

# Spatial uncertainty assessment in modelling reference evapotranspiration at regional scale

G. Buttafuoco<sup>1</sup>, T. Caloiero<sup>2</sup>, and R. Coscarelli<sup>2</sup>

<sup>1</sup>Institute for Agricultural and Forest Systems in the Mediterranean, National Research Council, Rende (CS), Italy

<sup>2</sup>Research Institute for Geo-Hydrological Protection, National Research Council, Rende (CS), Italy

Received: 1 July 2010 – Published in Hydrol. Earth Syst. Sci. Discuss.: 14 July 2010

Revised: 27 September 2010 – Accepted: 5 October 2010 – Published: 19 November 2010

**Abstract.** Evapotranspiration is one of the major components of the water balance and has been identified as a key factor in hydrological modelling. For this reason, several methods have been developed to calculate the reference evapotranspiration ( $ET_0$ ). In modelling reference evapotranspiration it is inevitable that both model and data input will present some uncertainty. Whatever model is used, the errors in the input will propagate towards the output of the calculated  $ET_0$ . Neglecting the information about estimation uncertainty, however, may lead to improper decision-making and water resources management. One geostatistical approach to spatial analysis is stochastic simulation, which draws alternative and equally probable, realizations of a regionalized variable. Differences between the realizations provide a measure of spatial uncertainty and allows to carry out an error propagation analysis.

The aim of this paper is to assess spatial uncertainty of a monthly reference evapotranspiration model resulting from the uncertainties in the input attributes (mainly temperature) at a regional scale. A case study was presented for the Calabria region (southern Italy). Temperature data were jointly simulated by a conditional turning bands simulation with elevation as external drift and 500 realizations were generated. Among the evapotranspiration models, the Hargreaves-Samani model was used.

The  $ET_0$  was then estimated for each set of the 500 realizations of the input variables, and the ensemble of the model outputs was used to infer the reference evapotranspiration probability distribution function. This approach allowed for the delineation of the areas characterised by greater uncertainty, to improve supplementary sampling strategies and  $ET_0$  value predictions.

## 1 Introduction

Reference evapotranspiration ( $ET_0$ ), defined as the evapotranspiration of a hypothetical surface of green grass of uniform height, actively growing and adequately watered, is one of the most important hydrological variables for scheduling irrigation systems, preparing input data for hydrological water-balance models and calculating actual evapotranspiration for a region and/or a basin (Blaney and Criddle, 1950; Dyck, 1983; Hobbins et al., 2001a,b; Xu and Li, 2003; Xu and Singh, 2005; Gong et al., 2006).

The concept of reference evapotranspiration was introduced to study the evaporative demand of the atmosphere independently of crop type, crop development and management practices. As water is abundantly available at the reference evapotranspiring surface, soil factors do not affect  $ET_0$ . The only factors affecting  $ET_0$  are climatic attributes. Consequently,  $ET_0$  is a climatic attribute and can be computed from weather data.  $ET_0$  expresses the evaporating power of the atmosphere at a specific location and time of the year and does not consider the crop characteristics and soil factors (Allen et al., 1998).

A multitude of methods exists to estimate reference evapotranspiration,  $ET_0$ , (e.g., Xu and Singh, 2002). The techniques for estimating  $ET_0$  are based on one or more atmospheric variables, such as air temperature, solar or net total radiation and humidity, or some measurement related to these variables, like pan evaporation ( $ET_{pan}$ ). Some of these methods are accurate and reliable; others provide only a rough approximation. Most of the methods were developed in specific studies and work better when applied to the climate for which they were developed (Penman, 1948; Jensen, 1973).

The Penman-Monteith equation is widely recommended because of its detailed theoretical base and its accommodation of small time periods. However, the detailed climatological data required by the Penman-Monteith are not often



Correspondence to: G. Buttafuoco  
(g.buttafuoco@isafom.cnr.it)

available especially in developing nations. This lack of meteorological data was resolved by Hargreaves et al. (1985) who devised an easy approach for calculating  $ET_0$ . The Hargreaves equation (Hargreaves and Samani, 1985) requires only daily mean, maximum, and minimum air temperature, usually available at most weather stations world-wide, and extraterrestrial radiation (Droogers and Allen, 2002). This method behaves best for weekly or longer predictions, although some accurate  $ET_0$  daily estimations have been reported in literature (Hargreaves and Allen, 2003).

To improve the prediction capacity of  $ET_0$  models for large areas, spatial data should be used as inputs because their continuous variation may reflect more appropriately the nature of the  $ET_0$  in comparison with the measurements made only at a few weather station locations. When the input data are sparse or poorly correlated in space, their direct measurements could be supplemented by secondary information originating from other related attributes (Goovaerts, 1997). The estimation generally improves when additional and denser information is taken into consideration.

Evapotranspiration depends mainly on temperature, which in turn is strongly controlled by topographic attributes, longitude, latitude and distance from the coast. Temperature varies both in space and time, and it is generally well correlated with elevation. Elevation can be regarded as completely known from accurate digital elevation models and can help to map temperature above the ground. A good example in which elevation was used as external drift to model temperature is reported in Hudson and Wackernagel (1994). Moreover, at regional scale, stationarity of temperature data cannot be assumed, so a non-stationary approach such as kriging with external drift must be used.

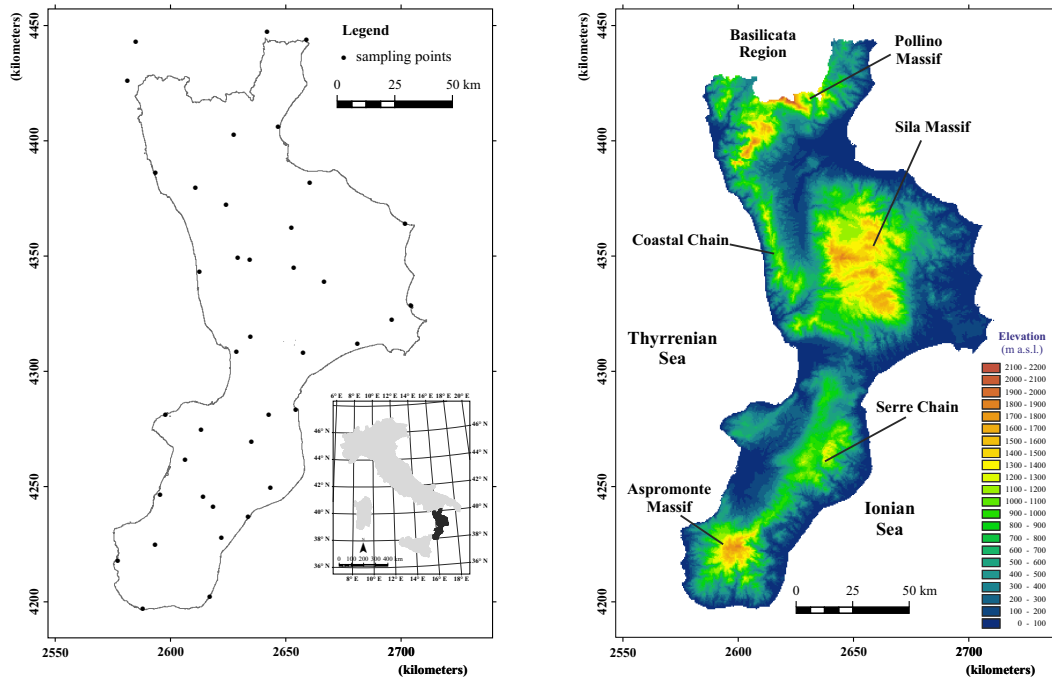
In most quantitative  $ET_0$  modelling with Geographical Information Systems (GIS), the calculation is very often assumed to produce an exact result, because most current GIS are intrinsically deterministic and cannot examine the impact of errors in input and output data. Knowing the quality of the model results is fundamental especially when they are used in spatial decision-making and water management. The quality of model predictions essentially depends on three main factors: (1) quality of data; (2) quality of the model and (3) the way data and model interact (Burrough, 2001). Therefore, it is very important to know how uncertainties in both model parameter and data propagate through the model. Model uncertainty is caused by: (1) the limitations in the mathematical models used to simulate the physical system imposed by the simplifying assumptions and/or (2) parameter errors in regression models. Model uncertainty is generally difficult to quantify and using a retrospective validation with independent data is a way of estimating the errors. Data uncertainty is instead caused by measurement errors, incomplete knowledge of spatial and temporal variations and heterogeneities at a spatial scale smaller than the sampling scale. To assess the model output error resulting from the uncertainties in the input attributes, a Monte Carlo analysis

(Heuvelink, 1998) can be used, an analysis which consists of the generation of an adequate random input data set realizations and considers the joint distribution of all input variables. The model is then run for each single set of realizations of the input variables and the ensemble of model outputs is used to infer the output probability function. A single Monte Carlo simulation consists of the model running at all locations of a fine grid covering the region of interest. The simplest way to store error surfaces of interpolated input data in a GIS is to assume that all data are normally distributed and, then, that the error is correctly expressed by the standard deviation. This method entails associating two numbers to each cell: the mean value and its standard deviation. However, a criticism to this approach is that spatial correlation is neglected and the spatially uncorrelated error is used for each realization. An alternative method consists of applying joint multivariate stochastic simulation (Gomez-Hernandez and Journel, 1992; Goovaerts, 1997), which is aimed at making predictions of cross-correlated variables. Such prediction is accomplished using a variogram matrix, which includes not only spatial autocorrelation, but also spatial cross-correlation between variables. The latter information is expected to improve the spatial prediction of  $ET_0$  by reducing its uncertainty, when compared with traditional Monte Carlo simulations. Stochastic simulation actually allows to estimate the cell-specific probability distribution functions which reflect the location of known data points and the spatial correlation structures of the variables. Although the stochastic simulation method is computer intensive, it has several advantages: unbiased predictions of model outputs; estimates of output uncertainties; assessment of error propagation in nonlinear and complex models; estimation of probability of exceeding a critical threshold.

The objective of this paper is to assess the spatial uncertainty of a monthly reference evapotranspiration model resulting from the uncertainties in the input attributes (mainly temperature data) at regional scale (Calabria region, southern Italy). In this study, it was focused on June mean reference evapotranspiration, and a subset of elevation as external drift variable at points on a 250-m square grid was used for simulation.

## 2 Materials and methods

The study case was the Calabria region located in the southern part of the Italian peninsula (Fig. 1) with an area of 15 080 km<sup>2</sup> and a coastline of 738 km on the Ionian and Tyrrhenian seas. To the North, it borders the Basilicata region for 80 km. Calabria has an oblong shape with a length of 248 km, and a width ranging between 31 and 111 km. Although Calabria does not have many high summits, it is one of the most mountainous regions in Italy (Fig. 1): 42% of the land is mountainous, 49% hilly, and only 9% is flat. The maximum elevation is 2267 m a.s.l., while the average elevation is 597 m a.s.l.



**Fig. 1.** Digital elevation model (right) of the study area and location (left) of the weather stations (pushpins).

Temperatures have been measured daily at 134 weather stations of the Italian Hydrographic Service (at present “Centro Funzionale Multirischi della Calabria” of the “Agenzia Regionale per la Protezione dell’Ambiente della Calabria” – Arpacal) during the period 1924–2009 (Fig. 1). Since it was focused on the month of June, only temperature data of this month were taken into account. Temperature time series having less than 30 years of observation were discarded and only temperature data from 42 weather stations were used. Some external stations were used to take into account the border effect. To map June temperatures, a subset of elevation at points on a 250-m square grid from an accurate digital elevation model was used in simulation as external drift variable.

The reference evapotranspiration was computed using the Hargreaves equation (Hargreaves and Samani, 1985), which can be expressed as:

$$ET_0 = 0.0023 R_a (T + 17.8) \sqrt{(T_{\max} - T_{\min})}. \quad (1)$$

where  $ET_0$  is the computed reference evapotranspiration ( $\text{mm d}^{-1}$ );  $R_a$  is the water equivalent of the extraterrestrial radiation ( $\text{mm d}^{-1}$ ) computed according to Allen et al. (1998);  $T_{\max}$ ,  $T_{\min}$  and  $T$  are the daily maximum, minimum and mean air temperature ( $^{\circ}\text{C}$ ), with  $T$  calculated as the average of  $T_{\max}$  and  $T_{\min}$ . 0.0023 and 17.8 are the original parameters proposed by Hargreaves and Samani (1985). The monthly mean reference evapotranspiration was obtained multiplying the result by 30 days.

### Stochastic simulation of the input attributes

In this paragraph, only a very brief introduction to the algorithm used in the case study will be given; for a detailed presentation, interested readers should refer to Goovaerts (1997), Chilès and Delfiner (1999), among others. Most geostatistics is based on the concept of a random function  $Z(x)$ , whereby the set of unknown values is regarded as a set of spatially dependent random variables  $Z(x_\alpha)$ . Each measurement of air temperature,  $z(x_\alpha)$ , at a different location  $x_\alpha$  ( $x$  is the location coordinates vector and  $\alpha$  the sampling points = 1, ...,  $n$ ) is interpreted as a particular realization of a random variable  $Z(x_\alpha)$ .

Geostatistical simulation, compared with an optimal procedure of estimation such as kriging, provides a more realistic means of evaluating the spatial variability of a variable (Castrignanò and Buttafuoco, 2004). Stochastic simulation results in a large number of equiprobable images, also called realizations, which honour the sample data and reproduce statistical characteristics and spatial features. Two types of simulations are available using geostatistics: unconditional and conditional. Unconditional simulations simply reproduce certain statistical measures (mean, variance, covariance function) of a variable without considering the observed data. Conditional simulations generate realizations that incorporate the correlation structure of the data, honour the data and reproduce some random component of variation, which is smoothed out by kriging.

There are several simulation techniques: the one used in this work is the turning bands method with external drift, which represents a reasonable trade-off between quality and computing time. This method was chosen because the weather stations having temperature data with more than 30 years of observations were sparse and few (only 42). Moreover, as it was mentioned before, temperature cannot be assumed stationary at a regional scale and elevation, which is an additional and denser information easily available, can improve the temperature estimation. The variation in temperature ( $Z(\mathbf{x})$ ) then comprises a deterministic and a stochastic components represented by the model:

$$Z(\mathbf{x}_\alpha) = m(\mathbf{x}) + \varepsilon(\mathbf{x}) \text{ and } E[Z(\mathbf{x})] = m(\mathbf{x}). \quad (2)$$

where  $\varepsilon(\mathbf{x})$  is the stochastic component with zero mean and variogram  $\gamma_\varepsilon(\mathbf{h})$  and  $m(\mathbf{x})$  is the drift which in Kriging with an External Drift (KED) is usually modelled as a linear function of a smoothly varying secondary (external) variable  $y(\mathbf{x})$  (elevation in this case):

$$m(\mathbf{x}) = a_0(\mathbf{x}) + a_1(\mathbf{x})y(\mathbf{x}). \quad (3)$$

where the two coefficients  $a_0(\mathbf{x})$  and  $a_1(\mathbf{x})$  are deemed constant within the search neighbourhood and are implicitly estimated through the kriging system (Goovaerts, 1997). The principle is to simulate a target variable using an auxiliary linear correlated variable known at the grid nodes of the result grid file. The auxiliary correlated variable in this case study was elevation because there is a good linear correlation between both June mean minimum ( $-0.90$ ) and June mean maximum ( $-0.87$ ) temperature data. The value of the secondary variable (elevation) must be known at all primary data locations  $\mathbf{x}_\alpha$  ( $\alpha = 1, \dots, n$ ) and at all locations  $\mathbf{x}_0$  being estimated. Moreover, the secondary variable should vary smoothly in space to avoid instability of the kriging with external drift system (Goovaerts, 1997). The simulation algorithm generates a 2D simulation from the 1-D simulations along the lines. The turning bands method is a powerful and useful mathematical operator (Christakos, 1987, 1992), however some authors (Deutsch and Journel, 1998) have criticised it because of the generation of artefacts in the simulated images. These artefacts are for the 3-D cases due to the limitation of a maximum of 15 lines which provide a regular partition of the 3-D space but there is no such limitation in 2-D. The turning band method consists of adding up a large number of independent 1-D simulations on the lines partitioning the plane. The value of the simulation at a point of the plane is the sum of the simulated values of the projected points on the different lines. In practice, the 2-D covariance is given and the 1-D covariance is obtained by a deconvolution process. The simulations along each line are discretized so that the same simulated value at a point is assigned to the “band” perpendicular to the line and containing the point. Hence, the name turning “bands” given to the method. The only parameter of this method is the count of bands that has

been fixed to 400 in this work. This was a good compromise to save computer time and obtain good results. Moreover, the number of realizations was fixed to 500 because high accuracies are reached only when the number of runs is sufficiently large.

The previous steps generate non-conditional realizations, which reproduce the given covariance function but do not honour the data. Conditioning is implemented in the software ISATIS<sup>®</sup> release 10.03 ([www.geovariances.com](http://www.geovariances.com)) by kriging (Journel and Huijbregts, 1978; Chiles and Delfiner, 1999). The conditional simulation at location  $\mathbf{x}_0$  is given by:

$$Z_{cs}(\mathbf{x}_0) = Z_s(\mathbf{x}_0) + \sum_{i=1}^n \lambda_i^0 [z(\mathbf{x}_i) - Z_s(\mathbf{x}_i)]. \quad (4)$$

where  $Z_{cs}(\mathbf{x}_0)$  is the conditional simulation at  $\mathbf{x}_0$ ;  $Z_s(\mathbf{x}_0)$  the non-conditional simulation at  $\mathbf{x}_0$ ;  $z(\mathbf{x}_i)$  the experimental value at experimental location  $\mathbf{x}_i$ ;  $Z_s(\mathbf{x}_i)$  the non-conditional simulation at experimental locations  $\mathbf{x}_i$ ;  $\lambda_i^0$  the kriging weight assigned at experimental location  $\mathbf{x}_i$  when estimating at location  $\mathbf{x}_0$ ; and  $n$  the number of experimental locations for kriging. As the turning bands method is a Gaussian simulation technique, it requires a multi-Gaussian framework. Therefore, each variable has initially been transformed into a normal distribution with zero mean and unit variance, and the simulation results have subsequently been back-transformed to the raw distribution. Such a procedure is known as Gaussian anamorphosis (Chilès and Delfiner, 1999; Wackernagel, 2003), and it is a mathematical function, which transforms a variable with a Gaussian distribution into a new variable with any distribution. The Gaussian anamorphosis can be achieved using an expansion into Hermite polynomials  $H_i(Y)$  (Wackernagel, 2003) restricted to a finite number of terms.

The joint turning bands simulation requires modelling and fitting a linear model of coregionalization (Goovaerts, 1997). The Linear Model of Coregionalization (LMC) is a quantitative measure of spatial correlation of the regionalized variable  $z(\mathbf{x}_i)$ . It provides a method for modelling the direct and cross-variogram(s) of two or more variables so that the variance of any possible linear combination of these variables is positive. Any experimental variogram is modelled as a combination of the same basic structures.

The aim was to build a model, which described the major spatial features of the attributes under study. The models used can represent bounded or unbounded variation. In the former models, the variance (known as the sill variance) has a maximum at a finite lag distance (range) over which pairs of values are spatially correlated. In this case, to model the coregionalization of the two attributes ( $T_{\max}$  and  $T_{\min}$ ), three ( $N(N+1)/2$ ) direct and cross variograms must be calculated and modelled jointly for the anamorphosed temperature data.

The best fitting function can be chosen by cross-validation, which checks the compatibility between the data and the model. It takes each data point in turn, removing it

**Table 1.** Basic statistics of mean maximum ( $T_{\max}$ ) and mean minimum temperature ( $T_{\min}$ ) data (Celsius degrees) for June and elevation data (m a.s.l.).

Statistics	$T_{\max}$	$T_{\min}$	Elevation
Mean	26.9	15.8	366.4
Stand. Dev.	2.5	3.0	386.5
Maximum	30.9	19.6	1315.0
Upper quartile	28.7	18.0	516.0
Median	27.9	17.1	242.0
Lower quartile	26.0	15.2	15.0
Minimum	20.1	7.7	2.0
Skewness	-1.0	-1.4	1.1
Kurtosis	3.3	4.2	3.4

temporarily from the dataset and using its neighbouring information to predict the value of the variable at its location. The estimate is compared with the measured value by calculating the experimental error, i.e., the difference between estimate and measurement, which can also be standardized by estimating the standard deviation. The goodness of fit was evaluated by the mean error (ME) and the mean squared deviation (MSDR). The mean error, which proves the unbiasedness of the estimate if its value is close to 0, is given by:

$$ME = \frac{1}{n} \sum_{i=1}^n [Z^*(\mathbf{x}_i) - z(\mathbf{x}_i)]. \tag{5}$$

where  $n$  is the number of observation points,  $Z^*(\mathbf{x}_i)$  is the predicted value at location  $i$ , and  $z(\mathbf{x}_i)$  is the observed value at location  $i$ . The mean squared deviation ratio MSDR, which is the ratio between the squared errors and the kriging variance ( $\sigma^2(\mathbf{x}_i)$ ) is expressed as:

$$MSDR = \frac{1}{n} \sum_{i=1}^n \frac{[Z^*(\mathbf{x}_i) - z(\mathbf{x}_i)]^2}{\sigma^2(\mathbf{x}_i)}. \tag{6}$$

If the model for the variogram is accurate, the mean squared error should equal the kriging variance and the MSDR value should be 1.

Then each geostatistical multivariate simulation has been used as an input to the  $ET_0$  model. Probabilistic information has been extracted from the set of simulated images. By averaging the simulated values at each cell, two different maps have been produced: the map of the expected value at any given location (E-type or Expected-value estimate; Journel, 1983) and the one of its standard deviation. The uncertainty in model predictions has been quantitatively evaluated from the replicate stochastic images.

In sum, the proposed approach consisted in the following steps:

1. generating a set of input attributes ( $T_{\min}$  and  $T_{\max}$ ) realizations  $a_i$  ( $i = 1, \dots, 500$ ) at nodes of a 250-m square

**Table 2.** Fitted linear model of coregionalization of anamorphosed  $T_{\max}$  ( $GT_{\max}$ ) and  $T_{\min}$  ( $GT_{\min}$ ). The coregionalization matrices, the eigenvalues, and the corresponding percentage of variance explained by each eigenvector for the three basic structures are reported.

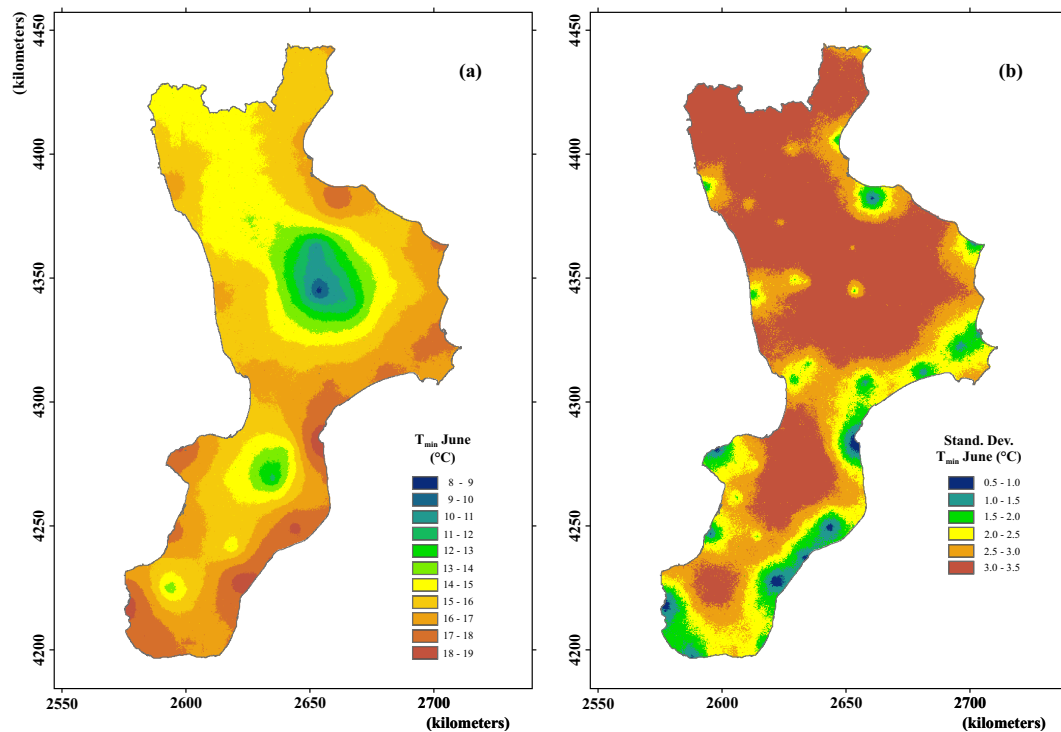
Variable	$GT_{\max}$	$GT_{\min}$
(1) Nugget effect		
$GT_{\max}$	0.5613	
$GT_{\min}$	0.2752	0.5784
Eigenvalue:	0.8452	0.2945
	(74.16%)	(25.84%)
(2) Exponential model (Pr. Range = 30 000 m)		
$GT_{\max}$	0.3108	
$GT_{\min}$	0.3589	0.5168
Eigenvalue:	0.7872	0.0403
	(95.13%)	(4.87%)
(3) Spherical model (Range = 60 000 m)		
$GT_{\max}$	0.2204	
$GT_{\min}$	0.0037	0.0001
Eigenvalue:	0.2205	0.0000
	(100.00%)	(0%)

grid using the joint stochastic simulation with elevation as external drift;

2. for this set of input realizations  $a_i$ , computing the reference evapotranspiration using Eq. (1);
3. for each input and output attributes, computing average and standard deviation of the simulated values at each cell to produce the maps of the expected values at any given location and the ones of their standard deviation.

### 3 Results and discussion

The summary statistics of mean maximum ( $T_{\max}$ ) and mean minimum temperature ( $T_{\min}$ ) data (Celsius degrees) for June and elevation data (m a.s.l.), measured at the weather stations, are reported in Table 1. The assumption of normal distribution was not accepted for both mean maximum ( $T_{\max}$ ) and mean minimum temperature ( $T_{\min}$ ) data at a probability level  $p > 0.10$ , and the data distributions showed long negative tails. Therefore, before conducting joint Gaussian simulation, we applied a Gaussian transformation to  $T_{\max}$  and  $T_{\min}$  data. No anisotropy was evident in the maps of the 2-D variograms (not shown) to a maximum lag distance of 100 km. A nested isotropic LMC (Table 2) was fitted to all the experimental direct and cross-variograms of the Gaussian transformed variables. The LMC (Table 2) includes



**Fig. 2.** Mean (a) and standard deviation (b) maps of simulations for the mean minimum temperature ( $T_{\min}$ ).

three different structures: a nugget effect, an exponential structure with a practical range of 30 000 m, and a spherical structure with a range of 60 000 m. The goodness of fit was evaluated by a cross-validation test and the results in terms of mean error (ME) and mean squared deviation (MSDR). ME ranged between  $-0.02$  and  $-0.07$ , while MSDR between 0.92 and 0.93. The optimal values for the two statistics are 0 for ME and 1 for MSDR, then the multivariate model of spatial correlation was unbiased and reproduced the experimental variance adequately. The sum of the eigenvalues at each spatial scale provides an estimate of the variance at that scale (Table 2). The nugget was about 50% of total variance (2.18), while the contribution of the shorter range component (30 000 m) of variation to the total variance was about 40% and the contribution of the longer range component (60 000 m) was 10%. The nugget effect component represents the unstructured spatial variation. It is mainly due to measurement errors and to the spatial variation at a scale lesser than the minimum distance of sampling. Moreover, in this study, the nugget ratio is due to the sparse and limited number of sampling locations. The variation at shorter scale (40%) is probably related to the local orographic characteristics of the region, while the longer scale of variation (10%) could be related to large scale factors.

The above LMC was used to generate the 500 simulations of  $T_{\min}$  and  $T_{\max}$  at nodes of a 250-m square grid. Afterwards, the expected values of  $T_{\min}$  and  $T_{\max}$  and their standard deviation were computed and mapped: Figs. 2 and 3

present a way to treat the jointly simulated images of the two variables, by calculating the mean and the standard deviation, respectively, of the 500 simulations at each grid node, and then mapping the results for each variable. The mean maps (Figs. 2a and 3a) show the complexity in spatial distribution of temperatures. The maps of the standard deviation (Figs. 2b and 3b), obtained by post-processing the simulations, have allowed to assess the uncertainties of non-Gaussian variables and to overcome the drawback of kriging variance of its independence from actual sample values. From a visual inspection, it shows clearly how the uncertainty distributions of temperature are mostly related to the density of the sample data (Fig. 1). Figures 2a and 3a also show, as expected, that lower values of mean temperatures are estimated in correspondence to the mountainous areas (Sila Massif, Serre Chain and Aspromonte Massif) (Fig. 1), which have also the higher uncertainties (higher values of standard deviation). Figures 2b and 3b also show a large area of the region (the northern part) characterised by high values of standard deviation. These high values occurred in an area with high variability in elevation and consequently in temperature data at short distance.

The 500 realizations of temperature data were used as input for  $ET_0$  model and then the maps of the mean and standard deviation of  $ET_0$  (Fig. 4) were obtained in a similar way to the maps of Fig. 2 and 3. A visual inspection of Fig. 4b shows clearly where the uncertainty in  $ET_0$  is high. The simulation has shown how the previous uncertainties in input



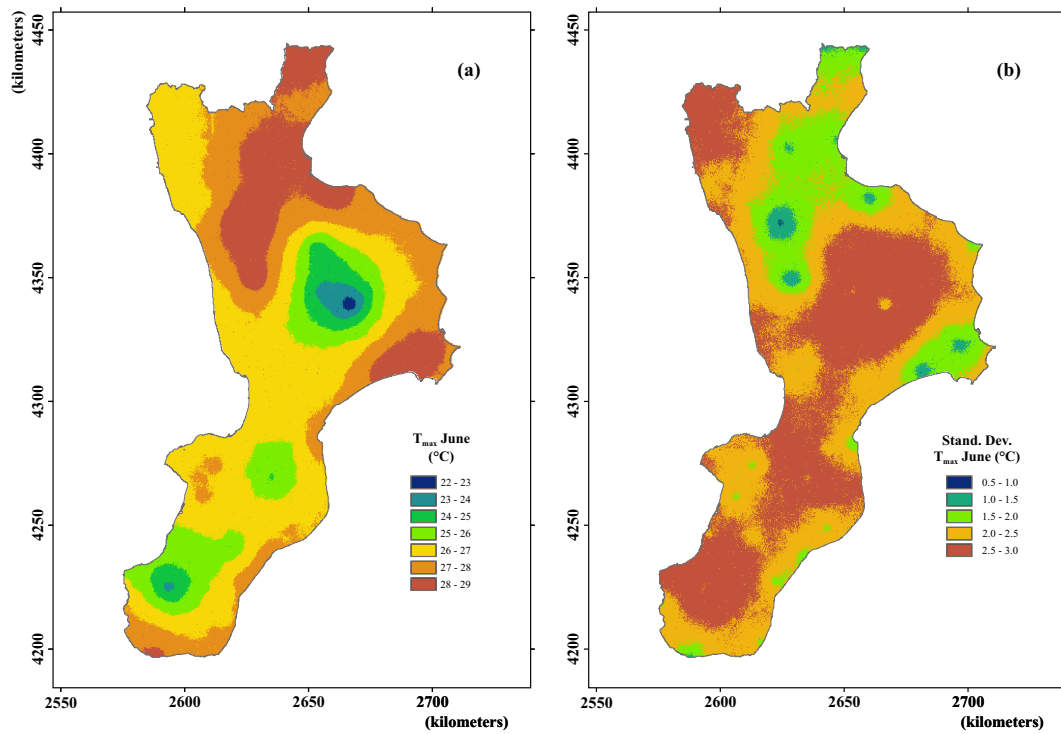


Fig. 3. Mean (a) and standard deviation (b) maps of simulations for the mean maximum temperature ( $T_{max}$ ).

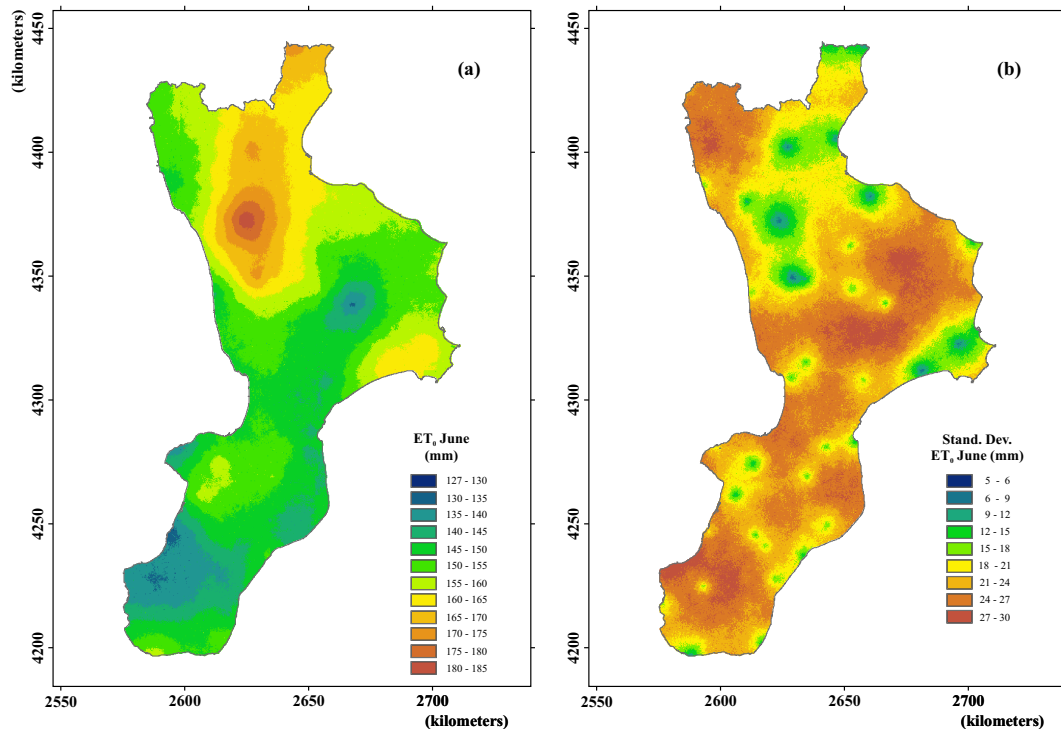


Fig. 4. Mean (a) and standard deviation (b) maps of simulations for  $ET_0$ .

variables can affect the predictions of  $ET_0$  model. In particular, one can note that there are extended areas characterised by high uncertainties localised on the Sila Massif, the Aspromonte Massif, the Serre Chain and the north-western portion of the region. The areas characterised by medium-high values of  $ET_0$  (Fig. 4a) present not very high values of standard deviation (Fig. 4b) and, therefore, less uncertainty.

This approach has demonstrated that it is possible to produce maps of uncertainty, which are more useful than the simple extrapolations of estimation points. Of course, it is important to evaluate how well the model approximates reality, i.e., model uncertainty. If the model has high uncertainty, a difference in model output may not indicate a real change and, thus, could be meaningless.

#### 4 Conclusions

The results of a spatial uncertainty analysis have shown that the prediction quality depends on the uncertainties of the data used in the analysis; therefore, map makers should convey the accuracy of the maps they produce (Heuvelink, 1998). A complete characterisation of the accuracy of spatial data should also include the spatial correlation of the attributes used for estimation and stored in a GIS. In the past, a single root mean squared error was sufficient to assess spatial accuracy, but now it is no longer sufficient, and much more information should be provided to characterise the quality of a map.

The objective of this study, however, was not to validate the model or assess the errors associated with the model type and coefficients, but rather to evaluate how the variability of inputs affects uncertainty of model prediction. By definition, a model is an approximation of reality and some models describe reality better than others. Therefore, the choice of model plays an important role in error prediction. In this paper, however, it was assumed that an appropriate model was selected and that the model errors were associated only with the spatial variation of the input attributes.

In order to obtain realistic values of the model output uncertainty, when the model outputs are supposed to be spatially correlated, it is critically important to model and assess spatial correlations of input variables (Heuvelink and Pebesma, 1999). Ignoring spatial correlation between input variables, as in the traditional Monte Carlo approach, implies modelling input variables as white noise. In this case, all uncertainty in  $ET_0$  predictions might vanish after mapping with a dense point grid. The required density depends on the estimate precision level, and it is of paramount importance to model spatial correlation correctly to separate input error from model uncertainty. The two types of variation are quite different. Spatial variation refers to the deterministic variation of  $ET_0$  or a single realization of the input attributes, whereas uncertainty refers to  $ET_0$  distribution for a single point obtained from the ensemble of Monte Carlo simula-

tions. The approach would put more emphasis on the quality of the input data and on how the input uncertainties may have a considerable impact on prediction uncertainty. Looking at the standard deviation map of  $ET_0$  (Fig. 4b), only a weak spatial pattern can be distinguished; the errors do not appear correlated with the estimates of  $ET_0$  (Fig. 4a), but with the density of weather stations.

Finally, it is worth pointing out the consequences of estimation uncertainty in the context of decision-making and water resources management. There is currently some reluctance to perform error recognition, possibly because of the greater analysis required. However, studying uncertainty leads to increased understanding of the roles played by the different input attributes, which also allows to evaluate the relative costs and benefits of using different scenarios.

*Acknowledgements.* The authors are grateful to the editor and to the anonymous reviewers for the precious remarks and comments which led to improve the initial version of this paper.

Edited by: S. Grimaldi

#### References

- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration – guidelines for computing crop water requirements, FAO Irrigation and Drainage, Paper no. 56, Rome, Italy, 1998.
- Blaney, H. F. and Criddle, W. D.: Determining water requirements in irrigated area from climatological irrigation data, US Department of Agriculture, Soil Conservation Service, Technical Paper no. 96, Washington, DC, USA, 1950.
- Burrough, P. A.: GIS and geostatistics: Essential partners for spatial analysis, *Environ. Ecol. Stat.*, 8, 361–377, 2001.
- Castrignanò, A. and Buttafuoco, G.: Geostatistical Stochastic Simulation of Soil Water Content in a Forested Area of South Italy, *Biosyst. Eng.*, 87, 257–266, 2004.
- Chilès, J. P. and Delfiner, P.: Geostatistics: Modelling spatial uncertainty, Wiley, New York, USA, 1999.
- Christakos, G.: The space transformations in the simulation of multidimensional random fields, *J. Math. Comp. Sim.*, 29, 13–319, 1987.
- Christakos, G.: Random field models in earth sciences, Academic Press, San Diego, CA, USA, 1992.
- Deutsch, C. V. and Journel, A. G.: GSLIB: Geostatistical software library and user's guide, Oxford University Press, New York, USA, 1998.
- Droogers, P. and Allen, R. G.: Estimating reference evapotranspiration under inaccurate data conditions, *Irrig. Drain.*, 16, 33–45, 2002.
- Dyck, S.: Overview on the present status of the concepts of water balance models, in: Proceeding of the Hamburg Workshop New Approaches in Water Balance Computations, Hamburg, 15–29 August 1983, IAHS Publication no. 148, 3–19, 1983.
- Gomez-Hernandez, J. J. and Journel, A. G.: Joint sequential simulation of multigaussian fields, in *Geostat Troia 1992*, Kluwer Publ. Dordrecht, The Netherlands, 1992.



- Gong, L., Xu, C. Y., Chen, D., Halldin, S., and Chen, Y. D.: Sensitivity of the Penman-Monteith reference evapotranspiration to key climatic variables in the Changjiang (Yangtze River) basin, *J. Hydrol.*, 329, 620–629, 2006.
- Goovaerts, P.: *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York, USA, 1997.
- Hargreaves, G. H. and Samani, Z. A.: Reference crop evapotranspiration from temperature, *Appl. Eng. Agric.*, 1, 96–99, 1985.
- Hargreaves, G. H. and Allen, R. G.: History and evaluation of Hargreaves evapotranspiration equation, *J. Irrig. Drain. E-ASCE*, 129, 53–63, 2003.
- Hargreaves, G. L., Hargreaves, G. H., and Riley, J. P.: Agricultural benefits for Senegal River basin, *J. Irrig. Drain. E-ASCE*, 111, 113–124, 1985.
- Heuvelink, G. B. M.: *Error propagation in environmental modeling with GIS*, Taylor, Francis, London, UK, 1998.
- Heuvelink, G. B. M. and Pebesma, E. J.: Spatial aggregation and soil process modelling, *Geoderma*, 89, 47–65, 1999.
- Hobbins, M. T., Ramirez, J. A., and Brown, T. C.: The complementary relationship in estimation of regional evapotranspiration: an enhanced advection-aridity model, *Water Resour. Res.*, 37, 1389–1403, 2001a.
- Hobbins, M. T., Ramirez, J. A., Brown, T. C., and Claessens, L. H. J. M.: The complementary relationship in estimation of regional evapotranspiration: the complementary relationship areal evapotranspiration and advection-aridity models, *Water Resour. Res.*, 37, 1367–1387, 2001b.
- Hudson, G. and Wackernagel, H.: Mapping temperature using kriging with external drift: theory and an example from Scotland, *Int. J. Climatol.*, 14, 77–91, 1994.
- Jensen, M. E.: *Consumptive use of water and irrigation water requirements*, American Society of Civil Engineers, New York, USA, 1973.
- Journel, A. G. and Huijbregts, C. J.: *Mining Geostatistics*, Academic Press, New York, USA, 1978.
- Journel, A. G.: Non-parametric estimation of spatial distributions, *Math. Geol.*, 15, 445–468, 1983.
- Penman, H. L.: Natural evaporation from open water, bare soil and grass, *P. Roy. Soc. Lond. A*, 193, 120–145, 1948.
- Wackernagel, H.: *Multivariate Geostatistics: an introduction with applications*, Springer-Verlag, Berlin, Germany, 2003.
- Xu, Z. X. and Li, J. Y.: A distributed approach for estimating catchment evapotranspiration: comparison of the combination equation and the complementary relationship approaches, *Hydrol. Process.*, 17, 1509–1523, 2003.
- Xu, C. Y. and Singh, V. P.: Cross-comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland, *Water Resour. Manag.*, 16, 197–219, 2002.
- Xu, C. Y. and Singh, V. P.: Evaluation of three complementary relationship evapotranspiration models by water balance approach to estimate actual regional evapotranspiration in different climatic regions, *J. Hydrol.*, 308, 105–121, 2005.