (T_a) describes the land's surface environmental conditions (Prihodko & Goward 1997) and is one of the most important climate variables in the study of terrain conditions. Climate risk models need daily T_a estimations with a considerable spatial resolution in order to provide more accurate conclusions. The Fireglove fire risk assessment project used an operation method developed by Recondo et al. (2013) in which the T_a was obtained with a spatial resolution of 1 km^2 based on MODIS data and information from 331 weather stations throughout the Iberian Peninsula.

Weather stations' measurements are highly accurate, however their distribution in an area is limited and it is not easy to extrapolate values to zones in which on-site measurement is not possible, and so we resort to spatial interpolation methods based on geostatistics like kriging, or spatial regression models such as Ordinary Least Square (OLS) and Geographically Weighted Regression (GWR), which allow us, based on their numerical algorithm, to estimate T_a in spots that do not have weather station registries (Montaner-Fernández et al., 2020). Certain areas have low weather station density, mainly mountain zones, which hinders the attainment of dense T_a spatial estimates in those areas. Despite the previous limitations, it is possible to use remote sensing data, which allows for improvement in spatial estimation of various meteorological variables, particularly the T_a based on satellite-obtained Land Surface Temperature (LST) (Montaner-Fernández et al., 2020; Huang et al., 2017). Nevertheless, T_a calculation based on satellite data is limited by the temporal resolution of higher spatial resolution satellites, such as MODIS, LANDSAT or SENTINEL, given that the time intervals between each successive image are much too broad.

To cover for the sparse spatial coverage (weather station density), we use interpolation between known points methods to generate the thermal field in a determined area of earth. These interpolation methods are effective in areas close to the weather stations, but present limitations in more remote areas (Herrera et al., 2016). The interpolation process errors reported in scientific literature generally range from 1 to 3 °C (Willmott & Robenson, 1995; Vogt et al., 1997).

In studies that have compiled land surface temperature (LST) time series data with weather station data to estimate the air temperature at the average monthly level, local regressions have been used. For example, through geographic weighted regressions (GWR) results have been obtained, using terrain elevation as independent variable, with values of R^2>0.91 and RMSE = 1.1-1.5 °C in comparison with the estimate obtained using Ordinary Least Square (OLS) (Yao & Zhang, 2013).

In the Iberian Peninsula there are studies available on the air temperature estimation derived from land surface temperature measured by the MODIS satellite by Recondo et al. (2011), who used the algorithms proposed by Sobrino et al. (2003) with $R^2 = 0.86$-0.88 and residual standard error (RSE) = 2.5-2.7 °C, but if variables such as the total atmospheric column water vapor content (W), the normalized difference vegetation index (NDVI), the year's Julian day, the elevation and the incline, are included, results improve with an R^2 of 0.92 and a RSE between 1.9 and 2.1 °C.

In the Cataluña region, northeast of the Iberian Peninsula, Spain, we also have at our disposition studies conducted by Cristóbal et al. (2008) based on MODIS data for the 2000 to 2005 period whose monthly model results range between 0.65 and 0.94 for R^2 with an RMSE average of 1.0 °C.

Additionally, other methods have been suggested by other authors to estimate air temperature through satellite data and auxiliary variables like predictors, using machine learning algorithms such as Support Vector Machines and Random Forest. However, these are not approached in this article (Ruiz-Álvarez et al., 2019; Otgonbayar et al., 2019; Marzban et al., 2018).

The general objective of this study is developing and implementing a simple method, based on spatially explicit linear regressions, that allows to obtain an estimate of the daily minimum and maximum air temperature level using information provided by weather stations and MODIS satellite images in the visible, near infrared and far infrared (thermal) spectrums.

2. Materials and methods

2.1. Study area

The Murcia region (MR) is located in the south of the European continent, southeast of the Iberian Peninsula between the 37° and 38° north latitude parallels and the 0° and 2° east longitude meridians. It is an area near subtropical latitudes, which conditions the regional climate. It possesses a complex and diverse landform, with coasts, plateaus, mountains, and valleys that create great landscape variety. The mountain landform usually surpasses altitudes of 1000 meters (Sierra de la Pila, Ricote, Carrascoy and other mountains) and even 2000 meters at Pico Revolcadores. There are hollows situated between 500 to 1000 meters high such as the Yecla-Jumilla Plateau to the northeast and the San Juan Valley to the northwest (Climate Atlas of the Murcia region, http://hdl.handle.net/20.500.11765/13220). In the central area we have the Río Segura Valley from Cieza to Murcia and coast-adjacent plains, all lower than 200 meters above mean sea level. Approximately, 27% of the Murcian territory corresponds to mountain landforms, 38% to intermontane depressions and corridors, and the leftover 35% to plains and plateaus. There is a Mediterranean climate, although there are important local contrasts and variations that make the Murcia region an ideal place to compare temperature trends and other climate parameters (Climate Atlas of the Murcia region, http://hdl.handle.net/20.500.11765/13220).

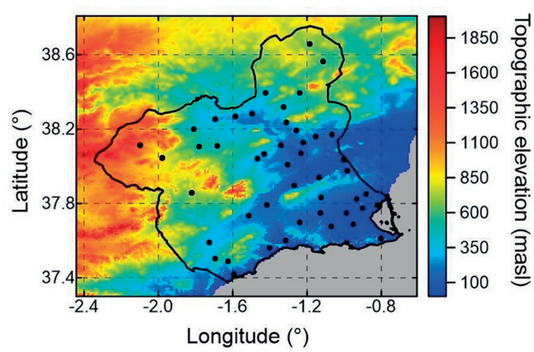

Figure 1. Study area and SIAM-IMIDA network composed of 47 weather stations, superimposed over the topography of the Murcia region and its surrounding areas, the Mar Menor and Mediterranean Sea (Mar Mayor) shown in grey.

2.2. Weather information

The utilized data correspond to the maximum temperature (°C) and minimum temperature (°C) in daily and mean monthly values, measured by sensors located in the 47 automatic weather stations (AWS) belonging to the Agrarian Information Service of Murcia (SIAM in its Spanish acronym) of the Murcian Institute of Agrarian and Alimentary Investigation and Development (IMIDA in its Spanish acronym), shown in Figure 1. The information was collected from hourly temperature data, where subsequently a database was created which allowed for verification of spatiotemporal variation between all the weather stations involved. Anomalous data was removed, and missing data were identified based on a systematic review of the data (data cleaning) and validation at several stages. The AENOR 500540 Spanish regulation was chosen, which in accordance with, consists of 7 kinds of consecutive weather data validations (Estévez et al., 2011). Thus, having the data filtered and the database systematized, we proceeded to fill in missing data using machine learning tools, specifically, the Support-Vector Machine (SVM) (Cortes & Vapnik, 1995; Zhao et al., 2006; Karatzoglou et al., 2006). To complete the missing data and outliers in the time series, some predictor variables were used, such as: (a) Altitude; (b) Location of the weather stations in the study area; (c) Month of the year; (d) Distance between stations; and (e) The 7 nearby stations that had at least 95% of available data.

For each of the 47 meteorological stations in the MR, different levels of validation were used on the meteorological information (AENOR 500540) to finally have a homogeneous database. During the validation process, 32008 records were calculated, equivalent to 9.3% of the total number of records in all the weather stations used.

2.3. Digital elevation model

Digital Elevation Models (DEMs) used correspond to the product ASTER GDEM, which is a global coverage grid DEM, with a spatial resolution of approximately 30 square meters depending on the location of each pixel.

2.4. Satellite data

The satellite data used were obtained from the TERRA and AQUA satellites and correspond to spatial information collected by the MODIS (Moderate Resolution Imaging Spectroradiometer) sensor, which was used by NASA to develop 34 products derived from its collected data (http://reverb.echo.nasa.gov/reverb). Of these, this study used the MOD11A2, MOD13A2 and MOD13Q1 products, which correspond respectively to land surface temperature and vegetation indices. The products associated with vegetation indexes contain the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI), obtained every 16 days, among other images, while the product associated with surface temperature contains daytime and night-time Land Surface Temperature and Emissivity, obtained every 8 days. All of these are sourced from the Earthdata NASA website and have a spatial resolution of 1 kilometre (km), except for MOD13Q1 which has a spatial resolution of approximately 250 meters (m).

The images correspond to an extent covering the MR for a time interval of 19 years (2001 and 2019) which were downloaded from MODIS archives and further processed. These were adjusted according to the scale factor recommended by MODIS (Wan, 2008), to later be averaged at the monthly-mean level for each year. Finally, the time series monthly mean values are calculated, obtaining the climatological values.

It is of importance to note that the acquisition time of the study area's LST images includes pixels with up to 3 different hours during daytime and 3 different hours during night-time. The obtained LST data were stored in °C to finally put together a mosaic with each variable's images, cropping the area of interest and then projecting to the geographic reference system with Datum WGS8 in spherical coordinates.

2.5. Topoclimatic modeling

To study the spatial distribution of minimum and maximum air temperatures it is necessary to have periodic, spatially distributed temperature data. The integration of weather station data and satellite images allows calibration of the LST and air temperature registry through the adjustment of statistical models, and subsequently extends the estimation to the entire study area via spatially explicit models and geostatistics (Littmann, 2008).

Traditionally global regressions are used which seek to understand the spatial behaviour of a variable through a unique equation, however, this equation's coefficients do not vary spatially (Aparecido et al., 2022; Lorençone et al., 2022). This search is done with a methodology named distance Weighted Least Squares (WLS); these weighted nonnegative constants being a function between each point and the rest (Fotheringham et al., 2002). Fundamentally due to statistical fit-based methods' defining parameters showing spatial variability (parametric instabilities), is that the Geographically Weighted Regression (GWR) method is used.

These sorts of methods are useful for performing spatial regressions affected by the parametric instability phenomenon with good results and the ability to generate maps of the adjusted parameters at a global scale (Fotheringham et al., 2002; Draper & Smith, 1981). The GWR method combines surface observation data (Automatic meteorological stations data) and the predictions to find a linear function that relates them, where the coefficients varying in the space adjusting to local effects (Equation 1). This method allows to calculate the coefficients of a multiple linear regression at each measurement point and to obtain the associated local errors. It is of importance to note that the coefficients calculated using GWR method must be spatially interpolated to obtain the selected variable's predictions on a full study area level. Equation 1 shows the applied linear estimation in GWR method.

$$y_i = a_0(u_i, v_i) + \sum_k a_k(u_i, v_i)x_{ki} + \delta_i \qquad (1)$$

Where (u_i, v_i) indicates the coordinates of the n-nth point in space, y_i is the dependent variable value, x_{ki} is descriptive independent variable in the i point, $a_k(u_i, v_i)$ is a parameter of the regression on each point of the independent variable, and δ_i is the error in the i point. On the other hand, the $a_k(u_i, v_i)$ coefficients are determined by a vector equation dependent on a $W_{i,j}$ weights function.

2.6. Downscaling LST

It is worth mentioning that once all the daytime and night-time LST climatic layers were obtained, their spatial resolution was augmented via the application of the previously described topoclimatic modelling method, using the monthly NDVI layers and the ASTER GDEM digital elevation model as descriptive variables. The topoclimatic model coefficients found after applying the GWR method correspondent to each pixel were interpolated using ordinary kriging with automatic variogram adjustment, thus obtaining monthly climatic layers with a 250 meter of spatial resolution.

2.7. Climatological maximum and minimum temperatures estimation

In the case of the minimum (T_N) and maximum (T_X) air temperature estimating, each month's climatological database information was used as dependent variable and the monthly daytime and night-time land surface temperature (LST) images of the pixels that span each respective station's area as descriptive variable. It should be noted that the MODIS sensor in its MOD11A2 product includes the survey of the study area 2 times a day in 3 different schedules, on board the TERRA satellite between 9:00 and 11:00, and on board the AQUA platform between 22:00 and 24:00, local time.

Considering the previous paragraph, Equation 2 allows the calculation of the maximum and minimum air temperature, using a multiple linear regression as a function of the predictive variable LST daytime and night-time, respectively.

$$T_a(x, y) = a_0(x, y) + a_1(x, y) \cdot LST(x, y) + \varepsilon(x, y) \qquad (2)$$

Where $a_0(x,y)$ and $a_1(x,y)$ correspond to the regression's coefficients and $\varepsilon(x,y)$ to the error or deviation, with (x,y) representing spatial variability. As a result of this process, whichever the data included in the models is, an equation that allows prediction of the local air temperature from the LST values of a pixel registered by the satellite over the weather station used for modelling will be obtained (Gutiérrez-Puebla et al., 2012). Geographic Weighted Regressions (GWRs) are local models that create an equation for each element of the dependent variable's data set (Soto-Estrada, 2013). In this case the equation's coefficients (a_0,a_1) were interpolated using ordinary cokriging with automatic variogram adjustment using altitude and distance to the Mediterranean coast as co-variables, thus obtaining monthly climatic layers adjusted to the station's collected values.

2.8. Statistical analysis

The analysis of the results was based on comparing each statistical model with the observed or measured value. Comparison of the results is at the mean monthly level and metrics

Table 1. Statistical criteria for evaluation performances of each model showed in Equation 1.

Description	Symbol	Formula	N°
Systematic error or Bias	BIAS	$\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)$	(3)
Mean Absolute Error	MAE	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\lvert O_i - P_i\rvert}$	(4)
Root Mean Square Error	RMSE	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2}$	(5)
Coefficient of Determination	r^2	$\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2}$	(6)
Index of Agreement	d	$1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(\lvert P_i - \bar{O}\rvert - \lvert O_i - \bar{O}\rvert)^2}$	(7)
Akaike Information Criterion	AIC	$2 \cdot k - N \cdot Ln\left(\frac{\sum_{i=1}^{N}(O_i - P_i)^2}{N}\right)$	(8)

such as Systematic Error or Bias (BIAS), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are used. Furthermore, a linear regression analysis was performed to test slope homogeneity between the methods displayed in Table 1 (Rawlings, 1988) and the observed value; and to calculate the Coefficient of Determination (R^2), which has been broadly used to evaluate the goodness of fit between the observed and estimated values. Additionally, other statistical indexes were calculated, such as the Index of Agreement (d) (Willmott et al., 1985; Legates and McCabe, 1999; Willmott and Matsuura, 2005; Meek et al., 2009; Willmott et al., 2012) and the Akaike Information Criterion (AIC), which is useful when comparing two or more statistical models that use the same dependent variable (Sakamoto et al., 1986; Burnham and Anderson, 2002;). A numerical expression used to obtain the AIC is the equation presented above (8) (Burnham and Anderson, 2002). Table 1 show the statistical criteria used to estimate each model's goodness of fit statistics, where N represents the number of observations, O the observed value, P is the predicted value, $\overline{O}$ the mean observed value, $\overline{P}$ the mean predicted value and k the number of parameters or independent variables used.

3. Results

The extreme air temperatures (minimum T_n and maximum T_x) estimation method was applied to the entire downloaded and processed MODIS image set for the estimation. In addition, a homogenous meteorological database was generated so that synergy between both types of data would be achieved for a better and more accurate estimation. In this section, the results are described showing a good representation of the temperature spatial patterns for the study area. Figures 2(a) and 2(b) show daytime and nighttime LST-MODIS' spatial variability for the spring, summer, winter, and autumn seasons in the MR.

3.1. Meteorological database

A total of 47 meteorological stations were collected in the MR, of which 40 had 20 years of data, which represents a total of 292200 records, of which 4281 correspond to missing data, equivalent to 1.4% of the total records. The other 7 remaining stations had between 3 and 15 years of data, with a total of 51135 records, of which 27727 correspond to missing data, which is equivalent to 54.2% of the total records. All the stations had periods of time in which no readings were recorded for hours or, in some cases, days, and prolonged periods of time.

During the data homogenization process, 32008 records were calculated, equivalent to 9.3% of the total records of all the meteorological stations considered for this study. Thus, missing records were completed for stations CA12, CA73, CI71, CR61, MO62, TP52 and TP91, up to the years 2005, 2016, 2015, 2006, 2015, 2009 and 2005 respectively. For the rest of the stations, the completed data corresponded to the period between the years 2000 and 2019.

3.2. Estimation of the air's minimum and maximum temperatures

Figures 2(a) and 2(b) respectively show the night-time and daytime LST-MODIS climatological maps corresponding to the 4 yearly seasons in the Murcia region, Spain. Figure 2(a) specifically shows the night-time LST, as the scene is captured at some point during the night and later converted to an approximately weekly product. On the other hand, Figure 2(b) shows the daytime LST, which corresponds to a scene captured at some point in the day and subsequently converted to an approximately weekly product. In both figures, the general patterns for the spatial distribution of night-time and daytime LST in the study area are clearly observed, fundamentally associated with topography and land use. A single-color scale is used to comparatively represent the LSTs for all the average seasons, since in this way it is possible to visualize the contrast between the different seasons.

The satellite information is used to obtain the spatially explicit linear regression between the air temperature (Minimum and Maximum), obtained from the network of meteorological stations, and the MODIS LST. The results of the topoclimatic modelling using GWR-Method for minimum and maximum temperature values related to the LST-MODIS values corresponding to night and day images, are shown in Tables 2 and 3, respectively. As shown in Tables 2 and 3, there is a strong linear correlation between LST from satellite images and Ta data obtained from weather stations in the study area, both for day and night LST. Specifically, Tables 2 and 3 show a summary of the results obtained for the minimum and maximum temperatures observed (Obs), modelled (Mod), and their difference (Diff=Mod-Obs), for mean values (ME) and the standard deviation (SD), respectively. In the same tables, the root mean square error (RMSE) and the Index of Agreement (d) are shown as indicators of the performance of the estimates made with GWR, for all the values extracted from the LST-MODIS images and the estimated values for each weather station and for the entire set of images used.

This numerical algorithm found values for the coefficients of determination (R^2) of the spatially explicit linear regression in the range of 0.82 to

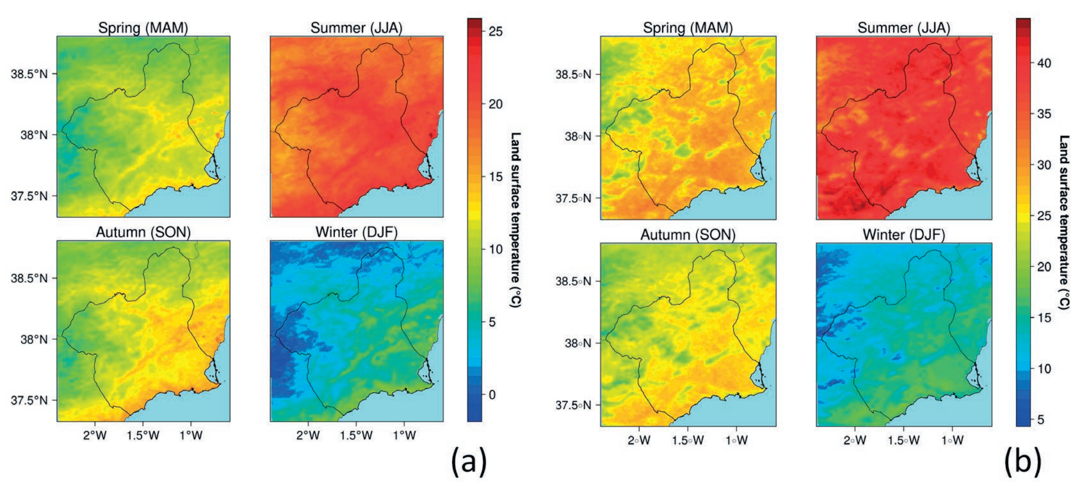

Figure 2. Maps for nocturnal (a) and diurnal (b) land surface temperature (LST) of climatological MODIS images for Spring, Summer, Autumn and Winter for the Murcia region, Spain.

Table 2. Summary statistics of GWR linear regression coefficients for estimation the spatial variation for the relationship between monthly mean LST and minimum air temperature (TN).

Month	a0	a1	RMSE	R^2	p	d	AIC
January	-0.5±0.5	1.2±0.1	1.0	0.82	**	0.94	138.26
February	-1.1±0.4	1.2±0.1	1.0	0.82	**	0.93	138.26
March	-0.7±1.2	1.0±0.1	1.0	0.82	**	0.92	137.09
April	-1.5±2.6	1.0±0.2	0.9	0.85	**	0.93	128.46
May	-1.7±4.6	1.0±0.3	0.9	0.85	**	0.93	121.99
June	4.5±7.3	0.6±0.4	0.8	0.85	**	0.93	118.36
July	1.7±9.9	0.8±0.4	0.8	0.84	**	0.93	116.05
August	1.2±10.1	0.9±0.5	0.8	0.86	**	0.94	117.50
September	0.9±8.3	0.9±0.5	0.8	0.89	**	0.96	112.18
October	1.2±5.5	1.0±0.4	0.9	0.88	**	0.95	122.83
November	-0.4±1.1	1.1±0.1	0.9	0.86	**	0.95	130.18
December	-0.1±0.7	1.1±0.1	1.0	0.84	**	0.95	133.61

RMSE units correspond to °C. The ** is a high statistical significance (p<0.01).

0.89 for the minimum temperatures, with an RMSE that varied between 0.8 and 1.0 °C. On the other hand, for maximum temperatures, the range for R^2 sits between 0.91 and 0.97, with an RMSE that ranges between 0.5 and 0.6 °C. In relation to the variability of the minimum and maximum air temperature, it is observed that these estimations represent adequately the seasonal patterns of temperature throughout the Murcia region, since the algorithm used represents local variabilities. The greatest difference in absolute values between the observed and modelled values at the monthly mean level is found at the El Chaparral (Cehegín) station with a value of 3 °C for the minimum temperature, while the smallest difference, corresponds about 0 °C, is found at the station Román (Jumilla) for the maximum temperature.

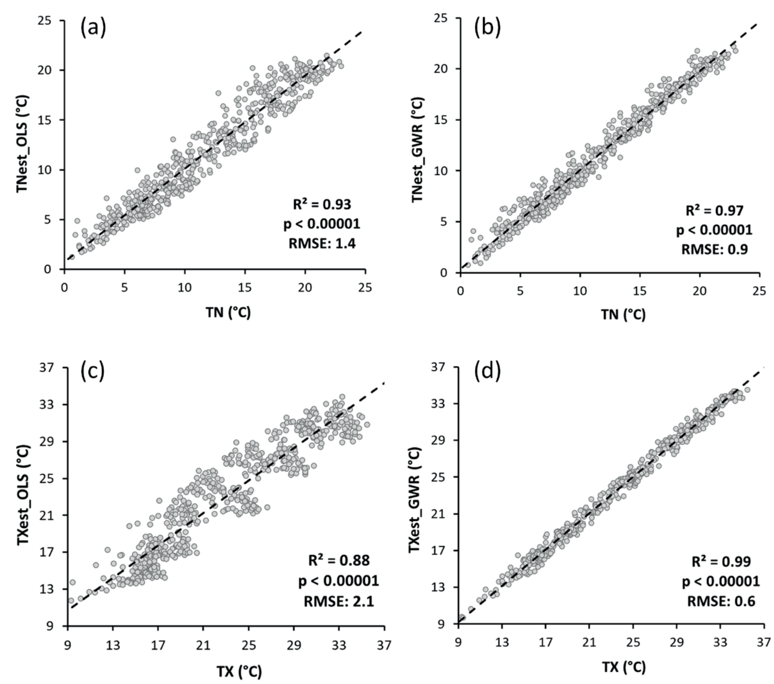

Figure 3. Graph: Estimated and observed temperatures for minimum air temperature OLS (a) and GRW (b), and maximum air temperature OLS (c) and GRW (d) in the Murcia region, Spain.

Table 3. Summary statistics of GWR linear regression coefficients for estimation the spatial variation for the relationship between monthly mean LST and maximum air temperature (TX).

Month	a0	a1	RMSE	R2	p	d	AIC
January	11.4±10.1	0.3±0.7	0.5	0.96	**	0.97	75.76
February	13.1±11.9	0.2±0.6	0.5	0.97	**	0.97	73.03
March	14.6±13.2	0.2±0.5	0.6	0.94	**	0.95	92.21
April	15.1±14.4	0.2±0.5	0.6	0.92	**	0.94	96.17
May	19.7±19.3	0.2±0.6	0.6	0.95	**	0.95	87.03
June	26.6±18.7	0.1±0.5	0.6	0.95	**	0.97	79.77
July	30.0±17.9	0.1±0.4	0.6	0.93	**	0.97	84.55
August	31.8±19.1	0.0±0.5	0.5	0.93	**	0.97	78.76
September	26.8±19.3	0.1±0.6	0.6	0.91	**	0.94	92.71
October	21.0±16.3	0.1±0.6	0.6	0.94	**	0.95	90.07
November	13.3±12.8	0.3±0.7	0.6	0.96	**	0.96	80.54
December	10.3±10.0	0.4±0.7	0.5	0.96	**	0.97	67.33

RMSE units correspond to °C. The ** is a high statistical significance ($p < 0.01$)

Figures 3(a) and 3(b) show the pairwise comparison between the observed and estimated values for the minimum temperature (T_n) through the (a) Ordinary Least Squares regression (OLS) and (b) GWR methods. The value of the coefficient of determination was 0.93 and 0.97 respectively, considered high, however the differences in the RMSE values, which were 1.4 °C and 0.9 °C respectively, indicate that the GWR method has a better performance than OLS for monthly mean values of minimum temperatures. In the case of maximum temperatures (T_x), the values of the coefficient of determination were 0.88 and 0.99, which correspond to the pairwise comparison between the observed and estimated values for the maximum temperature using the methods of (c) OLS and (d) GWR. The RMSE values in this case were 2.1 °C and 0.6 °C, respectively, which

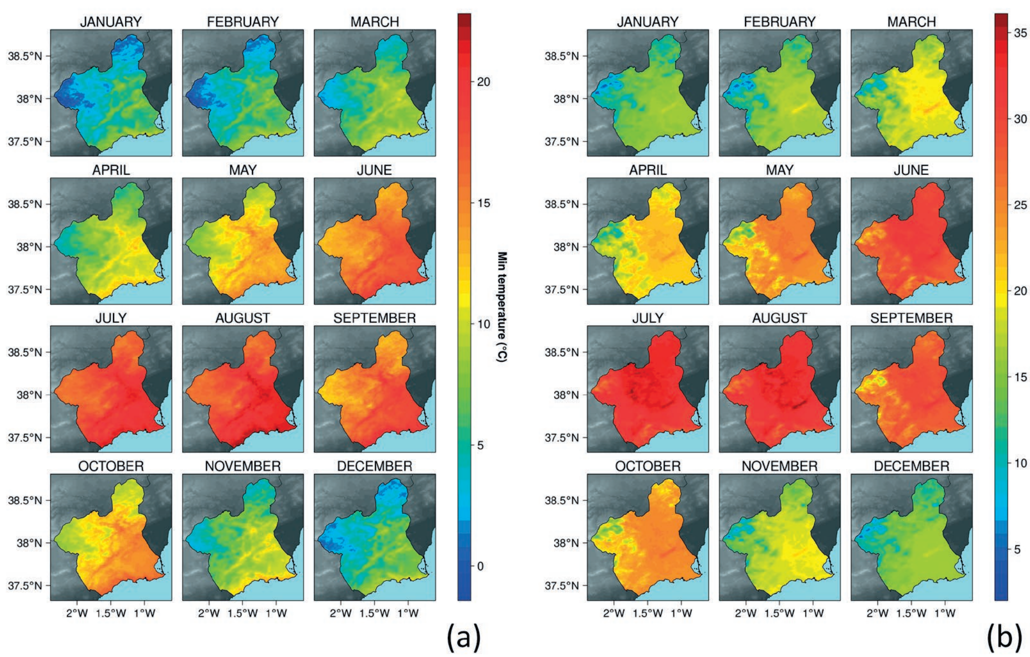

Figure 4. Minimum (T_N) and maximum (Tx) monthly mean air temperature values in the Murcia region, Spain.

indicates that in this case, too, the GWR method presents a better performance than OLS for mean monthly values of maximum temperatures. Figure 4 shows the spatial distribution of the monthly mean values of the minimum (T_N) and maximum (T_X) air temperature using the coefficients of the linear equations shown in Tables 2 and 3.

4. Discussion

The minimum (T_N) and maximum (T_X) monthly mean air temperatures estimated through the method used in the Murcia region, Spain, show a spatial stratification fundamentally due to the Mediterranean Sea and the physical geography's influence. This influence becomes more evident during the summer and spring months, when temperatures are higher in the coastal areas than in the mountain ones (Figure 4), which is also observed in the mean monthly LST-MODIS images (Figure 2). The spatial stratification of the minimum and maximum temperatures remains during the rest of the year (autumn-winter), however the differences between the coastal and mountain areas become more apparent.

In relation to the seasonal behaviour of the maximum temperatures in the 47 weather stations considered for the study, these display higher RMSE of 0.5 °C and 0.6 °C during winter and spring seasons, respectively, while the summer and autumn seasons showed RMSE values of 0.6 °C. In case of the minimum temperatures the behaviour is similar, with highest RMSE values in the summer and winter seasons at 0.8 °C and 1.0 °C, respectively, while the autumn and spring season values were correspondingly 0.8 °C and 0.9 °C.

For both T_N and T_X, it is observed that when calculating temperature estimates during the transitions between extreme seasons (summer and winter), the method used shows a greater margin of error. The coefficient of variation for T_N shows corresponding observed values of 41.04% and 31.96% during winter and autumn seasons, higher in comparison to the 28.51% and 12.56% values for spring and summer respectively. The situation is similar for TX, in which corresponding observed values for the autumn and spring seasons are 18.53% and 13.41%, higher than the values of winter and summer, at 10.08% and 6.73%, respectively. This error increase and spatial variability decrease in the estimations could be due to the rise in atmospheric activity during the autumns and spring transition months, which generates large air mass movement, causing more frequent episodes of higher and lower (though more spatially homogeneous) temperatures, and thus decreasing performance of the used simple method. Another source of errors could be the satellite observation timing, possibly due to the time for night and day observations. It is possible to correct this effect by adjusting a function that reproduces the behaviour of hourly temperatures, with the purpose of having values more adjusted to the satellite's time of passage (Bustos & Meza, 2015).

For maximum temperatures, the percentage relative differences between the observed and estimated values in the 47 weather stations display mean values lower than 7% (1.3 °C), showing a standard deviation of 0.3 °C and a spatial coefficient of variation of 77.8%. In the case of minimum temperatures, the percentage relative differences between the observed and estimated values in the 47 weather stations display mean values lower than 10.86% (2.5 °C) and a spatial coefficient of variation of 78.8%.

The results show, on average, lower error values compared to other studies in the field literature. The proposed method is based only on spatially explicit regression (GWR), a simple and easily reproducible numerical method with good results, as shown by some authors in different parts of the world (Yao & Zhang, 2013; Huang et al., 2017; Montaner-Fernández et al., 2020). The best adjustment and evaluation results of the linear models are obtained using MODIS nighttime and daytime together (mixed) to estimate minimum and maximum temperatures, respectively. These values show that the applied method managed to capture a significant percentage of the spatial behavior of temperatures, which is suitable for generating daily air temperature maps (Huang et al., 2017). According to the results described in the previous parahraph, it is possible to assert that the proposed interpolation method for estimating the spatial distribution of extreme temperatures in the Murcia region can be considered a valid and viable process to generate digital cartography for the T_N and T_X extreme temperatures climatological values (Jarvis & Stuart 2001; Bustos & Meza, 2015).

5. Conclusions

The medium resolution daily MODIS images (day and night) are valuable information that allow studies and investigations to be carried out on a regional and even global scale with different objectives.

In this study, as in previous ones, the linear relationship between air temperature and the LST-MODIS variable is verified, however, this relationship is improved by using the algorithm called Geographically Weighted Regression (GWR) since it allows for a better adjustment. The simple method implemented in the present study shows a good performance in the estimation of spatial and seasonal variability at the monthly mean level of air temperature, since it uses local adjustments instead of global adjustments. On the other hand, this good adjustment found is due to the integration of information from weather stations and the information provided by the MODIS satellite, which allowed a good estimation at the mesoscale territorial level of the maximum (T_X) and minimum (T_N) temperatures in the Murcia region, Spain.

The method and the spatial resolution developed in this study (250 m) can be adapted for finer spatial resolutions, such as by using LANDSAT or SENTINEL images to similar applications in estimating the spatial distribution of air temperature or variables related to it.

This linear relationship between the air temperature and the LST-MODIS variable can be used in other types of models, for example, the estimation of surface potential evapotranspiration (PET), which could improve our understanding of various spatial phenomena associated with vegetation at the mesoscale territorial level.

To finalize, this type of studies associated with the estimation of air temperature at the height of a weather station (approximately between 1 m and 2 m) through satellite images, can be an asset in relation to decision-making in territorial planning and definition of public policy, especially in studies of variability and climate change with applications to the agricultural sector.

Acknowledgments

This research was supported by the National Fund for Scientific and Technological Development (FONDECYT), Chile, project N° 1161809.

References

Akaike, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In: *2nd International Symposium on Information Theory*. Budapest, Hungary, 2-8 September. pp 267–281.

Aparecido, L.E.D.O., Moraes, J.R.D.S.C.D., Lima, R.F.D., Torsoni, G.B. 2022. Spatial Interpolation Techniques to Map Rainfall in Southeast Brazil. *Revista Brasileira de Meteorologia*, *37*(1), 141–155. https://doi.org/10.1590/0102-77863710015

Burnham, K.P., Anderson, D.R. 2002. *Model selection and inference: A practical information-theoretic approach*. New York: Springer-Verlag.

Bustos, E., Meza, F.J. 2015. A method to estimate maximum and minimum air temperature using MODIS surface temperature and vegetation data: application to the Maipo Basin, Chile. *Theoretical and Applied Climatology, 120*, 211–226. https://doi.org/10.1007/s00704-014-1167-2

Cortes, C., Vapnik, V. 1995. Support-Vector Networks. *Machine Learning*, *20*, 273-297. https://doi.org/10.1007/BF00994018

Cristóbal, J., Ninyerola, M., Pons, X. 2008. Modeling air temperature through a combination of remote sensing and GIS data. *Journal of Geophysical Research: Atmospheres*, *113*(D13). https://doi.org/10.1029/2007JD009318

Draper, N.R., Smith, H. 1998. *Applied regression analysis*. 326. New York: John Wiley & Sons. https://doi.org/10.1002/9781118625590

Estévez, J., Gavilán, P., Giráldez J.V. 2011. Guidelines on validation procedures for meteorological data from automatic weather stations. *Journal of Hydrology*. *402*(1-2), 144-154. https://doi.org/10.1016/j.jhydrol.2011.02.031

Fotheringham, S., Brundson, C.H., Charlton, M. 2002. *Geographically Weighted Regression: The analysis of spatially varying relationships*. Chichester: John Wiley & Sons.

Gutiérrez-Puebla, J., Daniel-Cardozo, O., García-Palomares, J.C. 2012. Regresión Geográficamente Ponderada (GWR) y estimación de la demanda de las estaciones del Metro de Madrid. In: *XV Congreso Nacional de Tecnologías de la Información Geográfica*. Madrid, Spain, 19-21 September. pp 327-338.

Herrera, S., Fernández, J., Gutiérrez, J.M. 2016. Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: assessing the effect of the interpolation methodology. *International Journal of Climatology*, *36*(2), 900–908. https://doi.org/10.1002/joc.4391

Huang, F., Ma, W., Wang, B., Hu, Z., Ma, Y., Sun, G., Lin Y. 2017. Air temperature estimation with MODIS data over the Northern Tibetan Plateau. *Advances in Atmospheric Sciences, 34*(5), 650–662. https://doi.org/10.1007/s00376-016-6152-5

Jarvis, C.H., Stuart, N. 2001. A Comparison among Strategies for Interpolating Maximum and Minimum Daily Air Temperatures. Part II: The Interaction between Number of Guiding Variables and the Type of Interpolation Method. *Journal of Applied Meteorology*, *40*(6), 1075-1084. https://doi.org/10.1175/1520-0450(2001)040<1075:ACASFI>2.0.CO;2

Karatzoglou, A., Meyer, D., Hornik, K. 2006. Support vector machines in R. *Journal of Statistical Software*, *15*(9), 1–28. https://doi.org/10.18637/jss.v015.i09

Legates, D.R., McCabe, G.J. 1999. Evaluating the use of goodness of fit measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, *35*(1), 233-241. https://doi.org/10.1029/1998WR900018

Littmann, T. 2008. Topoclimate and microclimate. In *Arid Dune Ecosystems*. Berlin: Springer. https://doi.org/10.1007/978-3-540-75498-5_12

Lorençone, P.A., Aparecido, L.E.D.O., Lorençone, J.A., Torsoni, G.B., Lima, R.F.D. 2022. Estimation of Air Temperature Using Climate Factors in Brazilian Sugarcane Regions. *Revista Brasileira de Meteorologia*, *37*(1), 121–140. https://doi.org/10.1590/0102-77863710008

Marzban, F., Conrad, T., Marzban, P., Sodoudi, S. 2018. Estimation of the Near-Surface Air Temperature during the Day and Nighttime from MODIS in Berlin, Germany. *International Journal of Advanced Remote Sensing and GIS*, *7*(1), 2478–2517. https://doi.org/10.23953/cloud.ijarsg.337

Meek, D.W., Howell, T.A., Phene, C.J. 2009. Concordance Correlation for Model Performance Assessment: An Example with Reference Evapotranspiration Observations. *Agronomy Journal*, *101*(4), 1012–1018. https://doi.org/10.2134/agronj2008.0180x

Montaner-Fernández, D., Morales-Salinas, L., Sobrino-Rodríguez, J., Cárdenas-Jirón, L., Huete, A., Fuentes-Jaque, G., Pérez-Martínez, W., Cabezas, J. 2020. Spatio-Temporal Variation of the Urban Heat Island in Santiago, Chile during Summers 2005–2017. *Remote Sensing, 12*(20), 3345. https://doi.org/10.3390/rs12203345

Otgonbayar, M., Atzberger, C., Mattiuzzi, M., Erdenedalai, A. 2019. Estimation of Climatologies of Average Monthly Air Temperature over Mongolia Using MODIS Land Surface Temperature (LST) Time Series and Machine Learning Techniques. *Remote Sensing*, *11*(21), 2588. https://doi.org/10.3390/rs11212588

Prihodko, L., Goward, S. 1997. Estimation of air temperature from remotely sensed surface observations. *Remote Sensing of Environment*, *60*(3), 335–346. https://doi.org/10.1016/S0034-4257(96)00216-7

Rawlings, J.O. 1988. *Applied regression analysis. A research tool*. Belmont: Wadsworth and Brooks/Cole.

Recondo, C., Zapico, E., Peón, J.J., Pendás, E., Aguirre, R., Abajo, A. 2011. Estimación de la temperatura del aire a partir de la temperatura de superficie obtenida con el sensor MODIS e información espacio-temporal. Una aplicación en modelos de riesgo de incendios forestales en la Península Ibérica. In: *XIV Congreso de la Asociación Española de Teledetección (AET)*. Mieres del Camino, España, 21-23 September. pp 457-460.

Recondo, C., Peón, J.J., Zapico, E., Pendás, E. 2013. Empirical models for estimating daily surface water vapour pressure, air temperature, and humidity using MODIS and spatiotemporal variables. Applications to peninsular Spain. *International Journal of Remote Sensing*, *34*(22), 8051-8080. https://doi.org/10.1080/01431161.2013.828185

Ruiz-Álvarez, M., Alonso-Sarria, F., Gomariz-Castillo, F. 2019. Interpolation of Instantaneous Air Temperature Using Geographical and MODIS Derived Variables with Machine Learning Techniques. *ISPRS International Journal of Geo-Information*, *8*(9), 382. https://doi.org/10.3390/ijgi8090382

Sakamoto, Y., Ishiguro, M., Kitagawa, G. 1986. *Akaike Information Criterion Statistics*. Dordrecht, The Netherlands: D. Reidel.

Sobrino, J.A., El Kharraz, J., Li, Z.L. 2003. Surface temperature and water vapor retrieval from MODIS data. *International Journal of Remote Sensing*, *24*(24), 5161–5182. https://doi.org/10.1080/0143116031000102502

Soto-Estrada, E. 2013. Regresión ponderada geográficamente para el estudio de la temperatura superficial en Medellín, Colombia. *Revista AIDIS de Ingeniería y Ciencias Ambientales*, *6*(3), 42–53.

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(14),1559–1579. https://doi.org/10.1002/(SICI)1097-0088(19971130)17:14<1559::AID-JOC211>3.0.CO;2-5

Wan, Z. 2008. New refinements and validation of the MODIS Land-Surface Temperature/ Emissivity products. *Remote Sensing of Environment*, *112*(1), 59–74. https://doi.org/10.1016/j.rse.2006.06.026

Willmott, C.J., Robeson, S.M. 1995. Climatologically aided interpolation (CAI) of terrestrial air temperature. *International Journal of Climatology*, *15*(2), 221–229. https://doi.org/10.1002/joc.3370150207

Willmott, C.J., Robeson, S.M. Matsuura, K. 2012. A refined index of model performance. *International Journal of Climatology*, *32*(13), 2088–2094. https://doi.org/10.1002/joc.2419

Willmott, C., 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*(1), 79–82. https://doi.org/10.3354/cr030079

Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., Rowe, C.M. 1985. Statistics for the evaluation of model performance. *Journal of Geophysical Research, 90*(C5), 8995–9005. https://doi.org/10.1029/JC090iC05p08995

Yao, Y., Zhang, B. 2013. MODIS-based estimation of air temperature of the Tibetan Plateau. *Journal of Geographical Sciences*, *23*(4), 627–640. https://doi.org/10.1007/s11442-013-1033-7

Zhao, C.Y., Zhang, H.X., Zhang, X.Y., Liu, M.C., Hu, Z.D., Fan, B.T. 2006. Application of support vector machine (SVM) for prediction toxic activity of different data sets. *Toxicology*, *217*(2-3), 105–119. https://doi.org/10.1016/j.tox.2005.08.019