Spatial Estimation of Reference Evapotranspiration in Andalusia, Spain

KARL VANDERLINDEN

Centro Las Torres-Tomejil, IFAPA, Junta de Andalucía, Seville, Spain

JUAN VICENTE GIRÁLDEZ

Departamento de Agronomía, Universidad de Córdoba/IAS-CSIC, Córdoba, Spain

MARC VAN MEIRVENNE

Department of Soil Management and Soil Care, Ghent University, Ghent, Belgium

(Manuscript received 19 January 2007, in final form 25 July 2007)

ABSTRACT

Knowledge of the spatial and temporal distribution of reference crop evapotranspiration (ET_0) is of interest for regional water resources management, especially in areas of the world where fine-tuning of agricultural water demands over large areas is required. This study provides a strategy for mapping ET_0 in regions with low meteorological data availability. For Andalusia, Spain, it involves estimating ET_0 from temperature data using a locally calibrated version of the Hargreaves equation and the application of geostatistical interpolation techniques that take into account elevation as secondary information. Average annual ET_0 at 191 observatories (with elevation between 0 and 1260 m) ranged from 954 to 1460 mm, with an average of 1283 mm, a standard deviation of 99 mm, and a correlation coefficient with elevation of −0.86. Simple kriging with varying local means (SKlm) and kriging with an external drift (KED)—two methods that take into account elevation as secondary information—increased spatial model efficiency by 30% as compared to ordinary kriging. SKlm was used for mapping ET_0 since it better reproduced the descriptive statistics of the point data and yielded slightly smaller root-mean-squared estimation errors than KED. The spatial correlation of annual and monthly ET_0 was well structured and anisotropic. Short-range variability, for separation distances up to 20–40 km, showed a strong linear increase with distance while long-range variability, up to 130–250 km, increased more gently with distance. The results of this structural analysis are relevant for the spatial optimization of a recently installed automated ET_0 observation network, while obtained maps constitute a valuable tool for regional water resources evaluation, planning, and management and contribute to optimizing water use in local irrigated agriculture.

1. Introduction

Periodic shortages in water resources cause great concerns among planners and administrators in regions in which seasonality of rain and its irregular occurrence demand rational management. This is the case in regions of Mediterranean climate, as in Spain and, especially, in Andalusia, where rainy periods are followed by long dry spells with high temperatures, causing large losses of water by evaporation. The spatial distribution of reference crop evapotranspiration (ET_0) is one of the key factors for water resources management in Andalusia. Recently, a network of automatic agrometeorological stations has been installed throughout the irrigated zones of Andalusia (IFAPA 2007). Data from this network became available since the year 2000, making them irrelevant for the calculation of long-term climatological averages or trends. Further research is necessary before these data can be integrated in the existing long-term observations provided by the network of the Spanish National Institute of Meteorology.

It is difficult to characterize variations in ET_0 on a regional scale in view of its dependence on many local factors such as net radiation, wind, relative humidity, and temperature, variables that are all interrelated and

Corresponding author address: Karl Vanderlinden, Centro Las Torres-Tomejil, IFAPA, Junta de Andalucía, Ctra. Sevilla-Cazalla, km 12, 2, 41200 Alcalá del Río, Seville, Spain.
E-mail: karl.vanderlinden@juntadeandalucia.es

DOI: 10.1175/2007JHM880.1

linked to the local geography. Knowledge of the spatiotemporal distribution of ET_0 at a regional scale enables the calculation of the required amount of water for adequate plant growth using established crop coefficients.

The Penman (1948) equation, or any of its variants (Allen et al. 1998), is generally considered as a reference method for calculating ET_0 but in many situations it is not applicable due to its high data demands. Methods based on temperature are more attractive when data availability is low. For years, the Thornthwaite (1948) method was very popular, but as a result of comments by Jensen et al. (1990) it was excluded from any subsequent use. The Hargreaves method (Hargreaves et al. 1985; Hargreaves 1994; Hargreaves and Allen 2003), only requiring minimum and maximum monthly mean temperatures, was selected as being the method with the most similar results to the Penman equation in a study made in 1992 by the Common Investigation Center of the European Commission and commented on by Hargreaves (1994). Shuttleworth (1993) recommends the Hargreaves method when temperature is the only variable available and cautions against making estimations of ET_0 for periods shorter than a month. Here we will use a previously calibrated version of this equation (Vanderlinden et al. 2004) to calculate the mean monthly and annual ET_0 in Andalusia.

Maps showing the spatial distribution of the annual and/or monthly ET_0 are highly useful and contribute to improved regional water resources management, including irrigation scheduling and water balance modeling. Vanderlinden et al. (2005) used the average annual ET_0 map for evaluating the average annual soil water balance in Andalusia.

In this work we deal with the spatial interpolation of mean monthly and annual values of ET_0, calculated from the mean monthly minimum and maximum temperatures, T_n and T_x, respectively. The spatial resolution (grid spacing) should depend on the degree of detail required to study the phenomenon, being conditioned by the number of available data, its density, and degree of spatial variability.

The theory of regionalized variables (Matheron 1965), and the geostatistical methods based on it, constitute an effective tool for assessing the spatial distribution of ET_0. Martínez-Cob (1996) used these methods to produce regional maps of ET_0 in northern Spain. The methodology developed in studies dealing with the spatial interpolation of precipitation, such as those by Phillips et al. (1992) or Goovaerts (2000), can also be useful for the spatial interpolation of ET_0 given the analogy of the problem.

There are other approaches to the regionalization of meteorological or climatological variables. Viau et al. (2000) determined a representative radius of the observations of the maximum daily temperature in Andalusia, by means of the surface temperature derived from National Oceanic and Atmospheric Administration (NOAA) images. They concluded that the mean representative radius of daily maximum temperature observations is 5 km, and under 3 km for over half the observatories. They also observed that the representative radius increased during dry periods or years. However, this could be an effect of the absence of vegetation during these periods, leading to an increased spatial homogeneity of the energy and water exchange at the land surface, and consequently an increased spatial extend of the flux footprint of the measurements (Schmid 2002). Kim et al. (2006) used a similar approach based on high-resolution *IKONOS* images to evaluate the flux footprint at a savannah and forest site. Stratification is also often used for regionalization of locally measured values (Stein et al. 1991; Van Meirvenne et al. 1994; Goovaerts 1997) since it provides the simplest way to incorporate secondary information into spatial estimation. A first regionalization of ET_0 can be made with the mean value of each stratum by dividing the population into subpopulations with a smaller variance than the original. In the case of ET_0 the stratification can be based on the elevation, the distance to the coast, or according to the boundaries of the eight provinces since the latter frequently coincide with geographical barriers such as mountain ranges or river valleys.

The objectives of this study are (i) to develop and validate a methodology for the spatial interpolation of ET_0 incorporating correlated secondary information, and (ii) to provide maps of mean monthly and annual ET_0 for Andalusia.

2. Materials and methods

a. Climatological data

The meteorological data used in this study were provided by the Spanish National Institute of Meteorology. Mean monthly minimum and maximum temperatures, T_n and T_x, respectively, were calculated from manually observed daily series of variable lengths and corresponded to the period between 1920 and 1998 (Vanderlinden 2002). Fifty percent of daily temperature records started before 1970. The average length of the records was 26 yr, and 50% of the records had a length between 19 and 36 yr. Observatories with more than a 15-yr record were retained to obtain a sufficiently dense spatial coverage of the study area only. Figure 1 shows a digital elevation model (DEM) for Andalusia and the location of the retained 191 climatological observato-

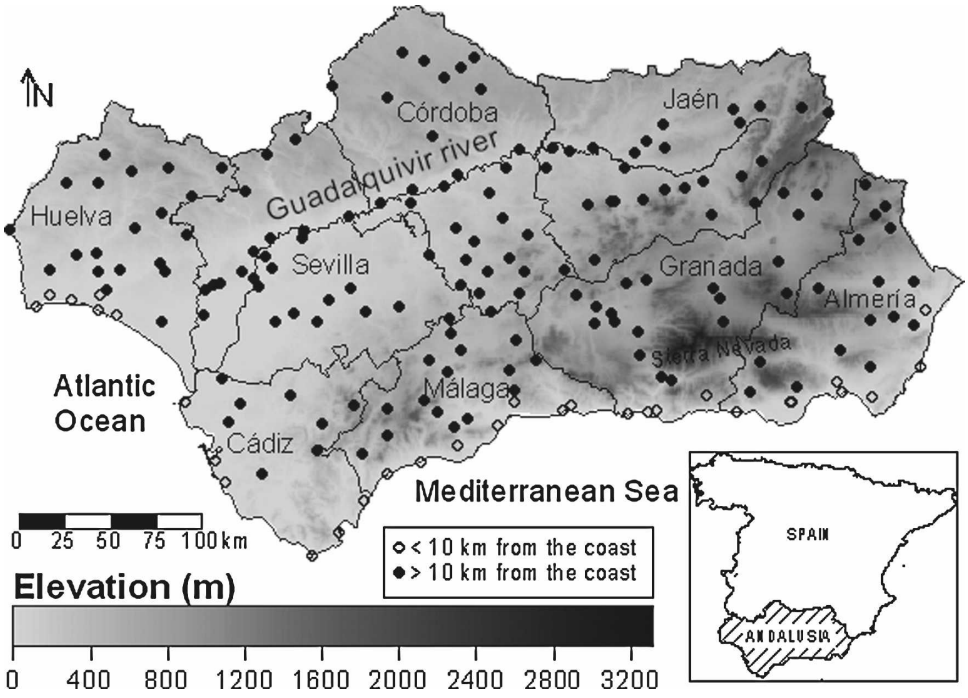

FIG. 1. Location of the 160 inland (>10 km from the coast) and 31 coastal observatories in Andalusia. Shades of gray represent elevation. Also, the political boundaries of the eight provinces of Andalusia are shown.

ries. The distance between observatories ranged from 1.3 to 31.5 km, with an average distance of 12.5 km. Thirty-one of them are situated at less than 10 km from the coast. These data were also used by Vanderlinden et al. (2005) for modeling the average annual water balance in Andalusia.

b. ET_0 estimation

The mean ET_0 of month j ($j = 1, \ldots, 12$), with $n(j)$ days, was calculated using the following equation (Hargreaves 1994):

$$\mathrm{ET}_0^j = \sum_{i=1}^{n(j)} CR_{i,j}(T_j + 17.8)\sqrt{\Delta T_j}\,(\mathrm{mm\ month}^{-1}), \quad (1)$$

where $R_{i,j}$ (mm day^{-1}) is the extraterrestrial radiation for day i of month j, T_j (°C) is the mean temperature of month j, and ΔT_j (°C) is the difference between the mean maximum and minimum temperature of month j. The extraterrestrial radiation was calculated using the procedure described by Shuttleworth (1993). Coefficient C, which was equal to 0.0023 in the original Hargreaves (1994) equation, was obtained through the empirical relationship developed by Vanderlinden et al. (2004):

$$C = 0.0005\frac{T_a}{\Delta T_a} + 0.00159 \quad R^2 = 0.90, \quad (2)$$

where T_a is the mean annual temperature and ΔT_a is the annual mean of the daily temperature range. Note from Eq. (2) that C can be obtained exclusively from temperature data. The mean annual ET_0^a was calculated as

$$\mathrm{ET}_0^a = \sum_{j=1}^{12} \mathrm{ET}_0^j. \quad (3)$$

c. Coastal and topographic effects on ET_0

According to Doorenbos (1976), the influence of the sea on meteorological variables is especially apparent in the first 2 km inland and gradually decreases up to distances of 10–15 km. The proximity of the sea affects wind speed, air humidity, and temperature. Depending on the local conditions, an increase in elevation usually reduces the temperature and evaporation and increases the amount of rainfall and wind speed. Barry (1981) reports how elevation influences the meteorological variables determining ET_0. Atmospheric pressure, vapor pressure, temperature, and its daily variation decrease with increasing elevation. It should be noted that wind speed depends more on slope and orientation than on elevation in general.

Solar radiation significantly increases up to heights of 2000 m and from then on, it only increases moderately.

In high-altitude, snow-covered areas net radiation decreases with increasing elevation due to increases in the albedo. There are few studies on evapotranspiration in high mountain areas, but according to Barry (1981) it can be assumed that ET_0 diminishes almost linearly with elevation up to 1000 m, and from then on it does so more gradually, remaining almost constant from 2000 m and above. In the next section, methods to include these coastal and topographical effects on ET_0 are discussed, using the 1-km DEM of Andalusia as exhaustive secondary information.

d. Spatial estimation

Geostatistics are based on the theory of regionalized variables (Matheron 1965) and the random function concept (Goovaerts 1997) where each point observation is considered to be the result of a spatially correlated random process. The structure of this spatial correlation can be expressed by means of the experimental variogram, $\gamma(h)$, which is estimated using (Goovaerts 1997)

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [y(\mathbf{x}_i) - y(\mathbf{x}_i + h)]^2, \tag{4}$$

where $N(h)$ is the number of data pairs separated by a lag distance, h, and $y(\mathbf{x}_i)$ is an observation of the random function in point $\mathbf{x}_i$. Authorized theoretical variogram models (Goovaerts 1997) can then be fitted to the experimental variograms. Here we will use spherical and exponential variogram models, respectively:

$$\begin{aligned} &\gamma(h) = c_0 + c\left[1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3\right] \quad h \le a \\ &\gamma(h) = c_0 + c \quad h \ge a \end{aligned}, \tag{5}$$

$$\gamma(h) = c_0 + c\left[1 - \exp\left(\frac{-3h}{a}\right)\right], \tag{6}$$

where c_0 is the nugget effect, the variogram value at lag 0, c is the contribution to the total variance or sill ($c_0 + c$), and a is the range, the lag distance at which the sill is reached. Spherical models are used for experimental variograms that have a linear behavior at the first lags, while experimental variograms are more appropriate for parabolic forms.

The random function $Y(\mathbf{x})$ can be broken down as

$$Y(\mathbf{x}) = m(\mathbf{x}) + \xi'(\mathbf{x}) + \xi'', \tag{7}$$

where $m(\mathbf{x})$ is the deterministic component, $\xi'(\mathbf{x})$ is the stochastic spatially dependent term, and ξ'' represents a normally distributed noise term.

Ordinary kriging (OK; Goovaerts 1997) estimates Y^*_{OK}, the expected value of Y at an unobserved point:

$$Y^*_{\mathrm{OK}}(\mathbf{x}_o) = \sum_{i=1}^{n} \lambda_i^{\mathrm{OK}} y(\mathbf{x}_i),$$

with

$$\sum_{i=1}^{n} \lambda_i^{\mathrm{OK}} = 1. \tag{8}$$

It assumes that the deterministic component in Eq. (7) is unknown, but stationary within the search neighborhood centered on the unobserved point, $\mathbf{x}_o$; $y(\mathbf{x}_i)$ are the n neighboring observations of Y. By imposing the condition that the n OK weights, λ_i^{OK}, should add up to unity the constant local mean is filtered out and the estimation is ensured to be unbiased. The corresponding system of equations can be solved either in terms of omnidirectional variogram models, in the case of isotropic variability, or with directional variogram models in the presence of anisotropy in the spatial correlation structure.

Simple kriging with varying local means (SKlm; Goovaerts 1997, 1999) estimates Y^*_{SKlm}, the expected value of Y at an unobserved point, while taking into account explicitly the deterministic component of Eq. (7):

$$Y^*_{\mathrm{SKlm}}(\mathbf{x}_o) = m(\mathbf{x}_o) + \sum_{i=1}^{n} \lambda_i^{\mathrm{SKlm}}[y(\mathbf{x}_i) - m(\mathbf{x}_i)], \tag{9}$$

where $m(\mathbf{x}_o)$and $m(\mathbf{x}_i)$ are the known local varying means, at the unobserved point $\mathbf{x}_o$, and at the n neighboring points $\mathbf{x}_i$, respectively; $y(\mathbf{x}_i)$ represents the n neighboring observations of Y, and $\lambda_i^{\mathrm{SKlm}}$ the SKlm weights. These weights can be obtained by solving the simple kriging system using the covariance function of the residual random function, $R(\mathbf{x})$,

$$R(\mathbf{x}) = Y(\mathbf{x}) - m(\mathbf{x}). \tag{10}$$

The study area can be stratified according to one, or several, exhaustively known secondary variables, assuming their mean to be constant within each stratum. After smoothing the transition between the boundaries of the strata, the resulting map provides local varying means for SKlm. In other cases it is more advantageous to establish a direct relationship between principal and secondary variables. Knowing the secondary variable everywhere, it is possible to estimate, through this relation, local means at every point and use it in SKlm.

Kriging with an external drift (KED; Goovaerts 1997, 1999) is similar to SKlm [Eq. (9)], but the difference is in the way the local nonstationary mean, $m(\mathbf{x})$, is modeled, that is, as a linear function of the secondary or external variable, $z(\mathbf{x})$:

$$m(\mathbf{x}) = a_0(\mathbf{x}) + a_1(\mathbf{x})z(\mathbf{x}). \tag{11}$$

Coefficients $a_0(\mathbf{x})$ and $a_1(\mathbf{x})$ are assumed to be constant within every search neighborhood and are estimated through the KED system, and not by a calibration or a regression as in the case of SKlm. KED is especially useful in the case that the relationship between the two variables changes with the spatial scale since the local mean is estimated by a regression within the search area. This procedure, however, is methodologically less elegant than SKlm. It assumes that the relationship between each of these variables is linear and requires the inference of the residual variogram without knowing the residues. In general, the original variogram is used, provisionally in the direction with the greatest continuity, where the drift is less strong, and assumes that the trend has little influence on the first lags of the variogram, which are of major importance to the interpolation. This means that the trend within the search area with regard to the interpretation of the variogram is ignored, while the methodology intrinsically acknowledges a linear relationship between the primary and the secondary variable within the search area. However, this apparent inconsistency seems to have little influence on the final results, which are usually similar to those of SKlm (Goovaerts 1999, 2000).

Cross validation is used here to compare the performance of OK, SKlm, and KED for interpolating ET_0. The results obtained for OK are used as a reference, since it does not take into account secondary information. Cross validation is done by consecutively eliminating each data value from the dataset and then estimating it. Different statistical parameters are then used to compare estimated, $y^*(\mathbf{x}_i)$, and observed values, $y(\mathbf{x}_i)$, at locations $\mathbf{x}_i$. Here we will use (i) the root-mean-square error (RMSE),

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[y^*(\mathbf{x}_i) - y(\mathbf{x}_i)]^2}, \qquad (12)$$

which gives the precision or dispersion of the estimates against the observations, and (ii) the coefficient of efficiency (CE),

$$\mathrm{CE} = 1 - \frac{\sum_{i=1}^{n}[y^*(\mathbf{x}_i) - y(\mathbf{x}_i)]^2}{\sum_{i=1}^{n}[y(\mathbf{x}_i) - \bar{y}]^2}, \qquad (13)$$

which ranges from $-\infty$ to 1 and has been widely used to evaluate the performance of hydrological models (Nash and Sutcliffe 1970; Legates and McCabe 1999). Since it is the ratio of the mean square error to the variance of the observed data, subtracted from unity, $\mathrm{CE} > 0$ indicates that the interpolation method is a better estimator than the observed mean. All geostatistical calculations (OK, SKlm, and SKlm) have been executed using the geostatistical software library, "Gslib" (Deutsch and Journel 1998), and the variogram calculation and modeling software, "Variowin" (Panatier 1996).

TABLE 1. Descriptive statistics for the mean monthly and annual maximum (T_x) and minimum (T_n) temperature (n = 191); m = mean.

		T_x			T_n		
	Month	m (°C)	Range (°C)	CV (%)	m (°C)	Range (°C)	CV (%)
Monthly	1	13.9	6.6–17.5	16.2	4.3	−2.5–11.3	59.3
	2	15.3	8.4–18.6	14.0	5.2	−1.8–11.4	46.0
	3	18.2	9.6–22.2	12.6	6.7	−0.7–13.2	35.9
	4	20.1	13.2–23.8	10.4	8.3	1.3–14.8	27.2
	5	23.9	16.8–27.8	8.8	11.3	4.0–16.8	19.0
	6	28.7	21.8–33.9	8.5	14.8	6.8–20.5	14.1
	7	33.3	24.2–38.5	8.3	17.9	9.4–25.8	11.9
	8	33.1	24.8–37.3	7.4	18.1	8.9–25.9	12.3
	9	29.3	23.3–34.3	7.5	15.7	6.2–22.0	15.2
	10	23.2	16.9–26.8	8.6	11.8	3.9–18.1	20.9
	11	17.9	11.0–21.1	12.3	7.9	0.1–14.4	33.0
	12	14.5	6.1–18.7	16.1	5.3	−3.1–12.2	49.8
Annual	Total	22.6	16.4–25.8	8.4	10.6	2.7–16.3	21.4
	Inland	22.7	16.4–25.8	8.9	10.7	2.7–16.3	19.6
	Coastal	22.3	19.6–23.7	4.2	13.5	9.9–16.1	10.1

3. Results and discussion

a. Univariate descriptive statistics

The monthly T_x and T_n ranged from 6.1° to 38.5°C and from −3.1° to 25.9°C, respectively, with minimums observed in December and maximums in July and August (Table 1). The monthly minimum temperature showed the highest coefficient of variation (CV) reaching values of more than 40% during the winter. On an annual basis and for both temperatures, the CV for inland observatories was about twice as large as the one observed at coastal observatories, due to the attenuating effect of the sea on the temperature dynamics throughout the year (smaller maximum and larger minimum temperatures) and the uniform low elevation of the coastal observatories.

The ΔT reached a maximum during summer (15.3°C in July) and a minimum during winter (9.2°C in December), as expected. The mean total annual ΔT was 12.0°C, and 12.6° and 8.9°C for the inland and coastal observatories, respectively. The lower ΔT in the coastal area is caused by the attenuating effect of the sea on the daily temperature range and the relatively flat topography. At the inland mountainous areas ΔT is enhanced by cold air movement toward the valleys during the

TABLE 2. Descriptive statistics for ΔT, $T/\Delta T$, and C (n = 191); m = mean.

	m	Range	CV (%)
ΔT (total)	12.0	4.4–17.5	19.3
ΔT (inland)	12.6	7.1–17.5	15.1
ΔT (coastal)	8.9	4.4–11.0	17.8
$T/\Delta T$ (total)	1.46	0.70–3.95	30.1
$T/\Delta T$ (inland)	1.33	0.70–2.73	20.3
$T/\Delta T$ (coastal)	2.11	1.45–3.95	25.6
C (total)	0.002 32	0.001 94–0.003 57	9.4
C (inland)	0.002 26	0.001 94–0.00 95	6.0
C (coastal)	0.002 64	0.002 31–0.003 57	10.2

night. The CV of this variable did not greatly fluctuate throughout the year, reaching a maximum value of 22.8% in July and a minimum of 18.3% in April.

The descriptive statistics for the C coefficient and for the $T/\Delta T$ ratio, from which it was calculated [Eq. (2)], are shown in Table 2. Taking into account all the observatories, the mean C coefficient equaled 0.0023, which was identical to the original Hargreaves (1994) coefficient. However, for the coastal area, a coefficient of 0.0026 was obtained. Also, the CV was higher for coastal stations due to differences between the Atlantic and the Mediterranean coasts, and more climatic unstable conditions such as cloud formation or high wind speeds, observed in some coastal observatories, especially for the ones near the Gibraltar Peninsula. Using principal component analysis (empirical orthogonal functions), several studies (e.g., Maheras 1989; Serrano et al. 1999; Nieto et al. 2004) show evidence of the climatological differences between the Atlantic and Mediterranean coasts in terms of temperature and rainfall caused by different prevailing synoptic situations during the autumn, winter, and spring. However, most of these circulation patterns disappear during late spring, summer, and early autumn, when evapotranspiration is at its maximum rate.

Table 3 shows the descriptive statistics for ET_0, calculated using Eqs. (1)–(3). The mean monthly values varied between 38 mm (December) and 200 mm (July). The maximum monthly value of 233 m was observed in July and the minimum value of 21 mm was observed in December. The standard deviation (s) reached a maximum of 17.8 mm in July and a minimum of 5.6 mm in December. The CV, however, reached its maximum (14.9%) during December and reached its minimum (7.7%) during September, which indicates that the relative spatial variability of ET_0 during the winter was almost twice as large as compared to late summer. The mean annual ET_0 value was 1283 mm, with minor differences for inland (1286 mm) and coastal (1270 mm) observatories. Both the minimum (954 mm) and the maximum (1460 mm) values were observed in inland stations. The minimum and maximum observed in the coastal areas was 1146 and 1357 mm, respectively. Considering all the stations, s of the annual ET_0 equaled 99.3 mm. It increased to 106.5 mm for inland stations, while for the coastal observatories an s of 45.2 mm was observed. Table 3 also shows the descriptive statistics for the elevation of the observatories. The mean elevation was 394 m, with a minimum of 2 m and a maximum of 1260 m. This means that no information was available on the relationship between the different variables and an elevation above 1260 m up to Mulhacen peak in the Sierra Nevada (3482 m). The CV (85%) indicated the important variability in observatory elevation.

TABLE 3. Descriptive statistics for the mean monthly and annual ET_0 (mm) and for the elevation of the observatories (n = 191). In the table, m = mean, s = standard deviation, and s^2 = variance.

	Month	m	Min	Max	s	s^2	CV (%)
Monthly	1	40.9	24.1	54.6	6.0	35.7	14.6
	2	50.9	31.6	63.7	6.5	42.8	12.8
	3	85.0	49.3	101.7	9.8	96.2	11.5
	4	108.9	72.0	126.2	10.8	115.9	9.9
	5	145.5	101.6	168.9	12.4	154.0	8.5
	6	172.0	132.2	201.2	14.4	206.9	8.4
	7	200.2	159.0	233.3	17.8	317.3	8.9
	8	180.5	145.1	206.5	14.5	209.1	8.0
	9	129.0	98.5	149.5	10.0	100.0	7.7
	10	82.8	59.4	95.5	7.5	56.8	9.1
	11	49.5	32.3	61.8	6.15	37.9	12.4
	12	37.9	21.4	50.8	5.65	31.9	14.9
Annual	Total	1283.0	953.9	1460.4	99.3	9857.3	7.7
	Inland	1285.7	953.9	1460.4	106.5	11 349.3	8.3
	Coastal	1269.1	1146.4	1357.3	45.2	2039.5	3.6
Elevation (m)		393.6	2.0	1260.0	333.0	110 883.0	84.6

The annual dynamics of the mean monthly ET_0 at coastal and inland observatories can be observed from the box-and-whisker plots of Fig. 2. As a consequence of the attenuating effect of the sea on T_x, T_n, and ΔT, during winter ET_0 was lower, and in the summer it was higher at the inland observatories as compared to the coastal ones. This feature will be incorporated into the spatial interpolation of the monthly ET_0 by means of stratification.

b. *Bivariate statistics*

The Pearson correlation coefficients between observatory elevation and monthly and annual ET_0 are shown in Table 4. To provide insight on the spatial variability of the correlation coefficients, results were also provided separately for the eight provinces of Andalusia (Fig. 1). Without considering the low correla-

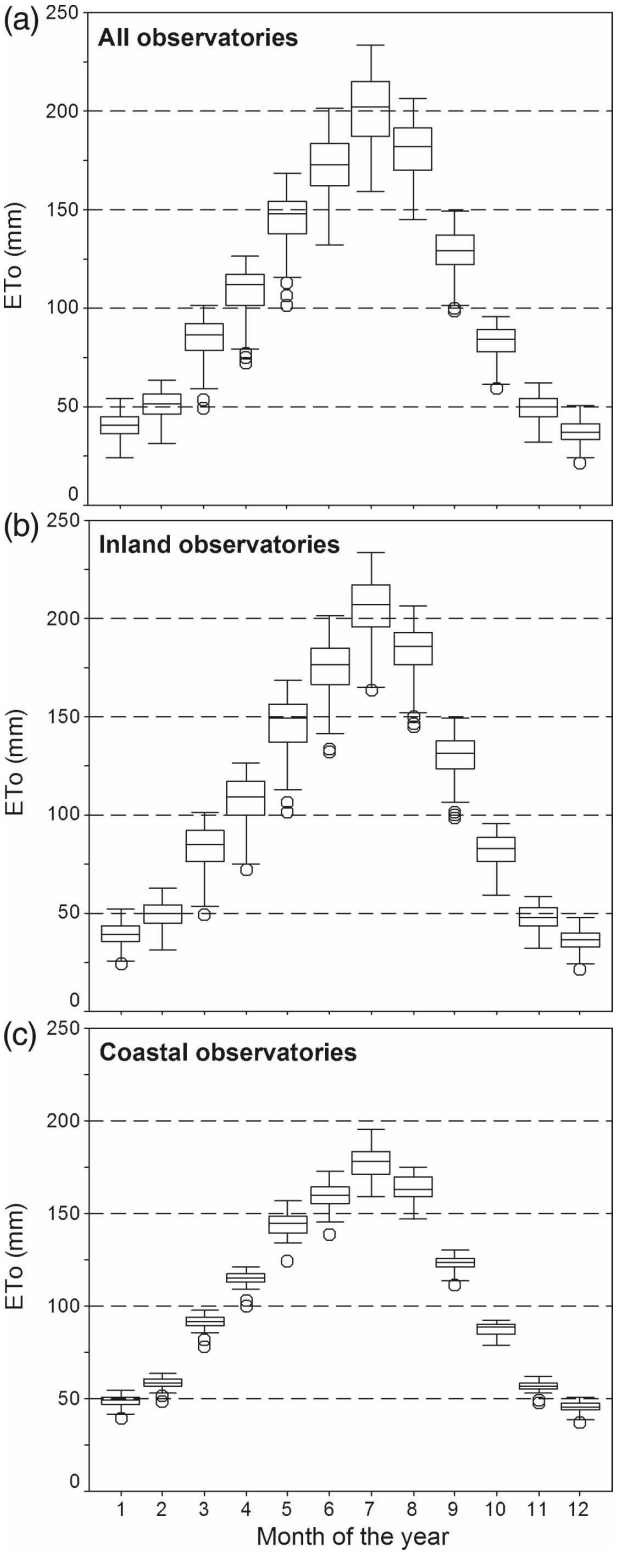

FIG. 2. Box-and-whisker plots for ET_0 for (a) all the observatories, (b) the inland observatories, and (c) the coastal observatories.

tion coefficients of the summer months (from June to September), correlations between −0.58 (at Córdoba in November and December) and −0.95 (Sevilla, in January) were observed. The spatial variability of ET_0 at the coastal observatories is mainly due to differences between the Atlantic and Mediterranean coasts or to site-specific windy or cloudy conditions (e.g., near the Gibraltar Peninsula), but in general not to elevation since all coastal observatories are located close to sea level. Therefore, eliminating these observatories, the overall correlation between ET_0 and elevation was improved from −0.75 to −0.86. The relationship of ET_0 with elevation was modeled using a quadratic polynomial function (Fig. 3). During the winter (e.g., January) the fitted functions behaved almost linearly, but in the summer (e.g., July) the curve flattened toward lower elevations, when the costal observatories were excluded. The goodness of fit (R^2) for these functions ranged from 0.78 in the spring (April) and autumn (October) to 0.46 in the summer (July). These relations were used to calculate the monthly local mean maps to be used in Eq. (9). None of the possible input variables to Eq. (1), T_x, T_n, T, ΔT, or C, showed a better relationship with elevation (not given), nor did the variable "distance to the coast" provide better relationships with these variables or ET_0 (not given). These results indicate that the coastal areas should be considered separately in the interpolation procedure using stratification. Doing so, better use could be made of secondary information (elevation) when interpolating the estimated ET_0 point data at the inland areas, without loss of information at the coastal areas.

c. *Variography and local means*

To interpret the evolution of the spatial correlation structure of ET_0 throughout the year, monthly normalized experimental variogram maps (i.e., the variogram divided by the sample variance) were calculated using lag intervals of 10 km up to a maximum lag of 140 km, and 180 km, in the north–south (N–S) and east–west (E–W) directions, respectively. These maps represent how the average variability among data pairs increases as a function of the length and the direction of the distance vector that separates them. As a consequence these maps are symmetric with respect to the (0, 0) separation distance point. Figure 4 shows the corresponding maps for the mean monthly ET_0 in January and July, and the annual ET_0. All the maps show a clear anisotropy in the spatial correlation structure, and a very low nugget effect. The nugget effect is the variogram map value at the (0, 0) point, which virtually represents the variability among observations separated by a zero distance. Theoretically this value should

TABLE 4. Pearson correlation coefficients between observatory elevation and mean monthly and annual ET_0. Values are given separately for the eight provinces, for all observatories, and for the inland observatories only. Bold entries are significant at the 0.05 probability level, regular roman at the 0.01 probability level, and italic entries are not significant.

	Almería	Cádiz	Córdoba	Granada	Huelva	Jaén	Málaga	Sevilla	All observatories	Inland observatories
Jan	−0.93	−0.90	−0.65	−0.85	−0.82	−0.76	−0.91	−0.95	−0.81	−0.79
Feb	−0.93	−0.90	−0.73	−0.87	−0.84	−0.79	−0.89	−0.95	−0.84	−0.82
Mar	−0.92	−0.75	−0.78	−0.86	−0.77	−0.81	−0.80	−0.83	−0.84	−0.84
Apr	−0.92	−0.78	−0.82	−0.85	−0.77	−0.88	−0.81	−0.91	−0.86	−0.88
May	−0.82	*−0.49*	−0.74	−0.61	**−0.39**	−0.88	−0.65	−0.81	−0.70	−0.82
Jun	−0.65	*−0.11*	−0.67	*−0.17*	*0.13*	−0.83	*−0.36*	−0.72	−0.42	−0.69
Jul	*−0.33*	*0.10*	−0.75	*−0.14*	*0.32*	−0.80	*0.06*	−0.68	−0.26	−0.66
Aug	−0.51	*0.09*	−0.75	*0.05*	*0.36*	−0.75	*−0.00*	−0.70	−0.32	−0.72
Sep	−0.72	*−0.20*	−0.70	−0.44	*0.01*	−0.77	**−0.47**	−0.85	−0.61	−0.83
Oct	−0.92	−0.73	−0.71	−0.85	−0.71	−0.81	−0.81	−0.94	−0.87	−0.88
Nov	−0.93	−0.96	−0.58	−0.87	−0.85	−0.81	−0.90	−0.92	−0.85	−0.83
Dec	−0.94	−0.90	−0.58	−0.83	−0.82	−0.79	−0.92	−0.92	−0.79	−0.77
Annual	−0.85	*−0.48*	−0.82	−0.65	*−0.38*	−0.86	−0.66	−0.89	−0.75	−0.86

be zero, but since there are no independent observations available at the same location, it is estimated from nearby observations. It accounts for the variability due to observational or procedural error and the variability that occurs beyond the smallest available separation distance between observations. In this case the low nugget effect indicates that the variability beyond the smallest separation distances between observatories was very small and that the temperature data used for ET_0 calculation were spatially consistent and of general good quality. This means also that the average annual and monthly ET_0 values were not affected by different record lengths and periods of the observatories. The long-distance structure is approximately maintained throughout the year, but there are notable variations in short-distance structure (dashed line). The direction with the greatest continuity varies between 15° in December and 34° in September (measured counterclockwise from the east). As expected, the correlation structure of the summer months dominates the annual pattern since these months contribute largely to the annual ET_0. The observed patterns are similar to the ones observed for elevation (Fig. 4d), indicating the impact of elevation on the spatial correlation structure of ET_0. The direction of major continuity (SW–NE) corresponds to the orientation of the Guadalquivir River valley, one of the areas with the greatest geographical and climatological continuity in Andalusia.

The observed anisotropy in the spatial correlation structure must be taken into account when modeling the spatial distribution of ET_0. We will model this large-scale directional pattern through the use of exhaustive secondary information (the observatory elevation and the 1-km^2 gridcell elevation data from the DEM) using KED and SKlm. In KED, the relationship with elevation is automatically incorporated during the interpolation, according to Eq. (11), while in SKlm secondary information is involved through local mean surfaces. Since the coastal observatories behaved differently and showed no relationship with elevation, the study area was divided into an inland and a coastal area. The local mean values for the inland area were calculated using the polynomial functions fitted in Fig. 3, while for the coastal area (<10 km from the coast) a constant value was used, equal to the mean coastal ET_0. However, this approach would imply an abrupt transition from one area to the other and would result in an artificial discontinuity in the final ET_0 map. Therefore we proceeded here as follows: pixels situated at less than 5 km from the coast were given a constant mean coastal ET_0 value while local mean values for pixels situated at more than 20 km from the coast were estimated according to the relationships shown in Fig. 3. For pixels situated between 5 and 20 km from the coast a progressive linear combination of the inland polynomial functions and the coastal observatory mean was calculated to create a smooth transition between the coastal and inland areas. No other approaches to calculate the varying local mean map were considered. The large-scale drift calculated in this way was used as the varying local mean in SKlm [Eq. (9)], using the variogram model of the residuals, which were calculated according to Eq. (10).

The variogram characterizes the spatial correlation structure of ET_0. Omnidirectional, directional, and residual experimental variograms were calculated and modeled for monthly and annual ET_0. The directional variograms were calculated in the direction of major continuity. For the mean annual ET_0 two scales of spatial variation could be distinguished (Fig. 5, top). As a consequence, an omnidirectional double spherical model was fit without a nugget, with a range of 29 km

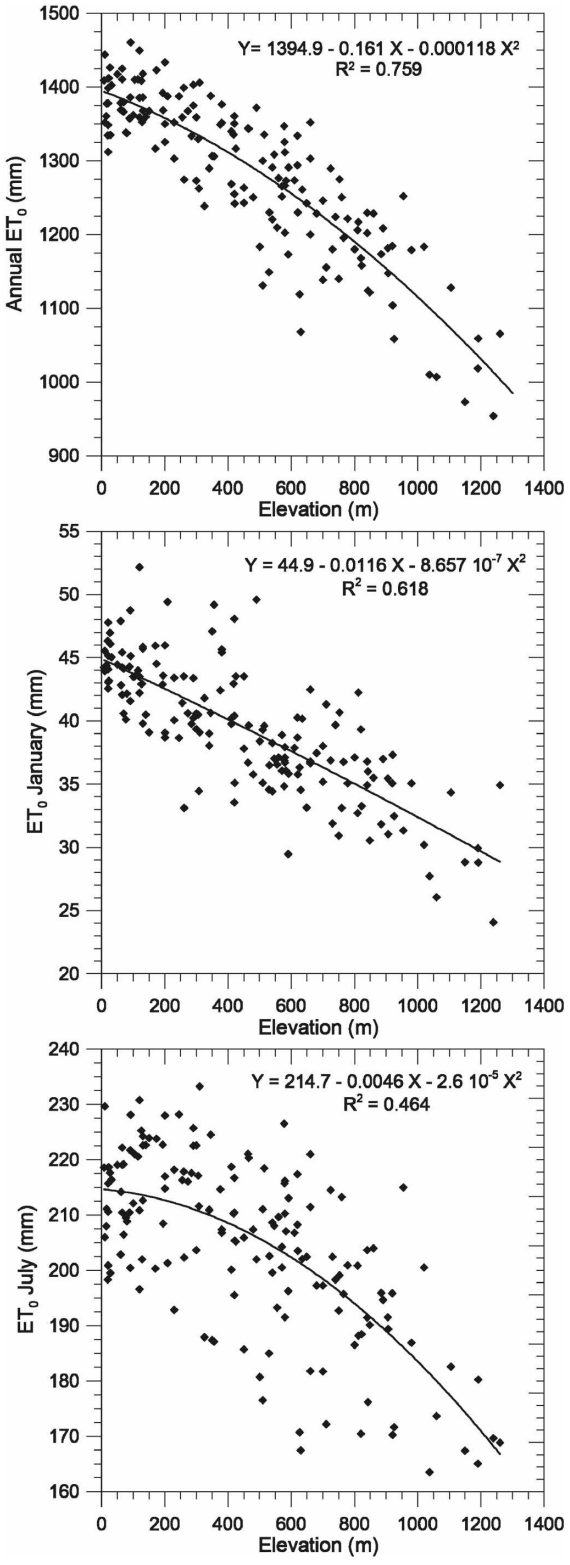

FIG. 3. Relationship between the annual and monthly (January and July) ET_0 and elevation for the inland observatories.

for the first and 144 km for the second structure. The respective contributions to the total variance (sill) were 3760 (38%) and 6050 mm^2 (62%). This means that on average approximately one-third of the total variance of ET_0 occurs within the first 30 km from each observatory. Observations of ET_0 separated by more than 144 km are statistically independent, while observatories separated by smaller distances are spatially correlated according to both scales of variation. For the residual variogram of the mean annual ET_0 the contributions to both structures were 2000 and 2400 mm^2, respectively, indicating that together with a small nugget of 506 mm^2, the secondary information reduced the total variance of the annual ET_0 by 50%. Most of the explained variance occurred at large lags. In the direction of major continuity (35°), a single exponential structure without a nugget effect, with a sill of 6500 mm^2 and a range of 59 km was fit.

For most of the monthly omnidirectional and residual variograms, nested models, consisting of an exponential and spherical structure, were used. The first structure of the omnidirectional variogram models had a range between 22 (August) and 38 km (May), and the second structure ranged between 133 (October) and 225 km (July). The residual variogram models consisted of a first structure with a range between 14 (December) and 35 km (September), and a second structure with a range between 194 (August) and 260 km (January). Only for the directional variograms (15°–34°) was a single structure sufficient, with a range that varied between 45 (September) and 96 km (March).

The results of this variogram analysis are not only useful for the geostatistical interpolation of ET_0, but they are also highly relevant for the further spatial optimization of a recently installed automatic ET_0 observation network in Andalusia (IFAPA 2007) that focuses on irrigated areas. Network design has been intensively studied in relation to accurate areal rainfall estimation (Bras and Rodríguez-Iturbe 1976; Lebel et al. 1987; Pardo-Igúzquiza 1998), but no direct applications to ET_0 were found. This is probably because rainfall is spatially more irregular than the other variables measured at meteorological observatories. More recently, de Gruijter et al. (2006) provided a wide range of methods and criteria for designing such monitoring networks, where the variogram of the variable under study plays a central role. Using the variogram or other covariance-based measures and geostatistical estimation or simulation tools, an optimal number of observatories and their location can be determined. This optimal configuration is generally found by minimizing an objective function that depends on the estimation accuracy and the total cost of the network installation,

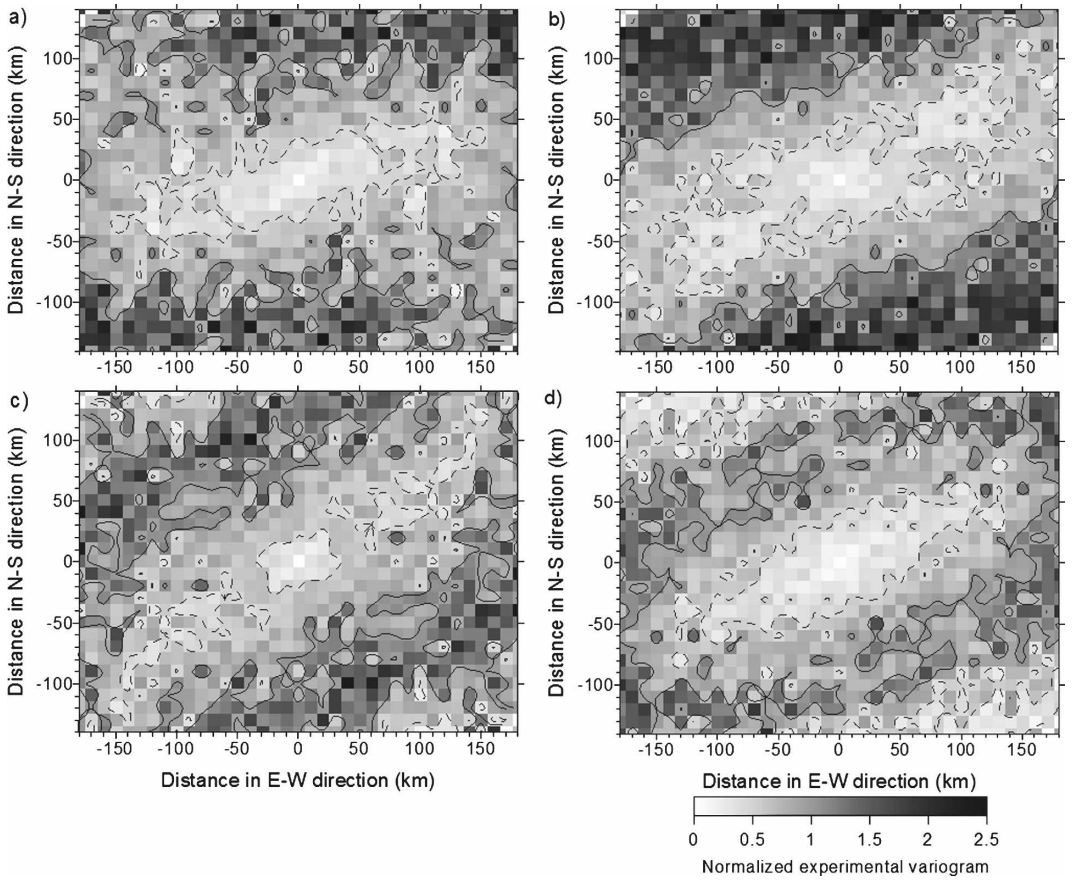

FIG. 4. Normalized experimental variogram maps for (a) ET_0 in January, (b) ET_0 in July, (c) annual ET_0, and (d) elevation. Contour lines correspond to values of 0.5 (dashed) and 1 (continuous).

exploitation, and maintenance. In most cases, at the time of designing the spatial configuration of a network, only limited information is available on the targeted attributes. However, in this case the information contained in the variograms of ET_0 can be used to decide on the optimal number and position of the observatories used for ET_0 calculation.

d. Validation of the geostatistical models

Table 5 shows that for the mean annual ET_0 the descriptive statistics of the predictions by SKlm were closest to the observed values, except with regard to the median and quartiles, for which the KED estimates were better. SKlm best reproduced the dispersion and range of the observations, although it was not capable of estimating accurately the extremes of the distribution. With respect to validation parameters, SKlm was also the best estimator, although the difference with KED was small. As expected, OK performed worse since elevation values were not used for the computations. The RMSE for OK, KED, and SKlm were 5.2%, 4.2%, and 4.1% of the mean annual ET_0, respectively. The incorporation of elevation into the spatial interpolation of mean annual ET_0 increased the model efficiency (CE) from 0.55 (OK) to 0.71 and 0.72 for KED and SKlm, respectively. It can be concluded from these results that SKlm is the best method for interpolating the mean annual ET_0 of the methods considered, while it provides an easy and straightforward way of incorporating secondary information into the interpolation. Also, for the average monthly ET_0 estimates, SKlm was superior. The cross-validation parameters in Fig. 6 showed that, especially during the spring months, SKlm performed better than KED. The RMSE for SKlm ranged from 2.3 mm (6.0% of the mean monthly ET_0) in December to 9.5 mm (4.7% of the mean monthly ET_0) in July. This means that larger relative precision can be achieved during the months with the largest ET_0, although the corresponding errors were larger. The fluctuations of the CE throughout the year ranged for SKlm from 0.84 in December to 0.65 in October. Also for this parameter SKlm performed better.

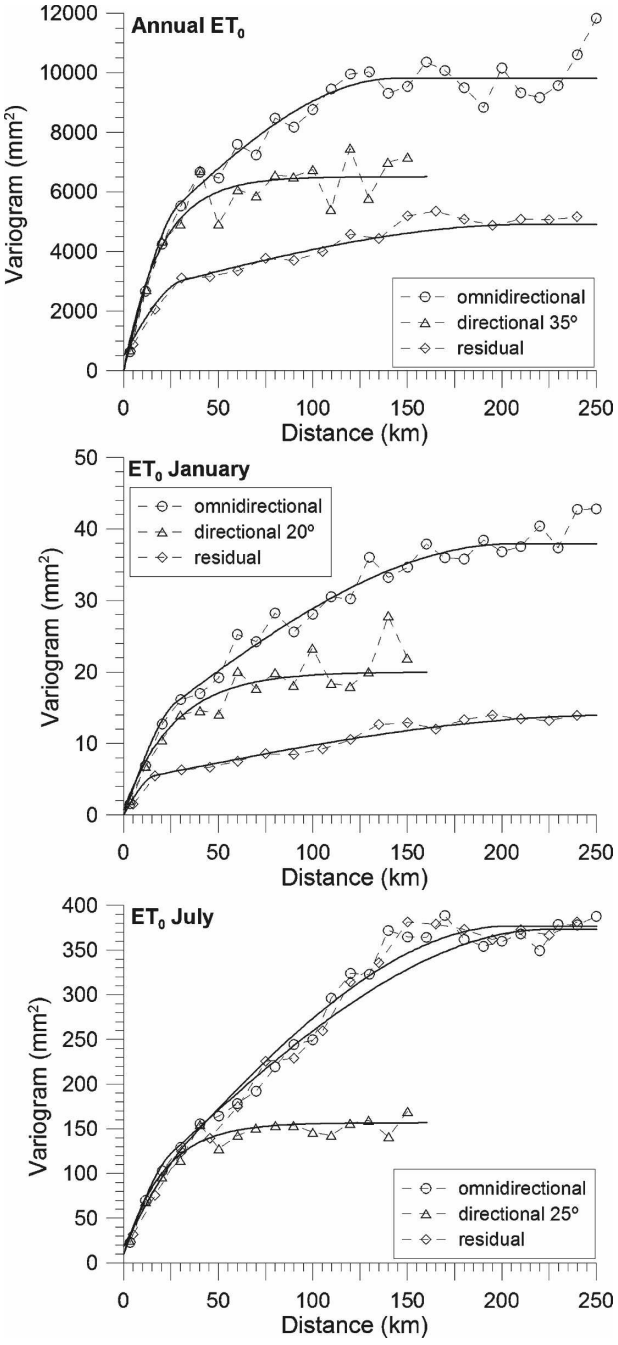

FIG. 5. Omnidirectional, directional, and residual experimental variograms for annual and monthly (January and July) ET_0, with fitted models.

TABLE 5. Descriptive statistics for observed and estimated mean annual ET_0, using OK, KED, and SKlm, and cross-validation parameters. In the table, m = mean, s = standard deviation, Q25% = lower quartile, med = median, and Q75% = upper quartile.

		Estimates		
n = 191 (mm)	Observations	OK	KED	SKlm
m	1283.0	1285.4	1284.9	1284.4
s	99.3	76.3	87.2	92.3
Min	953.9	1046.2	1029.5	936.6
Q25%	1228.5	1237.9	1230.0	1239.7
Med	1294.0	1288.0	1296.6	1299.4
Q75%	1359.0	1340.0	1357.5	1354.9
Max	1460.4	1427.5	1431.6	1434.2
RMSE	—	66.2	53.5	52.9
CE	—	0.55	0.71	0.72

e. ET_0 maps

In accordance with expectations, OK produced the smoothest maps of the mean annual ET_0 (Fig. 7). The Guadalquivir Valley especially stood out as a homogeneous area with high ET_0 values. The estimations obtained with KED and SKlm strongly reflect elevation, with the lowest ET_0 values occurring in the mountainous areas in east Andalusia. Since no data were available in these areas (above 1260 m), some care must be taken in interpreting ET_0 predictions, since these were mainly obtained through extrapolations. Since KED and SKlm use a different way of incorporating secondary information these extrapolations produce drastically different results.

The monthly maps showed how spatial ET_0 distribution changed throughout the year. Only the mean monthly ET_0 maps for January and July are represented in Fig. 8. The lowest ET_0 values were observed

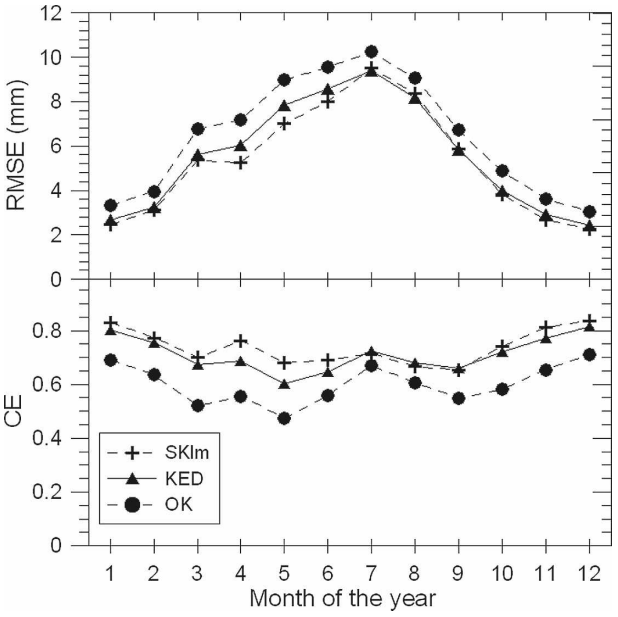

FIG. 6. Evolution of monthly RMSE and CE throughout the year, for SKlm, KED, and OK.

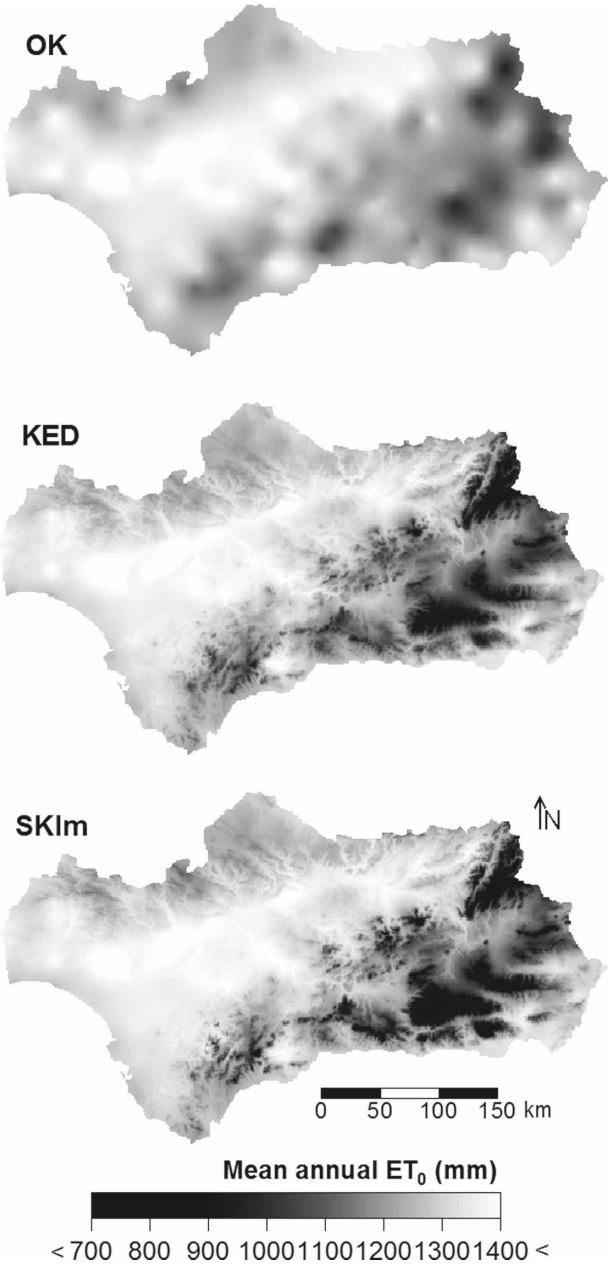

FIG. 7. Mean annual ET_0 maps produced with OK, KED, and SKlm.

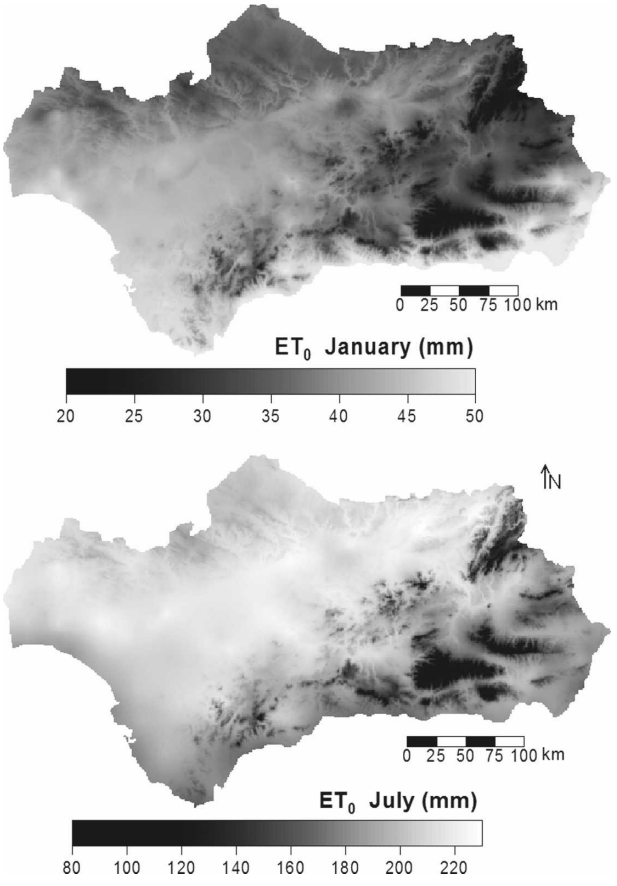

FIG. 8. Mean monthly ET_0 maps for January and December produced with SKlm.

for January and they gradually increased during the year until July, when they reached a maximum value. The maximum values for January, about 50 mm, occurred in the coastal areas, especially on the Mediterranean coasts. In the highest areas, values of less than 30 mm were observed. During February and November this spatial pattern persisted, and in addition, the values in the lower Guadalquivir Valley increased slightly. The mountainous areas became progressively separated from the rest. The area with high ET_0 values started extending throughout the Guadalquivir Valley and the coast during October, March, and April. During the remaining months, especially from June to August, the ET_0 decreased in the coastal areas and increased in the inland areas, farthest away from the coast or separated from the coast by mountain ranges. The highest values, of over 220 mm, were observed in July, in the middle and upper Guadalquivir Valley, and in the north of the province of Huelva. For these months, the mountainous areas were clearly differentiated, with ET_0 values below one-half of the maximum.

These maps were obtained from point estimations of ET_0 at the meteorological observatories [Eqs. (1) and (2)] and not from observed values. To evaluate the overall accuracy of the maps, the uncertainty of these calculations should be taken into account. Vanderlinden et al. (2004) evaluated Eqs. (1) and (2) daily estimations and obtained a mean error of 0.15%, a mean absolute error of 13.2%, and an RMSE of 17.7%. However, the accumulated error throughout the hydrological year was small. In the Almería airport observatory (southeast Andalusia), representing the coastal areas,

an accumulated error of 1% (underestimation) was observed for the mean annual ET_0 calculated with Penman–Monteith, and in the Seville airport observatory (southwest Andalusia) an accumulated error of 5% was observed (overestimation). Jensen et al. (1990) found an RMSE of 0.88 mm day^{-1} for monthly estimations of ET_0 with the original Hargreaves (1994) method. This shows that in general, the ET_0 calculation error will be larger than the spatial interpolation error.

4. Conclusions

Although a sparse automated agrometeorological network for accurate ET_0 estimation became operative at the beginning of this decade, long-term analysis, planning, and management of water resources dynamics throughout Andalusia still requires ET_0 estimations exclusively from temperature data. The Hargreaves method, improved by regional calibration, estimates the ET_0 from the maximum and minimum temperatures and is most suitable for achieving this objective, taking into account the availability of long-term temperature data in the region. Geostatistical methods taking into account elevation and distance to the coast as exhaustive secondary information (KED, SKlm) were compared with OK for interpolating ET_0.

Exploratory data analysis showed that during winter, higher ET_0 values were observed in the coastal areas, while in the summer the highest values were observed at inland stations. The ET_0 correlated well with elevation (<1260 m) after eliminating coastal observatories.

Variogram maps showed that the direction with the greatest spatial continuity of ET_0 corresponds to the orientation of the Guadalquivir Valley, one of the most geographically and climatologically uniform areas in Andalusia.

The cross-validation results for annual ET_0 indicated that the estimation improved significantly when elevation was incorporated. The difference between KED and SKlm was, however, very small but in favor of SKlm within the data limits.

A clear evolution of the spatial pattern of the mean monthly ET_0 throughout the year was observed. During the winter higher monthly values were predicted in the coastal areas and lower values predicted in the foothills. From February onward, ET_0 increased especially in the lower Guadalquivir Valley, while in the middle and upper valley this increase was observed only from April onward. In May, the ET_0 started to decrease in the coastal areas compared to the rest of Andalusia. For the subsequent months, the highest values were located in the middle and upper Guadalquivir Valley, far away from the influence of the sea or protected by mountains. From September onward, the area of high ET_0 values shifted again toward the coast.

The results of this spatial analysis of ET_0 in Andalusia are important for regional water management and for improving water use efficiency in irrigated agriculture through improved scheduling. In addition, the provided information on the spatial correlation structure of ET_0 can be essential for the further spatial optimization of the recently installed automated ET_0 observation network in the region.

REFERENCES

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith, 1998: Crop evapotranspiration: Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, Food and Agriculture Organization of the United Nations, 370 pp.

Barry, R. G., 1981: *Mountain Weather and Climate*. Methuen, 313 pp.

Bras, R. L., and I. Rodríguez-Iturbe, 1976: Network design for the estimation of areal mean of rainfall events. *Water Resour. Res.,* **12,** 1185–1195.

De Gruijter, J., D. Brus, M. Bierkens, and M. Knotters, 2006: *Sampling for Natural Resource Monitoring*. Springer, 332 pp.

Deutsch, C. V., and A. G. Journel, 1998: *GSLIB: Geostatistical Software Library and User's Guide.* 2nd ed. Oxford University Press, 369 pp.

Doorenbos, J., 1976: Agro-meteorological field stations. FAO Irrigation and Drainage Paper 27, Food and Agriculture Organization of the United Nations, 94 pp.

Goovaerts, P., 1997: *Geostatistics for Natural Resources Evaluation*. Oxford University Press, 483 pp.

——, 1999: Using elevation to aid the geostatistical mapping of rainfall erosivity. *Catena,* **34,** 227–242.

——, 2000: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.,* **228,** 113–129.

Hargreaves, G. H., 1994: Defining and using reference evapotranspiration. *J. Irrig. Drain. Eng.,* **120,** 1132–1139.

——, and R. G. Allen, 2003: History and evaluation of Hargreaves evapotranspiration equation. *J. Irrig. Drain. Eng.,* **129,** 53–63.

Hargreaves, G. L., G. H. Hargreaves, and J. P. Riley, 1985: Irrigation water requirements for Senegal River basin. *J. Irrig. Drain. Eng.,* **111,** 265–275.

IFAPA, cited 2007: Red de Información Agroclimática. [Available online at http://juntadeandalucia.es/innovacioncienciayempresa/ifapa/ria/servlet/FrontController.]

Jensen, M. E., R. D. Burman, and R. G. Allen, Eds., 1990: *Evapotranspiration and Irrigation Water Requirements: A Manual.* American Society of Civil Engineers, 332 pp.

Kim, J., Q. Guo, D. D. Baldocchi, M. Y. Leclerc, L. Xu, and H. P. Schmid, 2006: Upscaling fluxes from tower to landscape: Overlaying flux footprints on high-resolution (IKONOS) images of vegetation cover. *Agric. For. Meteor.,* **136,** 132–146.

Lebel, T., G. Bastin, C. Obled, and J. D. Creutin, 1987: On the accuracy of areal rainfall estimation: A case study. *Water Resour. Res.,* **23,** 2123–2134.

Legates, D. R., and G. J. McCabe Jr., 1999: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.,* **35,** 233–242.

Maheras, P., 1989: Principal component analysis of western Mediterranean air temperature variations 1866–1985. *Theor. Appl. Climatol.,* **39,** 137–145.

Martínez-Cob, A., 1996: Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. *J. Hydrol.,* **174,** 19–35.

Matheron, G., 1965: *La Théorie des Variables Régionalisées et Ses Applications* (*Regionalized Variable Theory and Its Applications*). Masson, 305 pp.

Nash, J. E., and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models. Part I: A discussion of principles. *J. Hydrol.,* **10,** 282–290.

Nieto, S., M. D. Frías, and C. Rodríguez-Puebla, 2004: Assessing two different climatic models and the NCEP–NCAR reanalysis data for the description of winter precipitation in the Iberian Peninsula. *Int. J. Climatol.,* **24,** 361–376.

Panatier, Y., 1996: *VARIOWIN: Software for Spatial Data Analysis in 2D*. Springer-Verlag, 91 pp.

Pardo-Igúzquiza, E., 1998: Optimal selection of number and location of rainfall gauges for areal rainfall estimation using geostatistics and simulated annealing. *J. Hydrol.,* **210,** 206–220.

Penman, H. L., 1948: Natural evaporation from open water, bare soil, and grass. *Proc. Roy. Soc. London,* **193A,** 120–148.

Phillips, D. L., J. Dolph, and D. Marks, 1992: A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. *Agric. For. Meteor.,* **58,** 119–141.

Schmid, H. P., 2002: Footprint modeling for vegetation atmospheric exchange studies: A review and perspective. *Agric. For. Meteor.,* **113,** 159–183.

Serrano, A., J. A. García, V. L. Mateos, M. L. Cancillo, and J. Garrido, 1999: Monthly modes of variation of precipitation over the Iberian Peninsula. *J. Climate,* **12,** 2894–2919.

Shuttleworth, W. J., 1993: Evaporation. *Handbook of Hydrology,* D. R. Maidment, Ed., McGraw-Hill, 4.1–4.53.

Stein, A., I. G. Staritsky, J. Bouma, A. C. van Eijnsbergen, and A. K. Bregt, 1991: Simulation of moisture deficits and areal interpolation by universal cokriging. *Water Resour. Res.,* **27,** 1963–1973.

Thornthwaite, C. W., 1948: An approach toward a rational classification of climate. *Geogr. Rev.,* **38,** 55–94.

Vanderlinden, K., 2002: Análisis de procesos hidrológicos a diferentes escalas espacio-temporales (Analysis of hydrological processes at different spatio-temporal scales). Ph.D. dissertation, University of Córdoba, 303 pp.

——, J. V. Giráldez, and M. Van Meirvenne, 2004: Assessing reference evapotranspiration by the Hargreaves method in southern Spain. *J. Irrig. Drain. Eng.,* **130,** 184–191.

——, ——, and ——, 2005: Soil water-holding capacity assessment in terms of the average annual water balance in southern Spain. *Vadose Zone J.,* **4,** 317–328.

Van Meirvenne, M., K. Scheldeman, G. Baert, and G. Hofman, 1994: Quantification of soil textural fractions of Bas-Zaire using soil map polygons and/or point observations. *Geoderma,* **62,** 69–82.

Viau, A. A., J. V. Vogt, and F. Paquet, 2000: Évaluation de la représentativité spatiale thermique des stations météorologiques du réseau d'Andalousie (Evaluation of the thermic spatial representativity of the meteorological stations in the region of Andalusia). *Int. J. Remote Sens.,* **21,** 3083–3113.