Agricultural and Forest Meteorology 149 (2009) 1343–1357

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain

Husain Alsamamra, Jose Antonio Ruiz-Arias, David Pozo-Vázquez *, Joaquin Tovar-Pescador

Department of Physics, University of Jaén, 23071, Jaén, Andalusia, Spain

ARTICLE INFO

Article history:
Received 19 May 2008
Received in revised form 10 March 2009
Accepted 15 March 2009

Keywords:
GIS
Kriging
Global solar radiation
Andalusia
Southern Spain
Complex topography

ABSTRACT

This study presents a comparative analysis of the ordinary and residual kriging methods for mapping, on a 1 km × 1 km grid size, the monthly mean of global solar radiation at the surface in Andalusia (southern Spain). The region of study is characterized by a wide range of topographic and climatic characteristic, which allows properly evaluating the two methods. The experimental dataset includes 4 years (2003–2006) of data collected at 166 stations: 112 stations were used to train the models and 54 in an independent validation procedure. Overall, the ordinary kriging method provide fair estimates: RMSE ranges from 1.63 MJ m^{-2} day^{-1} (6.2%) in June to around 1.44 MJ m^{-2} day^{-1} (11.2%) in October. In the residual kriging procedure, we propose using an external explanatory variable (derived just based on a digital elevation model) that accounts for topographic shadows cast, and that is able to explain between 13% and 45% of the spatial variability. Based on the combined used of the elevation and the former external variable, residual kriging estimates shows a relative improvement in RMSE values ranging from 5% in the summer months to more than 20% in the autumn and winter months. Particularly, RMSE is 1.44 MJ m^{-2} day^{-1} (5.5%) in June and 1.31 MJ m^{-2} day^{-1} (10.2%) in October. Explained variance also shows a considerable improvement compared to the ordinary kriging method, with all the months showing R^2 values above 0.92. Results show that most part of these improvements is associated with a better estimation of the minimum values, particularly during the winter part of the year. It is finally concluded that the proposed residual kriging method is particularly valuable when mapping complex topography areas.

1. Introduction

Global solar radiation is an important climate variable with respect to many fields as agriculture or renewable energy applications. In these fields, spatially continuous data set (maps) are needed. Along the last decades three different methodologies have been used for mapping the solar radiation: the use of satellite estimates, the use of Geographic Information System (GIS)-based solar radiation models and, traditionally, the use of the well known geostatistical methods.

Regarding the satellite estimates, geostationary satellites may provide spatially continuous irradiance values. Processing of satellite data provides less accurate values (compared to ground measurements), but has the advantage of the large spatial coverage that can provide. For instance, the Meteosat Second Generation (the second generation of the European meteorological satellite) has a temporal resolution of 15 min and 2.5 km of spatial resolution. In spite of the strong improvement compared to the previous Meteosat generation, many applications still need a better spatio-temporal resolution and the value of the solar radiation estimates for complex topography areas is limited. Nevertheless, several works have shown the usefulness of the satellite estimates as additional information in interpolation techniques for solar radiation estimation in mountainous regions (Zelenka et al., 1992; Beyer et al., 1997).

The use of GIS-based solar radiation model for solar radiation mapping has been developed along the last decades. These models uses the topographic information contained in a digital elevation model (DEM) to determine topographic features such as elevation, surface orientation and shadow casting. Based on this information and different physical parameterization, these models are able to estimate the incoming solar radiation at every point of the DEM (Tovar-Pescador et al., 2006). Several models such as SolarFlux (Hetrick et al., 1993; Dubayah et al., 1995), Solei-32 (Miklánek and Mèszáos, 1993; Mèszáos and Miklánek, 2002), Solar Analyst (Fu and Rich, 2000), SRAD (Wilson and Gallant, 2000), and r.sun (Hofierka and Súri, 2002) have been developed in the last decade.

* Corresponding author at: Departamento de Física, Universidad de Jaén, Edificio A3, Campus Lagunillas, 23071, Jaén, Spain. Tel.: +34 953 212783; fax: +34 953 212838.
E-mail addresses: husain@ujaen.es (H. Alsamamra), jararias@ujaen.es (J.A. Ruiz-Arias), dpozo@ujaen.es (D. Pozo-Vázquez), jtovar@ujaen.es (J. Tovar-Pescador).

0168-1923/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.agrformet.2009.03.005

Ruiz-Arias et al. (in press) provides an evaluation of some of these models. Recently, Pons and Ninyerola (2008) have proposed a new GIS-based solar radiation model, tuned with local measured meteorological data. The main problem of these models is that they need for external meteorological information (often difficult to obtain) to provide the solar radiation estimates. Particularly, depending on the model, meteorological parameters as the Linke turbidity, cloudiness, atmospheric transmittance, circumsolar coefficient or albedo, are needed.

The interpolation techniques allow obtaining spatially continuous databases from isolated-stations measurements based on spatially interpolation methods. There are two different groups of the interpolation procedures. The first one uses deterministic interpolation techniques, which use mathematical functions to calculate the unknown values based on the degree of similarity with respect to the known points and provides no reliability assessment errors of the predicted surface. Examples of these techniques are the spline-functions or weighted averages (Hulme et al., 1995; Zelenka et al., 1992). The second one is the stochastic methods, which uses both analytical and statistical methods to predict unknown values based on the spatial auto-correlation among data points and offers the reliability of the predicted surface (Burrough and McDonnell, 1998). Among these methods, one of the most widely used are the kriging methods (Webster and Oliver, 2001). The kriging methods have showed considerable advantages, compared to the deterministic interpolations procedures, in the estimation of the rainfall (Tabios and Salas, 1985; Buytaert et al., 2006) and the temperature (Jarvis and Stuart, 2001; Zhao Chuanyan et al., 2005). Regarding the solar radiation, Rehman and Ghori (2000) demonstrated the usefulness of the use of the kriging technique to estimate the global radiation in Saudi Arabia. In this study, the root mean square error (RMSE) of the estimates ranged from 0.5% to 1.7%. More recently, Ertekin and Evrendilek (2007) used the universal kriging for mapping daily global solar radiation in Turkey. Results showed that this procedure were reliable in predicting the spatial variability of the global solar radiation, with RMSE values ranging from 1.4 to 10.0 MJ m^{-2} day^{-1}.

The reliability of interpolation techniques is strongly dependent on the sample size (Hughes and Lettenmaier, 1981; Mubiru et al., 2006). Particularly, kriging may provide reliable estimates of climate variables, as the solar radiation, in homogeneous terrain with similar climate characteristics. Nevertheless, the reliability of the estimates decreases when the complexity of the topography increases or when the earth surface is heterogeneous (as along land-sea discontinuities). In such cases, stochastic interpolation processes may not provide meaningful spatially continuous estimates, since point-specific measurements can be affected by strong local variation. For the solar radiation, particularly, complex topography areas present a challenge. Discontinuity in elevation and surface orientation (slope and aspect), and shadows cast by topographic features can create strong local gradients in the solar radiation that interpolation processes may not properly account for.

Many techniques have been proposed to overcome this weakness. These techniques allow taking into account external variables that may provide complementary information for the interpolation and, therefore, compensate for the lack of data and the scarce sample size (Odeh et al., 1995). These external variables may be used locally or in the whole study area and, in most of the cases, are related to geographical or topographical characteristics. These characteristics are currently available across the world at 1 km × 1 km of resolution (for instance the USGS global data). There are numerous studies trying to identify external variables for temperature and precipitation (Chuvieco and Salas, 1996; Benzi et al., 1997; Hargy, 1997; Llasat, 1997; Menz, 1997; Jorge et al., 2003; Carrera-Hernández and Gaskin, 2007). A few also dealt with the solar radiation (Sen and Sahin, 2001; Uran and Yun, 2004). There are different ways in which the external variables can be taken into account in the kriging process. For instance, the information coming from the external variables can be considered during the interpolation process, using co-kriging methods (see Carrera-Hernández and Gaskin, 2007). This method is advantageous when the external variable is highly correlated to the studied variable (Gotway and Hartford, 1996), but becomes very complex when more than one covariables are considered (Ahmed and de Marsily, 1987). Instead of including the external information directly in the kriging process, it is possible to consider it during the first step, prior to the interpolation itself. There are different denominations for this technique, as '*kriging with a guess field*' (Ahmed and de Marsily, 1987) or '*residual kriging*' (Phillips et al., 1992; Martinez-Cob, 1996). We will use this last denomination hereinafter. Basically, in the first step, a multiple linear regression is fitted between the variable of interest and some external explanatory variables. Then, an ordinary kriging procedure is applied to the residuals of this multiple regression analysis. Finally, a map is obtained integrating both the multiple regression and the kriging results. This technique, although relatively simple, is powerful, since allows including in an easy way multiple sources of external information in the interpolation procedure that may compensates for the small sample size. The residual kriging procedure has been used in numerous works for precipitation and temperature estimation (Prudhomme and Reed, 1999; Ninyerola et al., 2000; Ustrnul and Czekierda, 2005).

In this work, we present an application of the residual kriging methodology for mapping monthly-averaged global radiation at the surface in Andalusia (southern Spain). The aim is to evaluate the potential usefulness of this methodology for mapping the solar energy resources in this region, characterized by a wide range of topographic and climatic characteristic. Particularly, a key part of this work is concerned with the identification topography-related external variables able to account for part of the observed solar radiation spatial variability and, then, useful to improve the reliability of the kriging estimates. The ordinary kriging method is also applied, to evaluate the improvement provided by the residual kriging method.

The paper is organized as follows. Section 2 describes the study region, the dataset and, finally, presents the methodology. Section 3 presents the results and, finally, some conclusions are highlighted in Section 4.

2. Methodology

In this section, the study area and the database used in this work are first introduced. Then, a brief description of the geostatistical methods used in this work is presented. A more detailed description of, particularly, kriging procedures can be found in (Burgess and Webster, 1980; Vic Barnett, 2004). Finally, at the end of the section, the models evaluation procedure is explained.

2.1. Study area

The region of the study is the area of Andalusia, in the southern part of Iberian Peninsula (Fig. 1), covering around 87,000 km^2. The region is located in a transition zone (latitudes 35°30′–38°30′N and longitudes 7°30′–1°30′W) from temperate to subtropical climates, with the Atlantic Ocean and the Mediterranean region in the southern bound. Two different parts, from the topographic point of view, can be considered in the region. The western part, covering around 30,000 km^2, is an almost homogeneous flat area, with about 100 m of mean elevation. On the other hand, the eastern part

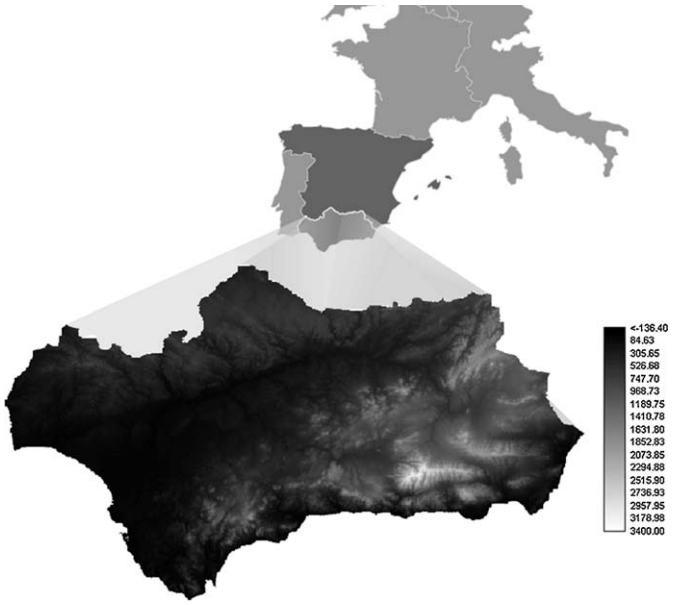

Fig. 1. Map showing the location of the study area in southern Spain. The color scale refers to the elevation in meters above sea level.

presents a very complex topography, with several mountain ridges, reaching up to 3470 m of elevation (the highest elevation of the Iberian Peninsula) in the Sierra Nevada National Park. The combination of the former factor give rise to the existence of several climates within the region, from pure Mediterranean climate conditions near the coasts, to continental climate in the interior and even mountain climate in some regions (Font-Tullot, 1988). The study area, therefore, presents a challenge for the kriging methodology, due to both the wide range of climatic characteristics and topographic variability, allowing to properly evaluate the ordinary and residual kriging methodologies. Particularly, the complexity of the study region provides the results with a greater representativeness than those derived from just a regional study with homogeneous physiographic conditions.

2.2. Database

The experimental dataset used in the study includes 4 years (2003–2006) of daily global solar radiation measures at the surface over 166 meteorological stations. The stations are owned and maintained by the Andalusian Regional Government, and are part of the Agricultural and Environmental monitoring network. Different kinds of instrument are included in this network, from Kipp and Zonen pyranometer to less accurate solid-state radiometers. Overall, an instrumental error of about 5% can be assumed. The location of the stations, displayed in Fig. 2, covers the region almost homogeneously. The 166 stations were divided into two groups: two-third of the stations (112) were used to train the models (training dataset) and one-third (54) was used for the validation process (validation dataset).

The mean density of meteorological stations is one station for every 482 km^2 (22 km × 22 km), the latitude and longitude differences between stations are no greater than 2.8° and 5.5°, respectively, and the distances between stations fluctuate between 5 and 450 km. The former topographic characteristic reduces the risk of mesoscale discontinuity or anisotropy (Merino et al., 2001). Finally, the elevation of the stations are in the range 4 to 1212 m, with about 7% of the study area are covered by mountains with elevation above 1212 m. Missing values represented less than 2% of the dataset, and were replaced using nearby stations. Several quality control procedures were performed at this stage. Particularly, the atmospheric transitivity, $\tau = MR/PR$, with MR the measured radiation and PR the potential extraterrestrial radiation, were computed each day. A maximum value of 0.8 (Luca et al., 2000) was assumed. Values above this maximum or below 0.01 MJ m^{-2} day^{-1} were considered wrong and therefore excluded (Iqbal, 1983). Excluded values (less than 1%) were replaced averaging the nearest four available values. The daily values were then averaged to obtain monthly-averaged global solar radiation at the surface. All the analyses were carried out independently for the 12 monthly values. Finally, 12 monthly averages were, again, obtained based on the 4-year-dataset. The kriging study was carried out for each of these 12 monthly spatial set of data.

2.3. Ordinary kriging

Kriging refers to a family of least-square linear regression algorithms that attempt to predict values of a variable at locations where data are not available based on the spatial pattern of the available data. The description of kriging theory and its application are given in detail by Delhomme (1978). Ordinary kriging is the only technique that considers two sources of information regarding the attribute, the variation and the distance between

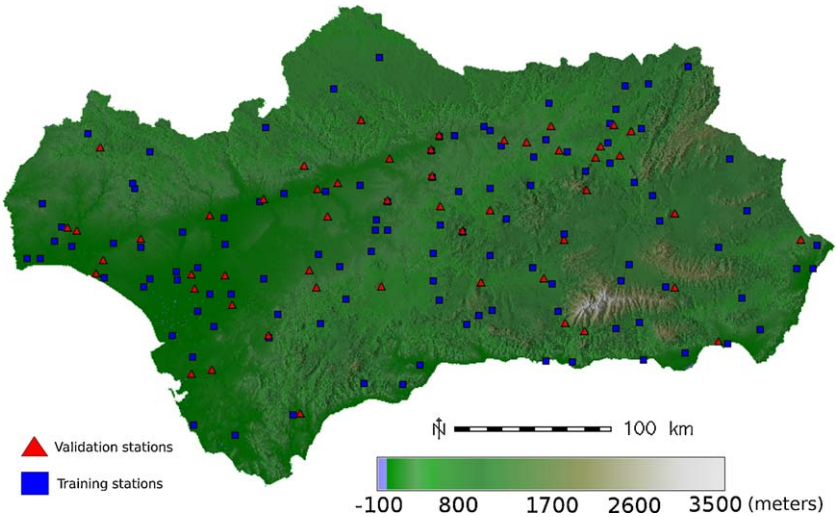

Fig. 2. Location of the 166 radiometric stations used in the study. Two-third of the stations (112) were used for the model training, while one-third (54) were used for validation purposes. The color scale refers to the elevation in meters above sea level.

points (Webster and Oliver, 2001). The basic assumption is that the data are realisations of a random function $\{Z(x):x \in D\}$, where "x" is a spatial index and "D" is a fixed domain in two dimensional space. An assumption of stationarity is made, expressed in two parts. First, the mean of the process is considered constant and should be a function of lag-distance, h. This means the expected difference in the values for two points x and $x + h$ (where h is the distance between the two points) is zero, i.e.:

$$E[Z(x) - Z(x+h)] = 0, \tag{1}$$

where E denotes expectation and $x \in D$, $x + h \in D$. Second, the variance of the difference between two values is assumed to depend only on the distance h between the two points, which means they are spatially auto-correlated and that close points are more likely to have similar values or similar differences than distant points, i.e.:

$$\mathrm{var}[Z(x+h) - Z(x)] = 2\gamma(h), \tag{2}$$

where "var" denotes variance. The function $\gamma(h)$ is called the semivariogram, often called the variogram (Burgess and Webster, 1980). The objective of ordinary kriging procedure is to estimate data values at unsampled locations x_0 using information available elsewhere in the domain ($x_1, x_2, \ldots, x_n$). This can be carried out by expressing $Z(x_0)$ as a linear combination of the data ($Z(x_1), Z(x_2), \ldots, Z(x_n)$):

$$\hat{Z}(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_0). \tag{3}$$

The optimal weights "λ_i" are calculated assuming that the estimation $Z(x_0)$ by $\hat{Z}(x_0)$ is unbiased, that is, the expected value of the estimates is the same as that of the known data. The condition needed for unbiased estimator is $\sum_{i=1}^{n} \lambda_i = 1$. The sum of the squares of the error estimation is minimised $E[\hat{Z}(x_0) - Z(x_0)]^2$.

The semivariogram is a statistical model which represents how the data vary spatially across the area of interest. The variation between points is measured using the semivariance. Pooling together pairs of data at a geographic distance h, the semivariance $\gamma(h)$ of the sample can be written as:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2, \tag{4}$$

where $N(h)$ is the number of pairs of points separated by distance h. Once the semivariogram function has been computed from the sampled values at different locations, the next step is to fit a parametric semivariogram $\gamma(h)$. A common method used to fit parametric semivariogram models to the sample semivariogram is the weighted least squares method, as proposed by Cressie (1985). This approach was taken in this work using IDRISI software (Clark Labs, 2006). In this study, an exponential model has been used to fit the sample semivariograms. Particularly, this model parameterizes the semivariogram in the following way:

$$\gamma(h) = C_0 + C_1 \left[1 - \exp\left(\frac{-h}{a}\right)\right], \tag{5}$$

where C_0, C_1, and a are called, respectively, nugget, sill, and range.

2.4. Residual kriging

The kriging procedure may provide reliable solar radiation estimates in homogeneous terrain with similar climate characteristics. Nevertheless, complex topography areas present a challenge. Variability in elevation, surface orientation (slope and aspect), and shadows cast by topographic features can create strong local gradients in the solar radiation. In such cases, kriging estimates reliability decreases, since point-specific measurements can be affected by strong local variation.

To overcome this weakness, different procedures have been proposed to include information coming from external explanatory variables during the kriging procedure. These external variables are selected in order to provide complementary information for the interpolation and, therefore, compensate for the relatively low sample size. The residual kriging procedure (Ahmed and de Marsily, 1987; Phillips et al., 1992; Martinez-Cob, 1996; Prudhomme and Reed, 1999) is, probably, the easiest of these techniques.

In the residual kriging procedure, instead of kriging $Z(x)$ directly, a regression analysis is first carried out between $Z(x)$ and some external variables $a_i(x)$, giving:

$$Z(x) = Z^*(x) + \varepsilon(x) = \sum \alpha_i a_i(x) + \varepsilon(x), \tag{6}$$

where $x \in D_{\mathrm{Obs}}$, D_{Obs} being the domain where the observations $a_i(x)$ are available. The coefficients α_i are fitted using the ordinary least squares (OLS) procedure. The spatial field $Z^*(x)$ can be estimated for each point x where $a_i(x)$ is known from Eq. (6). Where the true value $Z(x)$ is also known, a residual error, $r(x)$, can be defined as:

$$r(x) = Z(x) - Z^*(x), \tag{7}$$

where $x \in D_{\mathrm{Obs}}$.

The new variable $r(x)$ retains the spatial variability of $Z(x)$ (Odeh et al., 1995), but some of the variability has been removed, as a result of external information used in the regression model.

After this analysis, a kriging procedure is used for the residuals $r(x)$. As a result, a map of $\hat{r}(x)$, representing the corrections to apply to the regression model, is obtained. The final estimates $\hat{Z}(x)$ are obtained by combining both $Z^*(x)$ and $\hat{r}(x)$ estimates separately on the kriging grid:

$$\hat{Z}(x) = Z^*(x) + \hat{r}(x). \tag{8}$$

This technique, although relatively simple, is powerful, since it allows including, in an easy way, multiple sources of external information in the interpolation procedure that compensates the lack of local data.

In this study, the external variables used for the global radiation kriging procedure are related to topographic features. The use of GIS allows deriving a wide diversity of variables, readily available for the kriging grid. In Section 3.1, a detailed description of the external variables used in this work is presented.

2.5. Results evaluation procedure

Usually, cross-validation techniques are used to check the performance of the kriging procedure (Cooper and Istok, 1988). Particularly, sample values $Z(x_0)$ were deleted from the dataset one at a time and then the kriging process were carried out with the remaining sample values, in order to estimate the value of $Z(x_0)$ at the location of the deleted sample. In this work, we have rather preferred to use an independent data set for evaluating the models results. Particularly the set of 166 available measuring stations (Section 2.2) was divided into two groups: two-third of the stations (112, training dataset) were used to develop the models and one-third (54, evaluation dataset) was used for an independent validation process. This validation procedure, using an independent dataset is more stringent than the common cross-validation procedure.

Following Isaaks and Srivastava (1989), three different scores were used: the mean error (ME), the mean absolute error (MAE) and the RMSE, computed as follows:

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(\hat{Z}(x_i) - Z(x_i)), \tag{9}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{Z}(x_i) - Z(x_i)|, \tag{10}$$

$$RMSE = \sum_{i=1}^{n}\sqrt{\frac{1}{n}(\hat{Z}(x_i) - Z(x_i))^2}, \tag{11}$$

where "n" corresponds to the number of stations used in the study. Note, therefore, that these error measures are related to the spatial variability.

Also the percentage of error in terms of ME, MAE, and RMSE were obtained. These three scores provide different information about the method reliability. The ME is an indicator of mean bias, but it should be used cautiously as an indicator of accuracy, because negative and positive values counteract each other. MAE indicates the extent to which the process leads to error, ignoring sign. The RMSE is mainly a joint measure of bias in the mean and in the variance (spatial variance in our case), as obviously the square of individual differences between estimated and observed values put the emphasis on the errors in outliers or higher differences (Ashraf et al., 1997; Nalder and Wein, 1998). The combination of MAE and RMSE errors give the possibility to analyze if the errors in the estimates are only due to bias in the mean values (i.e. MAE $\approx$ RMSE) or if there is also a contribution of the errors in the spatial variability or in stronger values (i.e. RMSE $>$ MAE).

The combined evaluation of these three scores provides the best method of evaluating the models giving the kriged variable. Note that, for the residual kriging model, the kriged variable $r(x)$ is not the variable of interest $Z(x)$. Therefore, the evaluation is carried out for the final estimates of $\hat{Z}(x)$ obtained combining both the ordinary kriging and the residual kriging.

3. Results and discussion

The ordinary kriging procedure first involves (Fig. 3) computing the semivariogram and, then, the evaluation of solar radiation values based on this semivariogram. On the other hand, the residual kriging methodology consists of two stages, which involves both deterministic and stochastic estimations. In the first stage, the observed solar radiation values are parameterized using explanatory variables. This is the deterministic component of the model. In the second step the residuals of the regression analysis, treated as random spatial variables and which represent the stochastic component of the model, are interpolated using the ordinary kriging. The final solar radiation estimates are obtained by combining the results of the first and second step. This section describes the results of all these analysis. Particularly, in the first part, we introduce the explanatory variables used in the residual kriging procedure. In a second part, the ordinary and residual kriging procedures results are presented.

3.1. External explanatory variables for the residual kriging

As previously highlighted, a key part of residual kriging method consists on identifying external explanatory variables useful in providing additional information in the interpolation procedure. Many factors and processes interact to determine the amount of solar radiation received at a given point on the Earth's surface. First, the period of the year determines the incidence angle of the sunrays. Second, the atmospheric conditions change the solar fluxes through absorption and scattering processes. Finally, the variability in elevation, surface orientation (slope and aspect), and shadows cast by topographic features can create strong local gradients in the solar radiation. In this work, we have only considered explanatory variables that can be derived from DEM. This makes the results more general and useful, since these variables are readily available in a free-access. Particularly, in this study we have evaluated the effect of the elevation and shadows cast as external variables in the kriging procedure. To deal with the seasonality of the solar radiation variability, the study is carried out for each month independently. The topographic characteristics were derived based on a 1 km spatial resolution DEM. The IDRISI

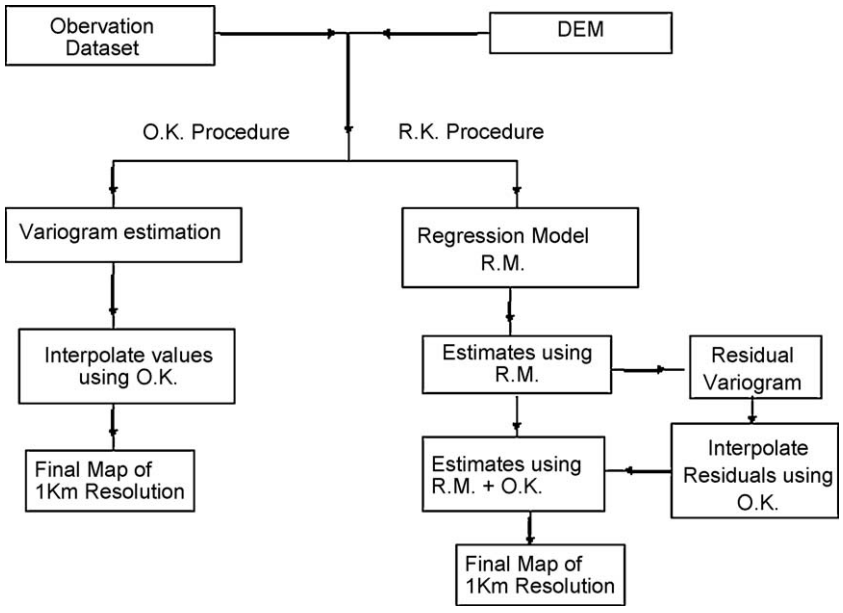

Fig. 3. Outline of the two different methodologies used in this work: ordinary kriging (OK) and residual kriging (RK). Note that the residual kriging method is a two-stage procedure.

GIS software (Clark Labs, 2006) was used to carry out the regression analysis and the whole residual kriging procedure.

Regarding the elevation, it is clear that this variable is related to the atmospheric attenuation of the incoming solar radiation at the earth surface. The higher the elevation, the lower the atmospheric layer thickness and, therefore, the lower the atmospheric attenuation. The importance of this explanatory variable is strongly dependent on the elevation gradient. In the present study, this gradient is relative small (minimum elevation is 4 m and maximum 1212 m).

The other external explanatory variable that has been explored in this study is related to the shadow cast caused by the topography. The slope and aspect effect on the measured values cannot be evaluated in this study since the stations are located on horizontal surfaces. The shadow cast influence, nevertheless, can be evaluated. Particularly, to assess this effect, a sky-view factor can be used. This sky-view factor can be defined as the ratio between radiation actually received by a planar surface and that received from the entire hemispheric radiating environment (without any obstruction) (Watson and Johnson, 1987). It represents the shadows cast by topographic features. Sky-view factor can be derived from DEM using several available procedures. In the present study the sky-view factor was computed using the GRASS GIS (GRASS, 2006) environment in the following way: given a location, the maximum elevation angle of sky obstruction was computed for 36 directions around this location. Then, a continuous curve of angle values was computed using linear interpolation. The area under this continuous curve, normalized to the entire hemisphere is considered the sky-view factor. Note that the value of the sky-view factor varies from 1, when the whole sky is obscured, to 0, when no obstructions are present. In actual cases, sky-view factor seldom is higher than 0.20. In this study, a modified version of this sky-view factor was used. Particularly, the sky-view factor was computed just using the topographic information in the south direction. That is, the angles were only computed for one half of the sky view, from solar azimuth angles between 90° (eastern) and −90° (western), instead of for the complete circumference (solar azimuth angles from −180° to 180°) (Fig. 4). We will call this new sky-view factor semi-sky-view factor hereinafter. As for the complete circumference, the semi-sky-view factor may take values between 0 and 1. Actual values for all the stations used in this study ranged between 0 and 0.195. This new explanatory variable revealed to be much more effective in explaining the solar radiation variability than a complete sky-view factor. The rationale behind this result is that, as it is known, in the northern hemisphere, the amount of radiation received from solar azimuth angles lower than −90° and greater than 90° is scarce.

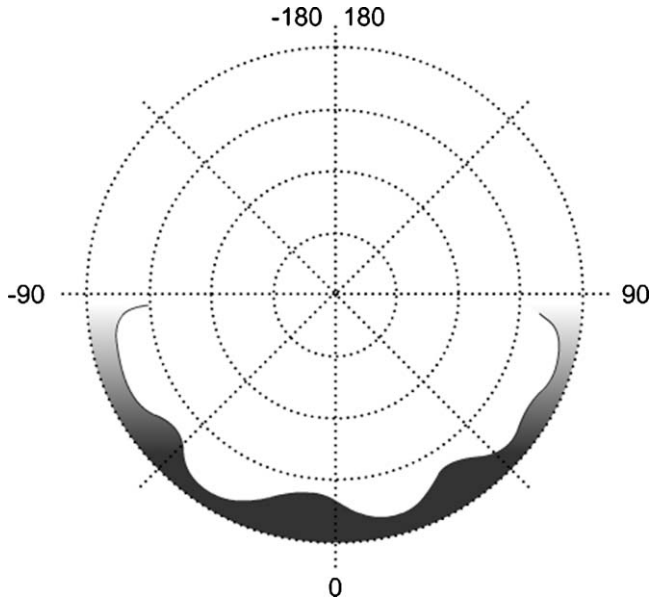

Fig. 4. Scheme showing the semi-sky-view factor used as explanatory variable in the residual kriging procedure. The values show the solar azimuth angle, the shaded area represent the area of sky occluded by the topography at the location of interest. Only the south hemi-half was used to compute the semi-sky-view factor.

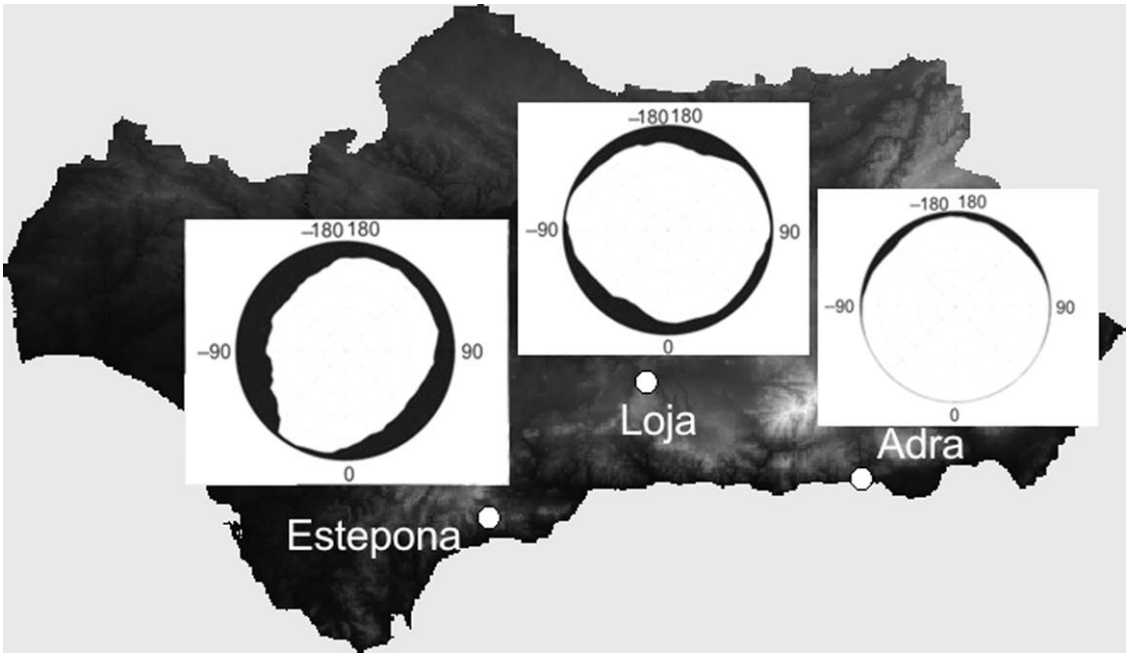

Fig. 5. Different examples of semi-sky-view factors. For the station of Adra, with no obstruction at solar azimuthal angles from 90° to −90°, the semi-sky-view factor is close to zero. For the Estepona case, with considerable obstruction in the range 90° to −90°, the semi-sky-view factor is 0.12. Finally, the Loja station presents an intermediate obstruction by the topography, with results in a 0.06 value of the semi-sky-view factor.

Therefore, the topographic information associated with these angles is not relevant in explaining the solar radiation variability and just increases the noise of the explanatory variable, without adding new information. This new semi-sky-view factor was used as independent variable in the residual kriging procedure.

Fig. 5 shows the map of the shadows caused by the topography for three stations representative of the study area. The Adra station (longitude 2°59′31″, latitude 36°44′52″, elevation 28 m) is located near the sea, with no obstruction at solar azimuthal angles between 90° and −90°. Therefore, the semi-sky-view factor value is close to 0. On the other hand, the Estepona station (longitude 4°27′30′, latitude 36°22′47″, elevation 310 m) is located in a valley with a high elevation ridge at the southern bound, resulting in a considerable obstruction at solar azimuthal angles between 90° and −90°. Particularly, the semi-sky-view factor value is 0.12 for this station (this means that 12% of the sky is obstructed at solar azimuthal angles between 90° and −90°). Finally, the Loja station (longitude 4°8′13″, latitude 37°10′14″, elevation 565 m) is surrounded by a relatively low elevation mountains range at the southern bound. This results in a semi-sky-view factor value of 0.06.

3.2. Regression analysis

A multiple regression analysis was carried out independently for each month using the elevation and the semi-sky-view factor as independent explanatory variables:

$$Y = b_1(\text{semi} - \text{sky} - \text{view factor}) + b_2(\text{elevation}) + c, \quad (12)$$

where Y is the estimated solar radiation value, b_1 and b_2 are the multiple regression coefficients for each independent variable, and c the intercept. A step-wise procedure was used for the regression parameter estimation. Particularly, the semi-sky-view factor was firstly regressed as independent variable and then the elevation was added. A t-test was carried for each step of the regression procedure and only parameters statistically significant at 5% level were further considered.

Table 1 shows the results of the multiple regression analysis for the training stations dataset. The most important explanatory variable is the semi-sky-view factor, which is statistically significant for all the months. This explanatory variable, which is negatively correlated with the solar radiation values, is able to explain from a minimum of 13% of the spatial variability (February) to a maximum of 45% in June. On the other hand, the elevation showed to be statically significant just from March to August and associated explained variance is considerable lower than for the semi-sky-view factor (values range from a minimum of 9.7% in August to a maximum of 15% in June). When considering both explanatory variables, explained variance ranges from a minimum of 13% in February to a maximum value of 46.7% in June. Considerable higher values are found during the summer than during the rest of the year, probably because during these months the amount of clear-sky days is higher than during the rest of the year. As a consequence, the topographic "signal" is "cleaner" than that during the rest of the months, when the highly variable weather influence (cloud cover, water vapour content, etc.) plays an important role. Nevertheless, the explained variance during the autumn and winter months is also relevant (about 33% in October). Therefore, it can be concluded that the effect of shadow cast by relief can be potentially higher in winter compared to the rest of the year due to the low elevation angle of the sun.

Table 1
Results of the multiple regression analysis for the training dataset. The first two columns show the explained variance provided by the semi-sky-view factor alone and the elevation alone (only during months with statistically significant correlations). The third to sixth columns show the linear regression parameters of multiple-regression analysis. Parameters b_1 and b_2 are the regression coefficient of, respectively, the semi-sky-view factor and elevation explanatory variables, while c stand for the intercept. Finally, the last column shows the explained variance of the multiple regression models. Only statistically significant values at the 5% level are displayed. Elevation data are in meters.

Month	R^2 (%) (sky-view factor)	R^2 (%) (elevation)	b_1	b_2	c	R^2 (%)
January	20.12		−17		9.721	20.12
February	13.00		−11.9		12.279	13.00
March	22.77	10.45	−23.8	0.001	16.984	24.2
April	30.21	11.38	−25.7	0.002	21.425	31.90
May	38.02	12.03	−42.8	0.003	24.894	39.00
June	45.40	14.84	−58.1	0.003	26.771	46.70
July	41.26	10.60	−51.3	0.002	27.814	42.69
August	44.80	9.79	−56.5	0.002	24.942	45.23
September	37.98		−51.2		20.896	37.98
October	33.02		−47		13.118	33.02
November	15.16		−16.1		10.46	15.16
December	22.12		−18.8		8.476	22.12

3.3. Semivariogram calculations

In this section, the semivariagrams obtained both for the ordinary and the residuals kriging methods are presented (Fig. 6). For the case of the ordinary kriging, the semivariagram is obtained based on, directly, the solar radiation training dataset. For the case of the residual kriging, the semivariagram is computed based on the residuals of the regression analysis carried out for the training dataset. Finally, 24 monthly semivariograms, 12 for the ordinary kriging method and 12 for the residual kriging methods, were estimated using the GSTAT package, integrated within IDRISI software (Clark Labs, 2006). None of the experimental semivariograms revealed the presence of parabolic shape near the origin. This means that the semivariograms appeared to be free of any structural component, as trend or drift. As a consequence, the ordinary krigring can be used, instead of universal kriging (which takes into account the spatial trend). Additionally, isotropy for spatial correlation was assumed in the entire study region. This

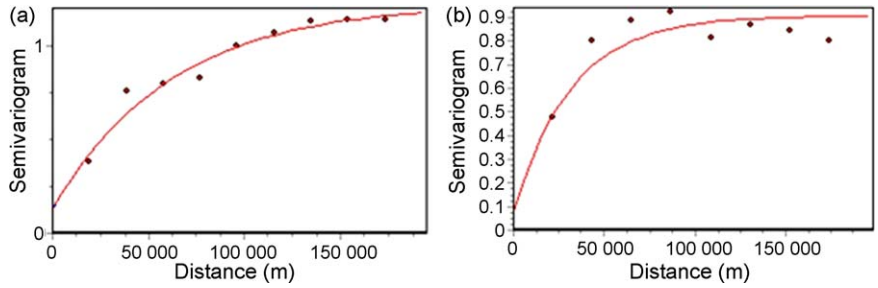

Fig. 6. January sample semivariograms, for the training data set, along with the fitted exponential model for raw data (a) and for the residuals data of the multiple regressions analysis (b).

Table 2
Semivariograms parameters for each month of the year computed based on the solar radiation training data set. Results are used for the ordinary krigring procedure.

Month	Nugget $(MJ\ m^{-2}\ day^{-1})^2$	Range (km)	Sill $(MJ\ m^{-2}\ day^{-1})^2$
January	0.19	63.5	1.4
February	0.1	18	1.84
March	0.2	29	3.2
April	0.2	20.5	3.3
May	0.1	37	7
June	0.18	38	4.05
July	0.7	25	5.4
August	0.5	28	6.3
September	0.2	29	8.25
October	0.54	25	5.3
November	0.13	24	2.1
December	0.12	23	1.3

Table 3
Semivariogram parameters for each month of the year computed based on the residuals time series resulting from the multiple regression analysis of the solar radiation training data set. Results are used for the residual krigring procedure.

Month	Nugget $(MJ\ m^{-2}\ day^{-1})^2$	Range (km)	Sill $(MJ\ m^{-2}\ day^{-1})^2$
January	0.08	32	0.83
February	0.15	21	1.65
March	0.13	23	1.8
April	0.23	21.7	2.1
May	0.18	19	2.7
June	0.15	22	1.85
July	0.31	21	3.05
August	0.33	17.5	3.2
September	0.2	24	5.5
October	0.3	18	4.3
November	0.1	31	1.9
December	0.11	30	1.13

means that semivariograms depend only on the distances between data pairs. According to Martinez-Cob (1996), the assumption of isotropy results in easier fitting of a theoretical model to experimental semivariogram. The weighted least-squares method was used to select the parameters (nugget effect, range and sill) of the theoretical semivariograms. In this method, the weights are the number of data pairs that provided the information to compute the experimental semivariogram. For all the cases, the exponential model was found suitable to fit the experimental semivariograms in our study area.

Overall, for the ordinary kriging, theoretical semivariogram models well fit the experimental data semivariograms. The model semivariogram parameters are listed, for each month, in Table 2, while in Fig. 6a, the case of the January month is presented. The nugget effect was rather small, with the lowest values, ~0.1 $(MJ\ m^{-2}\ day^{-1})^2$, recorded in February, May, November and December and the highest value, 0.7 $(MJ\ m^{-2}\ day^{-1})^2$, in July. Range values vary from a minimum of 18 km in February to a maximum of 63.5 km in January.

Table 3 present the results for the residual kriging procedure and Fig. 6b shows the case of January. Overall, the nugget, range, and sill parameter values tend to be (with some exceptions) considerably smaller in Table 3 than those in Table 2. This reflects the fact that explanatory variables used in the regression analysis partially account for the spatial variability of the solar radiation. The highest decrement in the parameter values is found for the sill; for this parameter an overall decrement of around 30% is found. Decrements of around 100% are found from May to September (the

Table 4
Evaluation results of the ordinary kriging procedure for the solar radiation at the training data set locations. The first column shows the observed spatial mean values. Units are all $MJ\ m^{-2}\ day^{-1}$ except for the R^2 coefficient. Values in brackets at the right show the error value in percentages. When applicable, the predicted-observed criterion was used.

Month		ME		MAE		RMSE		R^2
January	9.64	0.05	(0.5%)	0.81	(8.4%)	1.22	(12.6%)	0.94
February	12.0	0.03	(0.3%)	0.88	(7.3%)	1.23	(10.2%)	0.95
March	15.54	0.033	(0.2%)	1.11	(7.1%)	1.65	(10.6%)	0.91
April	20.31	0.022	(0.1%)	1.29	(6.4%)	1.92	(9.4%)	0.93
May	23.6	−0.1	(−0.4%)	1.3	(5.5%)	2.1	(8.9%)	0.94
June	26.48	−0.06	(−0.2%)	1.73	(6.5%)	2.31	(8.7%)	0.96
July	26.84	−0.018	(−0.1%)	2.07	(7.7%)	2.56	(9.5%)	0.95
August	24.0	−0.058	(−0.2%)	2.2	(9.2%)	2.78	(11.5%)	0.96
September	19.61	−0.05	(−0.3%)	2.27	(11.6%)	2.84	(14.4%)	0.96
October	13.11	−0.031	(−0.2%)	2.31	(17.6%)	2.81	(21.4%)	0.95
November	9.77	0.13	(1.3%)	1.37	(14.0%)	1.87	(19.1%)	0.94
December	8.15	0.06	(0.7%)	0.88	(10.8%)	1.29	(15.8%)	0.90

Table 5
As in Table 4 but for the validation dataset locations.

Month		ME		MAE		RMSE		R^2
January	9.9	−0.01	(−0.1%)	0.26	(2.6%)	0.60	(6.0%)	0.88
February	12.03	−0.01	(−0.08%)	0.28	(2.3%)	0.67	(5.61%)	0.77
March	15.53	−0.013	(−0.08%)	0.57	(3.7%)	1.44	(9.33%)	0.75
April	20.64	−0.02	(−0.1%)	0.50	(2.4%)	1.08	(5.20%)	0.94
May	23.87	−0.06	(−0.3%)	0.77	(3.2%)	1.52	(6.40%)	0.88
June	26.24	−0.046	(−0.2%)	0.81	(3%)	1.63	(6.20%)	0.96
July	26.68	−0.018	(−0.07%)	0.90	(3.4%)	1.80	(6.72%)	0.92
August	24.36	−0.03	(−0.1%)	0.81	(3.3%)	1.87	(7.70%)	0.94
September	19.68	0.01	(0.05%)	1.01	(5.1%)	2.09	(10.60%)	0.92
October	12.86	−0.03	(−0.2%)	0.78	(6%)	1.44	(11.21%)	0.94
November	10.05	−0.03	(−0.3%)	0.47	(4.8%)	1.05	(10.50%)	0.92
December	8.27	−0.03	(−0.4)	0.40	(4.8%)	0.93	(11.30%)	0.91

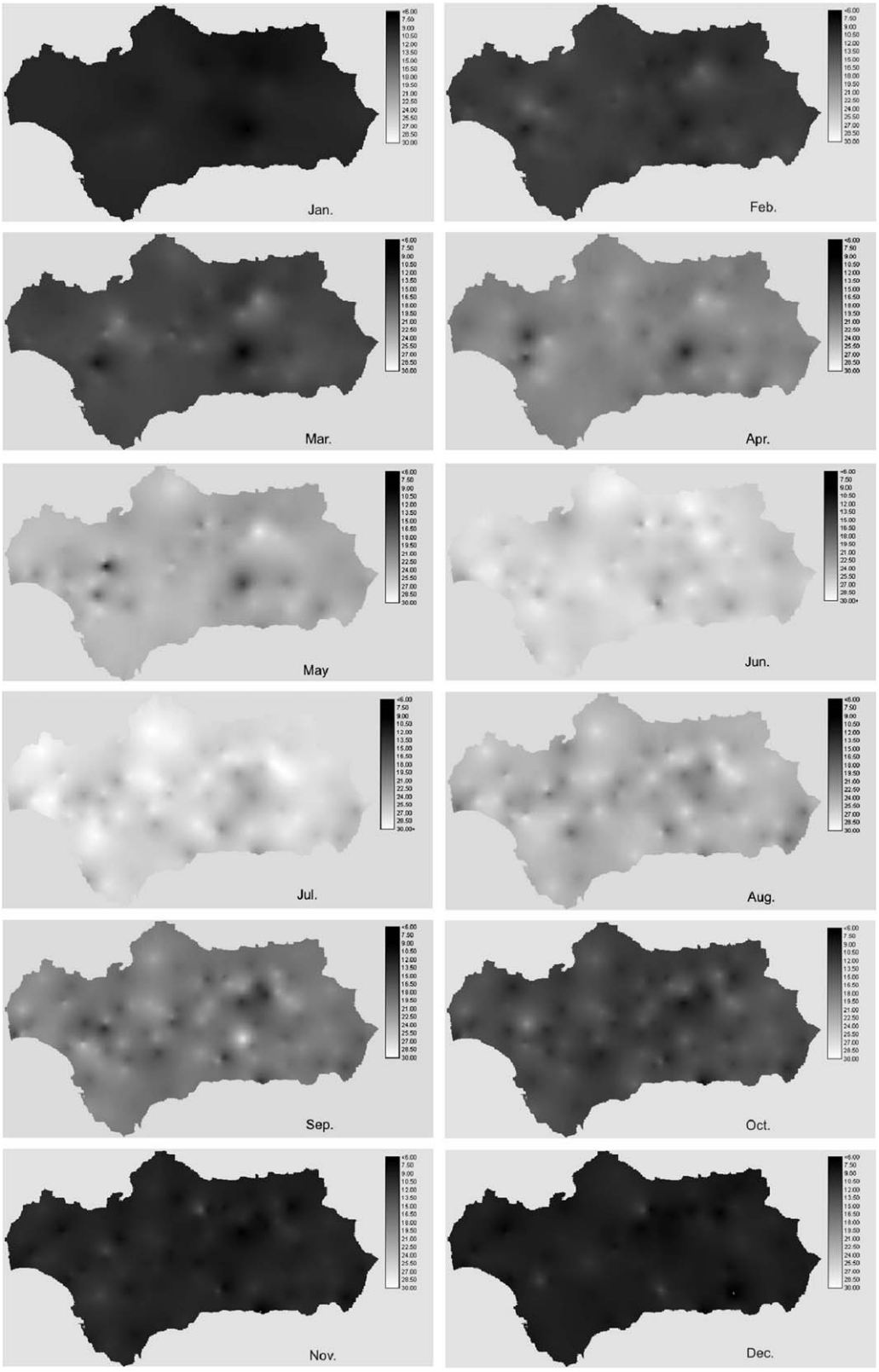

Fig. 7. Solar radiation maps obtained based on the ordinary kriging method on a 1 km × 1 km grid. Values (in MJ m^{-2} day^{-1}) are derived from the analysis of data corresponding to the period 2003–2006.

months with the highest explained variance in the multiple regression models). For the nugget and range parameters, the decrements in the values are not so important and, in some cases even smaller increments are found.

3.4. Models evaluation

Based on the results of previous sections, the monthly solar radiation has been estimated using both the ordinary and the

Table 6
As in Table 4 but for the residual kriging procedure.

Month		ME		MAE		RMSE		R^2
January	9.64	0.037	(0.4%)	0.72	(7.4%)	1.09	(11.30%)	0.96
February	12.0	0.021	(0.2%)	0.78	(6.5%)	1.11	(9.25%)	0.95
March	15.54	0.03	(0.2%)	1.01	(6.5%)	1.51	(9.71%)	0.96
April	20.31	0.017	(0.08%)	1.23	(6%)	1.78	(8.76%)	0.95
May	23.6	0.08	(0.3%)	1.21	(5.1%)	1.93	(8.18%)	0.96
June	26.48	0.05	(0.2%)	1.61	(6%)	2.15	(8.12%)	0.98
July	26.84	0.01	(0.04%)	1.91	(7.1%)	2.47	(9.20%)	0.97
August	24.0	0.047	(0.2%)	2.07	(8.6%)	2.59	(10.79%)	0.96
September	19.61	0.033	(0.2%)	2.14	(10.9%)	2.72	(13.87%)	0.97
October	13.11	0.03	(0.2%)	2.28	(17.4%)	2.76	(21.0%)	0.96
November	9.77	0.01	(0.1%)	1.19	(12.1%)	1.71	(17.5%)	0.96
December	8.15	0.047	(0.6%)	0.73	(8.9%)	1.2	(14.7%)	0.94

residual kriging procedure. Models estimates were evaluated in terms of the ME, MAE, RMSE, and the correlation coefficient.

3.4.1. Ordinary kriging model evaluation

Tables 4 and 5 presents the ordinary kriging method results for, respectively, the training and validation dataset. Training dataset results show, overall, that the ordinary kriging is able to provide fair estimates. Error values present a seasonal pattern: the summer months show the lowest RMSE values (in percentage) and the fall and the winter months the highest ones. Particularly, RMSE values ranges from around 2.31 MJ m^{-2} day^{-1} (8.7%) in June to around 2.81 MJ m^{-2} day^{-1} (21.4%) in October. Similarly, MAE ranges from around 1.73 MJ m^{-2} day^{-1} (6.5%) in June to a maximum of 2.31 MJ m^{-2} day^{-1} (17.6%) in October. Additionally, the ordinary kriging estimates show weak bias, the ME values being almost negligible for almost all the months. Only for November, a slight overestimation is found. Values of R^2 range from 0.9 in December to 0.96 in June. Regarding the validation dataset, Table 5 shows, again, a seasonal pattern in the errors. Nevertheless, RMSE values are considerably lower, especially for the winter months, with a decrement of about one-half compared to the values found for the training dataset. For instance, in October RMSE is 1.44 MJ m^{-2} day^{-1} (11.21%). For the summer months, the difference is just about one-third lower; particularly, for June RMSE is 1.63 MJ m^{-2} day^{-1} (6.20%). The MAE values decrement, compared to the training dataset, is even higher: values ranges from 0.26 MJ m^{-2} day^{-1} (2.6%) in February to 0.78 MJ m^{-2} day^{-1} (6%) in October. This means an overall reduction of about one half compared to the values found for the training dataset. Finally, ME values are similar for both dataset and R^2 shows slightly lower values than for the training data set. Fig. 7 shows the estimated solar radiation maps based on the ordinary kriging procedure. A fairly smooth pattern is observed, although some variability, associated with mountains areas, can be observed. The seasonal cycle is clear in the maps (that uses the same scale for all the months) with overall maximum radiation values in July and minimum in January.

3.4.2. Residual kriging model evaluation

Tables 6 and 7 show the evaluation results of the residual kriging models for, respectively, the training and validation datasets. For the training dataset, again, as for the ordinary kriging, a seasonal pattern is present, with the summer months showing the lowest RMSE values and the winter months the highest ones. Particularly, RMSE values range from around 2.15 MJ m^{-2} day^{-1} (8.12%) in the June to around 2.76 MJ m^{-2} day^{-1} (21%) in October. Similarly, MAE ranges from around 1.61 MJ m^{-2} day^{-1} (6%) in June to a maximum of 2.28 MJ m^{-2} day^{-1} (17.4%) in October. This error pattern seems to be related to the fact that the semi-sky-view factor explains more variability during summer months than during the winter months. Regarding the ME, the model tends to slightly overestimate the solar radiation, for all the months, but ME values are fairly low. Results for the validation dataset (Table 7) are considerably better than for the training dataset (similarly that in case of the ordinary kriging procedure). RMSE values are about a half of that found for the training period. Lowest RMSE relative value are found for February 0.54 MJ m^{-2} day^{-1} (4.5%) and the highest for October 1.31 MJ m^{-2} day^{-1} (10.2%). MAE values are also considerable lower for the validation dataset, showing an overall reduction of more than a half. Particularly, values ranges from 0.22 MJ m^{-2} day^{-1} (2.2%) in January to 0.70 MJ m^{-2} day^{-1} (5.4%) in October. ME values also are lower for most part of the months, while R^2 shows very similar values of those found in the training period. Fig. 8 shows the estimated solar radiation maps based on the residual kriging procedure. Visually, the surface obtained by the residual kriging was similar than the surface produced by ordinary kriging except that is more "spoty". This

Table 7
As in Table 6 but for the validation dataset locations.

Month		ME		MAE		RMSE		R^2
January	9.9	0.005	(0.05%)	0.22	(2.2%)	0.46	(4.61%)	0.94
February	12.03	0.01	(0.08%)	0.25	(2.0%)	0.54	(4.50%)	0.92
March	15.53	0.011	(0.07%)	0.49	(3.1%)	1.27	(8.22%)	0.94
April	20.64	0.018	(0.09%)	0.46	(2.2%)	0.93	(4.52%)	0.94
May	23.87	0.051	(0.2%)	0.63	(2.6%)	1.37	(5.71%)	0.95
June	26.24	0.031	(0.1%)	0.72	(2.7%)	1.44	(5.50%)	0.96
July	26.68	0.012	(0.04%)	0.78	(2.9%)	1.65	(6.20%)	0.95
August	24.36	0.021	(0.08%)	0.73	(3.0%)	1.57	(6.41%)	0.95
September	19.68	0.01	(0.05%)	0.92	(4.7%)	1.91	(9.70%)	0.95
October	12.86	0.024	(0.2%)	0.70	(5.4%)	1.31	(10.20%)	0.94
November	10.05	0.03	(0.3%)	0.40	(4.0%)	0.95	(9.43%)	0.94
December	8.27	0.042	(0.5%)	0.31	(3.7%)	0.81	(9.80%)	0.93

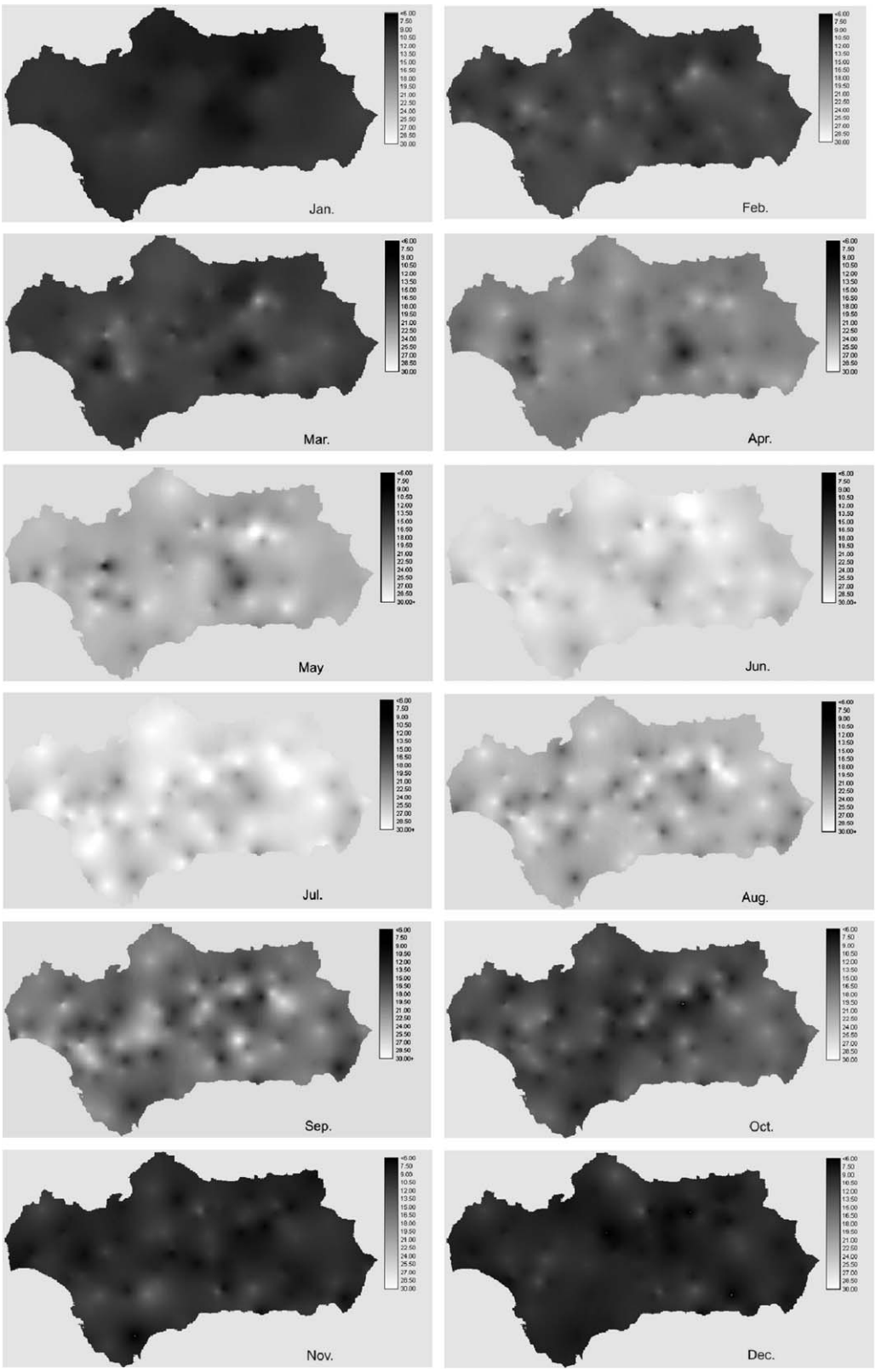

Fig. 8. . As in Fig. 7 but for the residual kriging method.

effect is related to the influence of the semi-sky-view factor, that adds spatial variability to the ordinary kriging maps.

3.4.3. Ordinary kriging versus residual kriging

The comparison of the ordinary and the residual kriging models results, in terms of the validation dataset (Tables 5 and 7), shows several interesting features. Residual kriging provides better estimates for all the months, with relative improvement in RMSE ranging from 5% to 20%. The maximum improvement is found for January (RMSE value of 0.46 versus 0.60, 23% for relative improvement). In June, relative improvement is about 5% and in October about 9%. Similar results are found for the MAE values,

with relative improvements ranging from 5% to 20%. Particularly, the maximum improvement is found during December (MAE 0.31 versus 0.4, 22% of relative improvement). A slight improvement is also found for the ME values, changing from a small underestimation using the ordinary kriging to almost negligible error values using the residual kriging. Finally, R^2 values show an overall improvement, being particularly high for January, February and March.

To analyze the rationale behind these results, the solar radiation predicted values, both using the ordinary kriging and residual kriging methods, were plotted against the observed values for the validation dataset locations. Particularly, Fig. 9 shows the June case (a month with a low RMSE value for the validation dataset using both ordinary and residual kriging methods, see Tables 5 and 7) and October (a month with a high RMSE value for the validation dataset, see Tables 5 and 7). Additionally, Tables 8 and 9 present different statistics (mean, standard deviation, maximum and minimum values) of the observed and estimated values (both using ordinary and residual kriging) for the validation dataset.

Both the ordinary and residual kriging solar radiation estimates fits quite well with the observed values (Fig. 9), but clearly the residual kriging provides better estimates. For June, the ordinary and residual kriging methods provide similar estimates for the highest values. Nevertheless, the intermediate and, particularly, the lowest values are clearly overestimated by the ordinary kriging, being fairly reproduced by the residual kriging model. This range of values (intermediate to low) seems to explain the overall reduction in RMSE and MAE values provided by the residual kriging. For instance, RMSE is 1.44 versus 1.63 MJ m^{-2} day^{-1} in June (Tables 5 and 7), that is a 5.5% of relative improvement in terms of RMSE. From a physical point of view, note that the semi-sky-view factor, as an explanatory variable in the residual kriging, acts just reducing the available solar radiation by the effect of the topography, while in no case increments the solar radiation. Therefore, for the summer months, characterized by high solar elevation angles, improvement provide by the residual kriging methodology should be associated with a better estimation of the lowest values. For the October month (Fig. 9), the improvement provided by the residual kriging method seems to be related to the whole range of values (although the most important improvement is found, again, for the lowest values). As a consequence, the improvement in the error scores is greater than for the summer case. For instance RMSE is 1.31 versus 1.44 MJ m^{-2} day^{-1} in October (Tables 5 and 7), that means a 9% of relative improvement in terms of RMSE. This result in an agreement with the fact that, due the lower elevation angles, the decrement in the solar radiation at the earth surface associated with shadows cast by topographic features is more important in the winter months that during the summer months.

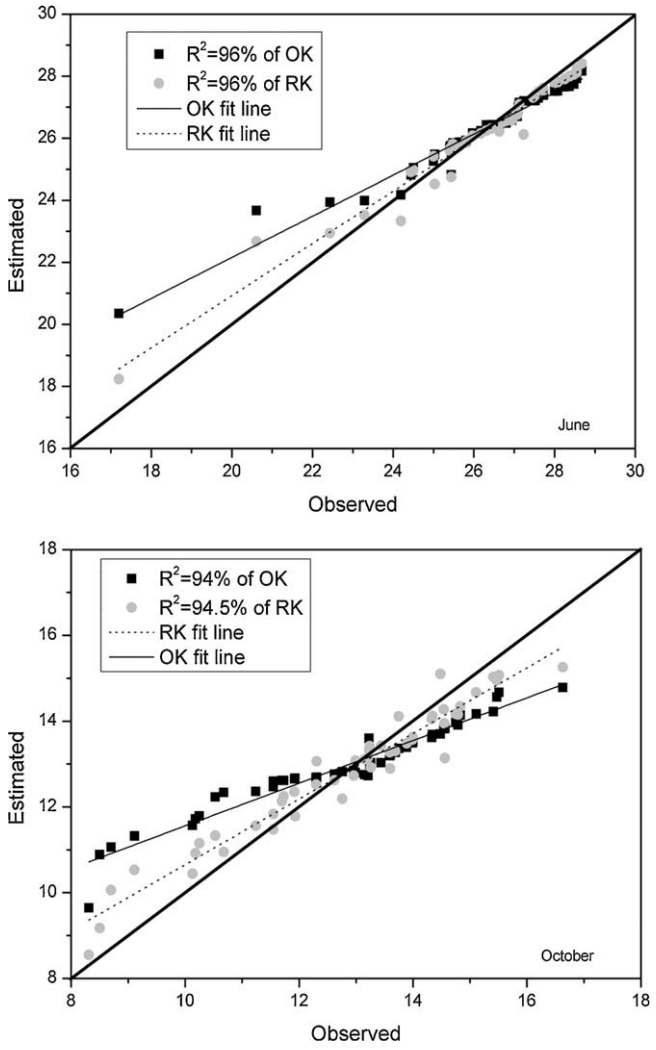

Fig. 9. Scarterplots of the observed and estimated, using both the ordinary kriging (OK) and residual kriging (RK) methods, global solar radiation values at the 54 locations of the validation data set. Top, for June, bottom, for October. The 1:1 line (thick solid line) is also displayed. Values are in MJ m^{-2} day^{-1}.

Solar radiation spatial mean values, averaged over the validation stations locations, are fairly reproduced by both

Table 8
Global solar radiation spatial mean and standard deviation (SD) values computed for the validation dataset locations. The observed values and the estimated using the ordinary kriging and the residual kriging methods are presented. Units are MJ m^{-2} day^{-1}. Values in brackets at the right show the absolute difference between the estimated and observed values.

Month	Observed values		Predicted values by OK				Predicted values by RK			
	Mean	SD	Mean		SD		Mean		SD	
January	9.64	0.99	9.65	(0.01)	0.62	(−0.37)	9.66	(0.02)	0.94	(−0.05)
February	12.0	1.36	11.98	(−0.02)	0.66	(−0.70)	12.06	(0.06)	1.07	(−0.29)
March	15.54	1.64	15.58	(0.04)	0.87	(−0.77)	15.61	(0.07)	0.97	(−0.67)
April	20.31	1.9	20.34	(0.03)	0.90	(−1.0)	20.36	(0.05)	1.71	(−0.19)
May	23.6	2.26	23.51	(−0.09)	1.54	(−0.72)	23.62	(0.02)	1.75	(−0.51)
June	26.48	1.89	26.41	(−0.07)	1.13	(−0.76)	26.46	(−0.02)	1.39	(−0.50)
July	26.84	2.36	26.8	(−0.04)	1.15	(−1.21)	26.78	(−0.06)	1.50	(−0.86)
August	24.0	2.53	23.94	(−0.06)	1.17	(−1.36)	23.97	(−0.03)	1.34	(−1.19)
September	19.1	2.72	19.59	(0.49)	1.23	(−1.49)	19.63	(0.53)	1.37	(−1.35)
October	13.11	2.41	13.07	(−0.04)	0.98	(−1.43)	13.14	(0.03)	1.10	(−1.31)
November	9.77	1.45	9.78	(0.01)	0.96	(−0.49)	9.81	(0.04)	1.19	(−0.26)
December	8.15	1.14	8.16	(0.01)	0.63	(−0.51)	8.21	(0.06)	0.87	(−0.27)

Table 9
As in Table 8 but for the maximum and minimum values.

Month	Observed values		Predicted values by OK				Predicted values by RK			
	Minimum	Maximum	Minimum		Maximum		Minimum		Maximum	
January	7.20	16.32	7.93	(0.73)	12.82	(−3.50)	7.47	(0.27)	14.22	(−2.10)
February	8.47	17.00	9.80	(1.33)	14.23	(−2.77)	9.21	(0.74)	15.61	(−1.39)
March	8.89	21.49	9.23	(0.34)	19.15	(−2.34)	9.11	(0.22)	19.94	(−1.55)
April	11.77	25.29	11.31	(−0.46)	23.78	(−1.51)	11.45	(−0.32)	24.57	(−0.72)
May	10.15	29.15	11.81	(1.66)	26.83	(−2.32)	10.97	(0.82)	28.10	(−1.05)
June	18.62	30.30	21.57	(2.95)	29.09	(−1.21)	19.21	(0.59)	29.17	(−1.13)
July	20.40	31.02	23.77	(3.37)	29.37	(−1.65)	21.76	(1.36)	29.88	(−1.14)
August	18.14	29.16	20.24	(2.10)	26.75	(−2.41)	19.16	(1.02)	27.12	(−2.04)
September	13.05	28.46	15.45	(2.40)	26.28	(−2.18)	14.32	(1.27)	27.31	(−1.15)
October	7.09	20.62	10.64	(3.55)	17.88	(−2.74)	8.37	(1.28)	18.91	(−1.71)
November	5.74	15.95	7.82	(2.08)	12.72	(−3.23)	6.21	(0.47)	14.68	(−1.27)
December	4.82	12.67	6.22	(1.40)	10.55	(−2.12)	6.10	(1.28)	11.37	(−1.30)

ordinary and residual kriging methodologies (Table 8). Particularly, both methodologies provide a maximum error (overestimation) of about 0.5 MJ m^{-2} day^{-1} (2.6%) during September. On the other hand, more substantial differences are found for the (spatial) standard deviations. Particularly, the residual kriging method provided considerable better estimates than the ordinary kriging (as can be expected based on results from Fig. 9). Maximum improvements are found for the winter months. For instance, for January, the observed standard deviation is 0.99 MJ m^{-2} day^{-1}, while the value provide by the ordinary kriging is 0.62 MJ m^{-2} day^{-1} (37% error) and the estimated using the residual kriging is 0.94 MJ m^{-2} day^{-1} (5% error). For the summer months, residual kriging also provides better standard deviation estimates, but improvements are lower.

Regarding the maximum values (Table 9), and as can be expected from Fig. 9, similar estimates are provided by both kriging methods during the summer months. On the other hand, residual kriging estimates are better for the winter months. For instance, for October, the observed maximum value is 20.62 MJ m^{-2} day^{-1}, while the value provide by the ordinary kriging is 17.82 MJ m^{-2} day^{-1} (13.1% error) and the estimated using the residual kriging is 18.91 MJ m^{-2} day^{-1} (8.3% error).

As far as minimum values are concerned, as can be expected from Fig. 9 and given that the semi-sky-view factor acts reducing the minimum value, residual kriging perform considerable better than the ordinary kriging method for all the months. Maximum differences are found for the summer months. For instance, for June, the observed minimum value is 18.62 MJ m^{-2} day^{-1}, while the value provide by the ordinary kriging is 21.57 MJ m^{-2} day^{-1} (15.8% error) and the estimated using the residual kriging is 19.21 MJ m^{-2} day^{-1} (3.1% error).

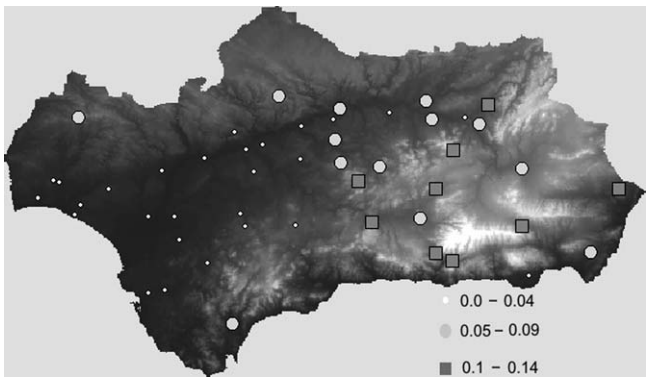

Fig. 10. Differences between ordinary kriging and residual kriging global solar radiation estimates, for October and at the evaluation dataset locations. Values are in MJ m^{-2} day^{-1}.

To summarize, it seems that most part of the improvement provided by the residual kriging method compared to the ordinary kriging method is associated with a better estimation of the minimum values, mainly for the winter months. Particularly, the ordinary kriging method overestimates the values in the lower range, since it does not account for the effect of shadows cast. From a spatial point of view, this effect is more important in complex topography areas. Fig. 10 presents, for the validation stations locations and for the October month, the difference between the ordinary kriging and residual kriging estimates. As expected, the stations located in complex topography areas, principally those with mountain ranges at the southern bound, shows higher ordinary kriging estimates than residual kriging estimates. This applies for most part of the stations located in the eastern part of the study region. On the other hand, stations located in the western part of the domain, with no major topographic features, show similar ordinary and residual kriging estimates.

4. Conclusions

Overall, the ordinary kriging method is able to provide fair estimates of the global solar radiation in southern Spain, with RMSE values ranges from 1.63 MJ m^{-2} day^{-1} (6.2%) in June to around 1.44 MJ m^{-2} day^{-1} (11.2%) in October. Nevertheless, by the inclusion of external information in the interpolation procedure, the residual kriging estimates shows lower errors than the ordinary kriging. Particularly, the inclusion of a semi-sky-view factor as explanatory variable (which accounts for topographic shadows cast) improve the RMSE values ranging from 5% in the summer months to more than 20% in the autumn and winter months. Explained variance also shows an improvement compared to the ordinary kriging method, with all month showing R^2 values above 0.92.

The results showed that both methods provide similar estimates during the summer months and for higher values of the global solar radiation. Nevertheless, intermediate and, particularly, lower values are clearly overestimated by the ordinary kriging, being fairly reproduced by the residual kriging model. This suggests that the semi-sky-view factor, as an explanatory variable, acts mainly reducing the available solar radiation by the effect of the topography. Therefore, for the summer months, characterized by high solar elevation angles, improvement provided by the residual kriging methodology should be associated with a better estimation of only the weaker values of the solar radiation. Regarding the autumn and winter months, the improvement provided by the residual kriging method, although is more important for the lower values, now is associated with the whole range of values. This explains the

greater improvement in the error scores found during the winter and autumn months as compared to summer months. This is in agreement with the fact of that, due to the lower elevation angles, the relative decrement in the solar radiation at the earth surface associated with shadows cast by topographic features is more important during the winter part of the year.

To sum up, it seems that most part of the improvement in the estimates provided by the residual kriging method compared to the ordinary kriging method is associated with a better estimation of the minimum values, particularly for the winter part of the year. The ordinary kriging method tends to overestimate the values in the lower range, since it does not account for the effect of shadows cast. From a spatial point of view, this effect is more important in complex topography areas. As expected, the stations located in complex topography areas, and particularly those with mountain ranges at the southern bound, show considerably higher ordinary kriging estimates than residual kriging estimates. This applies for most part of the stations located in the eastern part of the study region, where topographic discontinuities are higher than in other areas of Andalusia. On the other hand, stations located in the western part of the domain, with no major topographic features, show similar ordinary and residual kriging estimates. Then, proposed residual kriging method is particularly valuable when mapping complex topography areas.

The main utility of the solar radiation maps, as those here derived, are related to environmental and renewable energy applications. There are many solar resources spatial products in Europe for this kind of applications. For instance, the PVGIS Europe (Šúri et al., 2005) that uses data from meteorological stations and a relatively complex GIS-based solar radiation model, reports RMSE values about 5% for global radiation at the surface. The ESRA (Greif and Scharmer, 2000; Page et al., 2001) that uses both meteorological stations and satellite estimates, reports RMSE of the order of 8% for global radiation at the surface. Other products, as the Helioclim (Rigollier et al., 2004) based on satellite estimates, reports higher errors (RMSE about 25% for global radiation at the surface). Compared to these products and although its relatively simple formulation, our methodology provide similar error estimates. Nevertheless, the difficulties in a fair comparison between methodologies due the different spatial and temporal resolution, should be highlighted. In addition, the uncertainty in the solar radiation estimates based on the here-proposed model is of the same order of magnitude (about 5%) as of the measurement errors.

Finally, the authors believe the methodology here-presented could be improved. For instance, further work could be carried out to include new explanatory variables as the temperature. It is know that this variable have a relatively important relationship with the global solar radiation while there are, usually, a better spatial coverage of temperature measurements than solar radiation measurements.

Acknowledgements

The Spanish Ministry of Science and Technology (Project ENE2007-67849-C02-01) and the Andalusian Ministry of Science and Technology (Project P07-RNM-02872) financed this study. The data were kindly provided by the Andalusian Ministry of the Environment. We want to thank Mr. Mariano Corzo Toscano (Andalusian Regional Government) for helping in the data processing. H. Alsamamra is supported by a grant from the Spanish International Cooperation Agency (Spanish Foreign Office Ministry). The authors would like to acknowledge the very helpful comments made by two anonymous reviewers, which helped to substantially improve the manuscript.

References

Ahmed, S., de Marsily, G., 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resource Res. 23, 1717–1737.
Ashraf, M., Lofits, J.C., Hubbard, K.G., 1997. Application of geostatistics to evaluate partial weather station networks. Agric. Forest Meteorol. 84, 255–271.
Benzi, R., Deidda, R., Marrocu, M., 1997. Characterization of temperature and precipitation fields over Sardinia with principal component analysis and singular spectrum analysis. Int. J. Climatol. 17, 1231–1262.
Beyer, H.G., Czeplak, Terzenbach, U., Wald, L., 1997. Assesment of the method used to construct clearness index maps for the new European solar radiation atlas (ESRA). Solar Energy 61, 389–397.
Burgess, T.M., Webster, R., 1980. Optimal interpolation and isarithmic mapping of soil properties. Eur. J. Soil Sci. 31, 333–341.
Burrough, P.A., McDonnell, 1998. Principles of Geographical Information Systems. Claredon Press, Oxford, p. 333.
Buytaert, W., Célleri, R., Biévre, B., Cisneros, F., Wyseure, G., Deckers, J., Hofstede, R., 2006. Human impact on the hydrology of the Andean páramos. Earth-Sci. Rev. 79, 53–72.
Carrera-Hernández, J.J., Gaskin, S.J., 2007. Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico. J. Hydrol. 336, 231–249.
Chuvieco, E., Salas, J., 1996. Mapping the spatial distribution of forest fire danger using GIS. Int. J. Geogr. Inf. Syst. 10, 333–345.
Clark Labs, 2006. IDRISI Software. Clark University, MA, USA.
Cooper, R.M., Istok, J.D., 1988. Geostatistics applied to groundwater contamination. J. Environ. Eng. 114, 287–289.
Cressie, N., 1985. Fitting variogram models by weighted least squares. Math. Geol. 17 (5), 563–586.
Delhomme, J.P., 1978. Kriging in the hydrosciences. Adv. Water Resource 1, 251–266.
Dubayah, Ralph, Rich Paul, 1995. Topographic solar radiation models for GIS. Int. J. Geogr. Inf. Sci. 9, 405–419.
Ertekin, C., Evrendilek, F., 2007. Spatio-temporal modeling of global solar radiation dynamics as a function of sunshine duration for Turky. Agric. Forest Meteorol. 145, 36–47.
Font-Tullot, I., 1988. Historia del Clima de España: Cambios Climáticos y sus Causas. Instituto Nacional de Meteorología, Madrid.
Fu, P., Rich, P.M., 2000. The Solar analyst, User Manual. USA.
Gotway, C.A., Hartford, A.H., 1996. Geostatistical methods for incorporating auxiliary information in the prediction of spatial variables. J. Agric. Biol. Environ. Stat. 1, 17–39.
Greif, K., Scharmer, K., 2000. ESRA: European Solar Radiation Atlas, Fourth edition. Commission of the European Communities by Presses de l'Ecole Ecole des Mines de Paris, France.
GRASS (Geographical Resources Analysis Support System), 2006. GIS, http://grass.itc.it/.
Hargy, V.T., 1997. Objectively mapping accumulated temperature for Ireland. Int. J. Climatol. 17, 909–927.
Hetrick, W.A., Rich, P.M., Barnes, F.J., Weiss, S.B., 1993. GIS-based solar radiation flux models. In: Proceedings of the ASPRS-ACSM Annual Convention, vol. 3, Bethesda, MD, pp. 132–143.
Hofierka, J., Súri, M., 2002. The Solar radiation model for open source GIS: implementation and application. In: Proceedings of the Open source GIS-GRASS Users Conference. pp. 1–19.
Hughes, J.P., Lettenmaier, D.P., 1981. Data requirements for kriging: estimation and network design. Water Resource Res. 17, 1641–1650.
Hulme, M., Conway, D., Jones, P.D., Jiang, T., Barrow, E.M., Turney, C., 1995. Construction of a 1961–1990 European climatology for climate change modelling and impact applications. Int. J. Climatol. 15, 1333–1363.
Iqbal, M., 1983. An Introduction to Solar Radiation. Academic Press, Toronto, pp. 106–1067.
Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University, New York, p. 561.
Jarvis, E.H., Stuart, N., 2001. A comparison among strategies for interpolating maximum and minimum daily air temperatures. J. Appl. Meteorol. 40, 1075–1084.
Jorge, M., Javier, L., Pilar, G., 2003. Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. J. Hydrol. 270, 1–11.
Llasat, M.C., 1997. Metteorological conditions of heavy rains. FRIEND Projects H-5-5 and 1-1, UNESCO, 269–276.
Luca, B., Ducco, G., Donatelli, M., Stein, A., 2000. Modelling, interpolation and stochastic simulation in space and time of global solar radiation. Agric. Ecosyst. Environ. 81, 29–42.
Martinez-Cob, A., 1996. Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. J. Hydrol. 174, 19–35.
Menz, G., 1997. Regionalization of precipitation models in east Africa using meteosat data. Int. J. Climatol. 17, 1011–1027.
Merino, G.G., Jones, D., Stooksbury, D.E., Hubbard, K.G., 2001. Determination of semivariogram models to krige hourly and daily solar irradiance in western Nebraska. Am. Meteorol. Soc. 40, 1085–1194.
Mèszáos, I., Miklánek, P., 2002. Solar energy income modelling in mountainous areas. In ERB and NEFRIEND Proj.5 Conference, Interdiscplinary Approaches in Small Catchment Hydrology. Bratislava, Slovakia, 127–135.
Miklánek, p., Mèszáos, I., 1993. The estimation of energy income in grid points over the basin using simple digital elevation model. Ann. Geophys. 11 (Suppl. II), 296.

Mubiru, J., Karume, K., Majaliwa, M., Banda, E.J., Otiti, T., 2006. Interpolating methods for solar radiation in Uganda. Theor. Appl. Climatol. 88, 259–263.
Nalder, I.A., Wein, R.W., 1998. Spatial interpolation of climatic normals: test of a new method in Canadian boreal forest. Agric. Forest Meteorol. 92, 211–225.
Ninyerola, M., Pons, X., Roure, M.J., 2000. A methodological approach of clomatological modelling of air temperature and precipitation through GIS techniques. Int. J. Climatol. 20, 1823–1841.
Odeh, I.O., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes. Geoderma 67, 215–226.
Page, J., Albuisson, M., Wald, L., 2001. The European Solar Radiation Atlas: a valuable digital tool. Solar Energy 71, 81–83.
Phillips, D.L., Dolph, J., Marks, D., 1992. A comparison of geostatistical procedures for spatial analysis for precipitation in mountainous terrain. Agric. Forest Meteorol. 58, 1031–1047.
Pons, X., Ninyerola, M., 2008. Mapping a topographic global solar radiation model implemented in a GIS and refined with ground data. Int. J. Climatol. 28, 1821–1834.
Prudhomme, C., Reed, D.W., 1999. Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. Int. J. Climatol. 19, 1337–1356.
Rehman, S., Ghori, S.G., 2000. Spatial estimation of global solar radiation using geostatistics. Renewable Energy 21, 583–605.
Rigollier, C., Lefèvre, M., Wald, L., 2004. The method Heliosat-2 for deriving shortwave solar radiation data from satellite images. Solar Energy 77 (2), 159–169.
Ruiz-Arias, J.A., Tovar-Pescador, J., Pozo-Vazquez, D., Alsamamara, H., in press. A comparative analysis of DEM-based models to estimate solar radiation in mountainous terrain, Int. J. Geogr. Inf. Syst.
Sen, Z., Sahin, A., 2001. Spatial interpolation and estimation of solar irradiation by cumulative semivariograms. Solar Energy 71, 11–21.
Šúri, M., Huld, T.A., Dunlop, E.D., 2005. PVGIS: a web-based solar radiation database for the calculation of PV potential in Europe. Int. J. Sustain. Energy 24, 55–67.
Tabios, G.Q., Salas, J.D., 1985. A comparative analysis of techniques for spatial interpolation of precipitation. Water Resource Res. 21, 365–380.
Tovar-Pescador, J., Pozo-Vázquez, D., Ruiz-Arias, J.A., Batlles, J., Lóoez, G., Bosch, J.L., 2006. On the use of the digital elevation model to estimate the solar radiation in areas of compex topography. Meteorol. Appl. 13, 279–287.
Uran, C., Yun, J., 2004. Solar irradiance-corrected spatial interpolation of hourly temperature in complex terrain. Agric. Forest Meteorol. 126, 129–139.
Ustrnul, Z., Czekierda, D., 2005. Application of GIS for the development of climatological air temperature maps: an example of Poland. Meteorol. Appl. 12, 43–50.
Vic Barnett, 2004. Enviromental Statistics. John Wily and Sons, Chichester, England, p. 235.
Watson, I.D., Johnson, G.T., 1987. Graphical estimation of sky view-factors in urban environments. Int. J. Climatol. 7, 193–197.
Webster, R., Oliver, M.A., 2001. Geostatistics for Environmental Scientist. John Wiley and Sons, Chichester, England, p. 149.
Wilson, J.P., Gallant, J.C., 2000. Terrain Analysis: Principles and Applications. John Wiley and Sons, New York, p. 212.
Zelenka, A., Czeplak, G., D'Agostino, V., Josefson, W., Maxwell, E., Perez, R., 1992. Techniques for Supplementing Solar Radiation Network Data. Technical Report, International Energy Agency # IEA-SHCP-9D-1. Swiss Meteorological Institute, Switzerland.
Zhao Chuanyan, Nan Zhongren, Cheng Guodong, 2005. Methods for modelling of temporal and spatial distribution of air temperature at landscape scale in the southern Qilian mountains. China Ecol. Model. 189, 209–220.