Author accepted version. Final publications as:
Carrera-Hernandez, J.J & Gaskin, S.J. (2007) Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico. Journal of Hydrology. 336(3-4): 231-249. doi:10.1016/j.jhydrol.2006.12.021

Spatio temporal analysis of precipitation and temperature in the Basin of Mexico

J. J. Carrera-Hernández *, S. J. Gaskin

McGill University, Department of Civil Engineering and Applied Mechanics, 817 Sherbrooke Street West, Montreal QC, H3A 2K6, Canada

Abstract

The spatial distribution of climatological variables such as rainfall and temperature is needed whenever hydrological modelling is undertaken at the watershed scale. These models can be used to simulate hydrological processes at a daily or hourly time step and the interpolation of climatological variables (in particular precipitation) at this time scale poses a particular problem due to its large spatial variation. This work analyzes the temporal variation of both minimum and maximum temperature and rainfall, its correlation with elevation and whether or not this relationship should be used when daily data are interpolated. In order to achieve this, the monthly distribution of these variables is derived from daily interpolations, which is compared to their monthly accumulated value for each climatological station. The interpolation methods used to undertake the analysis were Ordinary Kriging (OK), Kriging with External Drift (KED), Block Kriging with External Drift (BKED), Ordinary Kriging in a local neighborhood (OK_l) and Kriging with External Drift in a local neighborhood (KED_l). This analysis used daily climatological data from approximately 200 stations located in the Basin of Mexico for June 1978 and June 1985, from which accumulated monthly data were derived. The results of this analysis show that the interpolation of daily events is improved by the use of elevation as a secondary variable even when these variables show a low correlation.

Key words: Rainfall, temperature, interpolation, Kriging, Mexico City, Basin of Mexico

1 Introduction

The spatial distribution of climatological variables is needed as input in distributed hydrological models and different authors have undertaken interpolations in different geographical locations and for different time periods. Temperature has been interpolated at a daily time step (Jarvis and Stuart, 2001), while rainfall has been interpolated using averaged values ranging from daily (Kyriakidis et al., 2001), monthly (Lloyd, 2005; Hudson and Wackernagel, 1994) or annual (Hofierka et al., 2002; Goovaerts, 2000; Martinez-Cob, 1996; Phillips et al., 1992; Dingman et al., 1988; Tabios and Salas, 1985) aggregation levels.

The analysis of the spatial distribution of daily rainfall is difficult mainly because of intermittence and large variability. The use of daily climatological data is needed in (semi)arid regions when developing water balances or when studying aquifer recharge, as annual evapotranspiration may greatly exceed rainfall. Thus in order to improve the estimates of aquifer recharge in these regions, daily data should be used.

Spatial interpolation can be undertaken through the use of various algorithms, and their evaluation has been addressed by several authors among them Tabios and Salas (1985) who compared Kriging with Thiessen, Inverse Distance Weight (IDW), Polynomial trend surfaces and inverse square distance, and by Jarvis and Stuart (2001) who compared Ordinary Kriging (OK), partial thin plate splines, inverse distance weighting and trend surface analysis. These studies concluded that the Kriging method yields a more realistic spatial behaviour of the climatological variable of interest. However, Kriging comprises different interpolation methods which can be differentiated by whether or not they use an external variable and if this variable is used in a global or local neighborhood. Among the Kriging methods that do not make use of an external variable are the following: Simple Kriging (SK), Ordinary Kriging (OK), Kriging with varying local means (Klm) and Block Kriging (BK), while among the methods that make use of a secondary variable are Factorial Kriging (FK), Kriging with External Drift (KED) and Cokriging

* Corresponding author.
Email addresses: jaime.carrera@mail.mcgill.ca (J. J. Carrera-Hernández), susan.gaskin@mcgill.ca (S. J. Gaskin).

(Cok). The performance of different Kriging methods has been reported in previous works such as Lloyd (2005) who used Moving Window Regression (MWR), IDW, OK, SKlm and KED to interpolate monthly precipitation values in England for 1999 for which the use of elevation as an auxiliary variable through the application of KED provided more accurate estimates from March to December.

The use of auxiliary variables in order to improve the spatial interpolation of climatological variables has been analyzed by Jarvis and Stuart (2001). When interpolating minimum and maximum temperatures they concluded that including values recorded at nearby stations provides greater accuracy than the selection of auxiliary or guiding variables. The only study that uses daily rainfall data is the Spatial Interpolation Comparison of 1997, summarized in Dubois (1998) and which consisted of the comparison of different interpolation techniques applied to a set of 100 rainfall measurements in Switzerland on May 8th, 1986. The interpolation methods applied were Ordinary and Indicator Kriging (Atkinson and Lloyd, 1998), Inverse Distance Weighting (Tomczak, 1998), Linear and Zone Kriging (Saveliev et al., 1998), Neural Network Residual Kriging (Demyanov et al., 1998), Multiquadratic functions (Thieken, 1998) and Probability Class Kriging (Allard, 1998). In addition, Hofierka et al. (2002) applied the Regularized Spline with Tension (RST) method of Mitasova and Mitas (1993) to the SIC 97 data set, comparing their results with those obtained by the previously mentioned authors using the Root Mean Square Error (RMSE) as a benchmark. The RMSE values ranged from 5.20 (mm) to 6.14 (mm) (Hofierka et al., 2002); the lowest RMSE value was obtained by the Regularized Spline with Tension method, without using elevation as an auxiliary variable.

Unfortunately, a clear answer can not be found on whether or not the use of elevation as a secondary variable can improve the spatial interpolation of daily rainfall. According to some authors (Lloyd (2005); Goovaerts (2000)) although the use of elevation as an auxiliary variable improves the spatial interpolation of monthly rainfall data, the relationship between rainfall and elevation is less useful when interpolating daily data.

This study examines whether or not the relationship between three different climatological variables (rainfall, minimum and maximum temperature) and elevation should be used when interpolating daily climatological data using data for four different days, two days in June 1978 and two days in June 1985 thus analyzing the effect of temporal variations in the relationship between elevation and the climatological variable of interest. In addition, the accumulated monthly values for June 1978 and 1985 obtained from the daily interpolations are analyzed. Five different methods are investigated: Ordinary Kriging (OK), Kriging with External Drift (KED), Block Kriging with External Drift (BKED), Ordinary Kriging on a local neighborhood (OK_l) and Kriging with External Drift on a local neighborhood (KED_l). Kriging with External Drift was chosen among the methods that consider elevation as a secondary variable as Goovaerts (2000) found that this method provides slightly better results than cokriging while not being as computationally demanding. Although other methods such as RST (Mitasova and Mitas, 1993) can also be used, it was decided to use only Kriging methods, in order to analyze a method that can be automatically applied to large datasets (e.g. daily interpolations for long term analysis). Those such as RST need additional parameters to be "tuned" (Hofierka et al., 2002).

2 Kriging

This section presents a brief overview of the interpolation methods used. It is based on Wackernagel (2003), Deutsch and Journel (1998), Goovaerts (1997) and Isaaks and Srivastava (1989), which provide an in depth view of the Kriging method along with its variants. Kriging methods use the available data $z(\mathbf{u}_\alpha)$ at n points in a specified search neighborhood in order to determine the values at unsampled locations $z(\mathbf{u})$, where the $z(\mathbf{u})$ value is a realization of a stationary Random Function (RF) that comprises $n+1$ Random Variables (RV). The $z(\mathbf{u})$ values are estimated through a linear estimator $Z^*(\mathbf{u})$ in the following way (Goovaerts, 1997):

$$Z^*(\mathbf{u}) - m(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u})[Z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)] \qquad (1)$$

where $\lambda_\alpha(\mathbf{u})$ is the weight assigned to datum $z(\mathbf{u}_\alpha)$, interpreted as a realization of the RV $Z(\mathbf{u}_\alpha)$ while $m(\mathbf{u})$ and $m(\mathbf{u}_\alpha)$ are the expected values of the RVs $Z(\mathbf{u})$ and $Z(\mathbf{u}_\alpha)$. The Kriging methods aim to minimize the estimation variance $\sigma_E^2(\mathbf{u}) = Var\left(Z^*(\mathbf{u}) - Z(\mathbf{u})\right)$ where the RF $Z(\mathbf{u})$ is decomposed into a residual ($R(\mathbf{u})$) and trend ($m(\mathbf{u})$) component:

$$Z(\mathbf{u}) = m(\mathbf{u}) + R(\mathbf{u}) \tag{2}$$

where the residual component is modeled as a stationary RF with zero mean and covariance $C(\mathbf{h})$. In order to model $Z(\mathbf{u})$ the Kriging methods use the semivariogram ($\gamma(\mathbf{h})$), which considers the spatial relation of data and which is related to the covariance by:

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h}) \tag{3}$$

where $C(0)$ is the covariance when $\mathbf{h} = 0$, or the variance of Z. In order to apply the Kriging method, a theoretical variogram model ($\gamma(\mathbf{h})$) is fitted to the experimental variogram ($\widehat{\gamma}(\mathbf{h})$) which is computed by considering the difference between observations separated by a distance $\mathbf{h}$:

$$\widehat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} \left[z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h})\right]^2 \tag{4}$$

Three Kriging variants can be distinguished according to the way in which the trend $m(\mathbf{u})$ (eq. 2) is handled (Goovaerts, 1997): Simple Kriging (SK) which considers the trend to be known and constant on the study area, Ordinary Kriging (OK) which considers that the trend is unknown and constant on a specified search neighborhood and Universal Kriging or Kriging with a Trend model (KT) which considers that the local mean varies within each local neighborhood and on which the trend is modelled as a function of coordinates. Kriging with External Drift is an extension of KT, but in this case the trend is a function of one or more secondary variables (such as elevation) (Deutsch and Journel, 1998).

2.1 Ordinary Kriging

Ordinary Kriging (OK) accounts for local variation of the mean as it uses a local neighborhood $w(\mathbf{u})$ centered on the location $(\mathbf{u})$ being estimated. This Kriging technique considers the trend component $m(\mathbf{u})$ to be stationary, thus the linear estimation is expressed as a linear combination of the $n(\mathbf{u})$ RVs $Z(\mathbf{u}_\alpha)$ and the mean value m:

$$Z^*_{OK}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) Z(\mathbf{u}_\alpha) \tag{5}$$

where

$$\sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) = 1 \tag{6}$$

These equations yield the Ordinary Kriging system in terms of the semivariogram (Goovaerts, 1997; Wackernagel, 2003):

$$\begin{cases} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{OK}(\mathbf{u}) \gamma(\mathbf{u}_\alpha - \mathbf{u}_\beta) - \mu_{OK}(\mathbf{u}) = \gamma(\mathbf{u}_\alpha - \mathbf{u}) \text{ for } \alpha = 1, \dots, n(\mathbf{u}) \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{OK}(\mathbf{u}) = 1 \end{cases} \tag{7}$$

where $\mu(\mathbf{u})$ is a Lagrange parameter used to constraint the weights and $\gamma(\mathbf{u}_\alpha - \mathbf{u})$ represents the semivariogram for different lags, as described by (4). The term $\gamma(\mathbf{u}_\alpha - \mathbf{u})$ expresses the dissimilarities between each data point $(\mathbf{u}_\alpha)$ and the estimation point $(\mathbf{u})$, while $\lambda_\beta^{OK}(\mathbf{u})$ represents the weight values obtained by solving (7), which are then used in (5) to determine Z^*_{OK}.

2.2 Block Kriging

Block Kriging is an extension of Ordinary Kriging and uses a moving neighborhood or block of given dimensions to estimate the mean. For a block V of known dimensions, on which $\mathbf{u}_i$ samples are found, the block mean $z_v(\mathbf{u})$ can be computed as:

$$z_v(\mathbf{u}) = \frac{1}{N}\sum_{i=1}^{N} z(\mathbf{u}_i) \tag{8}$$

The block ordinary system is written as (Goovaerts, 1997):

$$\begin{cases} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta v}(\mathbf{u})\gamma(\mathbf{u}_\alpha - \mathbf{u}_\beta) + \mu_v(\mathbf{u}) = \gamma(\mathbf{u}_\alpha, V(\mathbf{u})) \text{ for } \alpha = 1,\ldots,n(\mathbf{u}) \\ \sum_{\beta=1}^{n(\mathbf{u})} = 1 \end{cases} \tag{9}$$

2.3 Kriging with External Drift

Kriging with External Drift (KED) (sometimes called Universal Kriging) should be used when a secondary variable is highly correlated with the variable of interest (Hudson and Wackernagel, 1994).

KED evaluates the correlation between the climatological variable and the secondary variable (in this case, elevation) within neighborhoods, providing information about the primary trend at location $\mathbf{u}$ (Goovaerts, 2000). The trend $m(\mathbf{u})$ is modeled as a linear function of one or more secondary variables $y(\mathbf{u})$:

$$m(\mathbf{u}) = a_0(\mathbf{u}) + a_1(\mathbf{u})y(\mathbf{u}) \tag{10}$$

The trend coefficients $a_0(\mathbf{u})$ and $a_1(\mathbf{u})$ are constant within the search neighborhood. In order to apply this Kriging technique the relation between the two variables must be linear (if not an appropriate transformation of the secondary variable is needed) and the value of the secondary variable must be known throughout the modeling domain. The KED estimator is (Goovaerts, 1997):

$$Z^{*}_{KED}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_{\alpha}^{KED}(\mathbf{u}) Z(\mathbf{u}_{\alpha}) \tag{11}$$

where the Kriging weights λ_{α}^{KED} are the solution of the $(n(\mathbf{u})+2)$ equation system (Goovaerts, 2000):

$$\begin{cases} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta}^{KED}(\mathbf{u})\gamma_R(\mathbf{u}_{\alpha}-\mathbf{u}_{\beta}) + \mu_0^{KED}(\mathbf{u}) + \mu_1^{KED}(\mathbf{u})y(\mathbf{u}_{\alpha}) = \gamma_R(\mathbf{u}_{\alpha}-\mathbf{u}) \text{ for } \alpha = 1,\ldots,n(\mathbf{u}) \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta}^{KED}(\mathbf{u}) = 1 \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta}^{KED}(\mathbf{u})y(\mathbf{u}_{\beta}) = y(\mathbf{u}) \end{cases} \tag{12}$$

The term $\gamma_R(\mathbf{u}_{\alpha}-\mathbf{u}_{\beta})$ represents the semivariogram of the residuals which are equal to the original values with the trend or drift removed (i.e. the semivariogram of the residuals).

3 Study area

The area under study is the Basin of Mexico, home to one of the largest metropolitan areas in the world: Mexico City and its Metropolitan Area (MCMA). The Basin has a mean elevation of 2240 meters above sea level (masl) and an approximate area of 9600 km^2. The analysis presented here used a rectangular area that encloses the Basin, as well as those climatological stations located near to it, as is illustrated in Fig. 1, thus using a total area of 16,800 km^2. This figure also shows that the Basin is surrounded by mountains and that its elevation ranges from 2000 masl up to 5500 masl on the high peaks located in

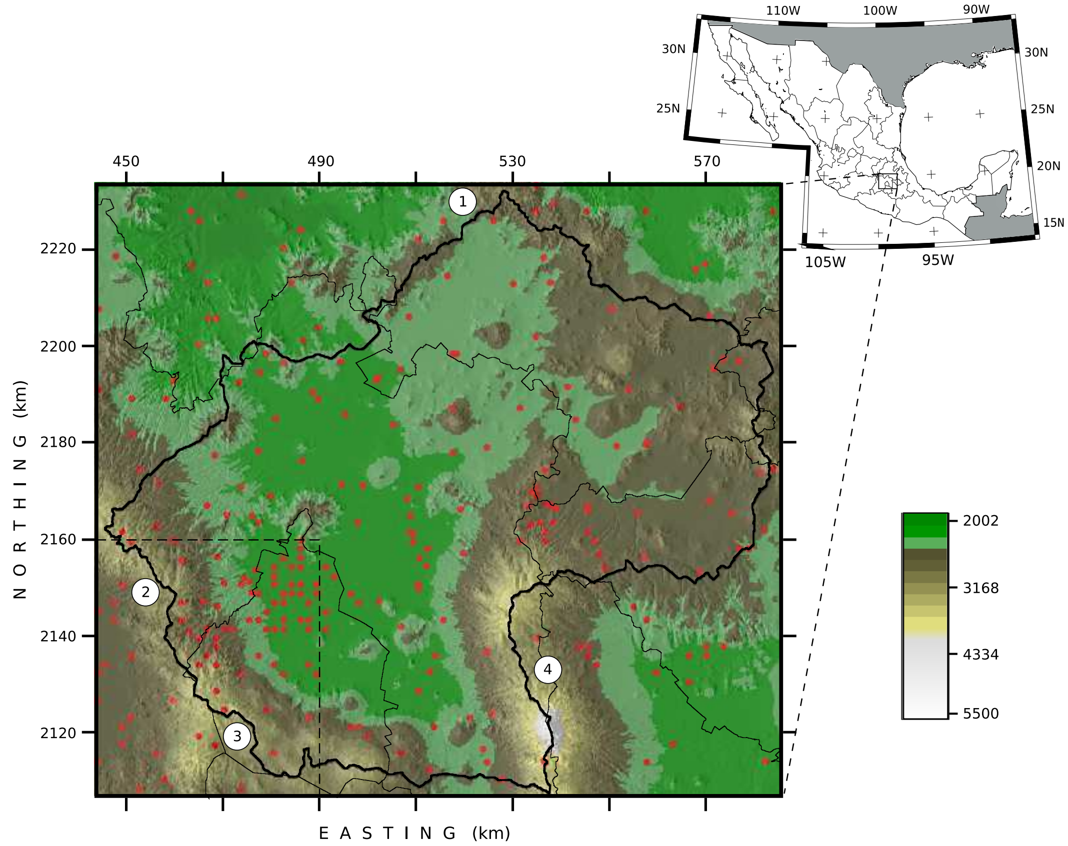

Fig. 1. Location and elevation of the study area; red dots represent the climatological stations used in the daily analysis. Circled numbers represent: 1) Sierra de Pachuca, 2) Sierra de las Cruces, 3) Sierra Chichinautzin and 4) Sierra Nevada. The dashed rectangle represents the Southwestern subarea.

the southeastern region.

The climatological data used in this study are stored in the Basin of Mexico Hydrogeological Database (BMHDB) (Carrera-Hernández and Gaskin, 2005), which is structured in such a way that its data can be directly used in (geo)statistical analyzes without further processing, as illustrated in Fig. 2, which shows how the database is accessed. The external variable used in this study (elevation) is stored as a Digital Elevation Model in the form of a raster map in GRASS, which is linked to R (R Development Core Team, 2005) through its GRASS library (Bivand, 2000). The climatological variables can be interpolated using a secondary variable through R's GSTAT library (Pebesma, 2004) which is used to fit the experimental semivariogram and to undertake the spatial interpolation; after this is achieved, the resulting distribution is written in the GRASS database as a raster map.

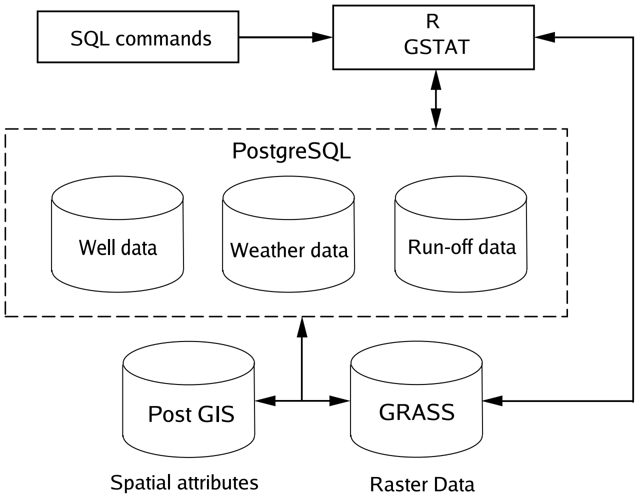

Fig. 2. Structure of the Basin of Mexico Hydrogeological Database and its interaction with (geo)statistical tools

The daily analysis presented used data from June 1978 and June 1985; these years were selected in order to use these results in a soil water balance, as there are other data such as land cover available for those two years. The number of stations with data within the Basin was also considered to select these years, as the quantity of climatological stations is not constant in time.

3.1 Data analysis

The spatial variation of the three climatological variables was undertaken at a grid resolution of 200 m as a resolution between 200 m and 5 km is considered to be appropriate in order to represent the variation of topographically dependent variables (Hutchinson and Galland, 1999). The 200 m resolution was chosen in order to account for the effects of topography on both temperature and rainfall. In order to analyze the effect of elevation as a secondary variable on spatial interpolation, daily data from June 1978 and June 1985 were used, as June is the month in which precipitation is largest in the study area. The daily correlation between the climatological variables and elevation is shown in Fig. 3(a) for June 1978 and Fig. 3(b) for June 1985. This figure shows that rainfall has a very low correlation with elevation, ranging from 0.00 to a maximum of 0.25 in 1978 and 0.18 in 1985. Maximum temperature exhibits a better correlation with elevation (in general above 0.3), although for two days this correlation is as low as 0.17 but reaches a maximum value of 0.60 on

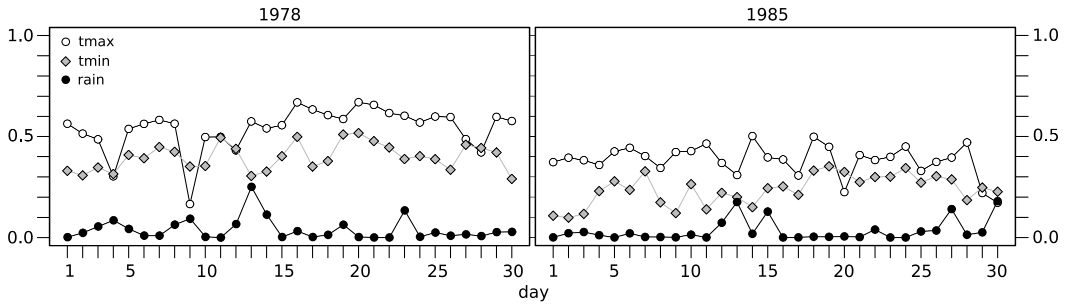

Fig. 3. Correlation of daily climatological variables for (a)June 1978 and (b) June 1985

some days. As the present study comprises a large area, the minimum value of rainfall for all days is zero while its maximum value exceeds 100 mm in three days during June 1978 and only once in 1985 (Fig. 4).

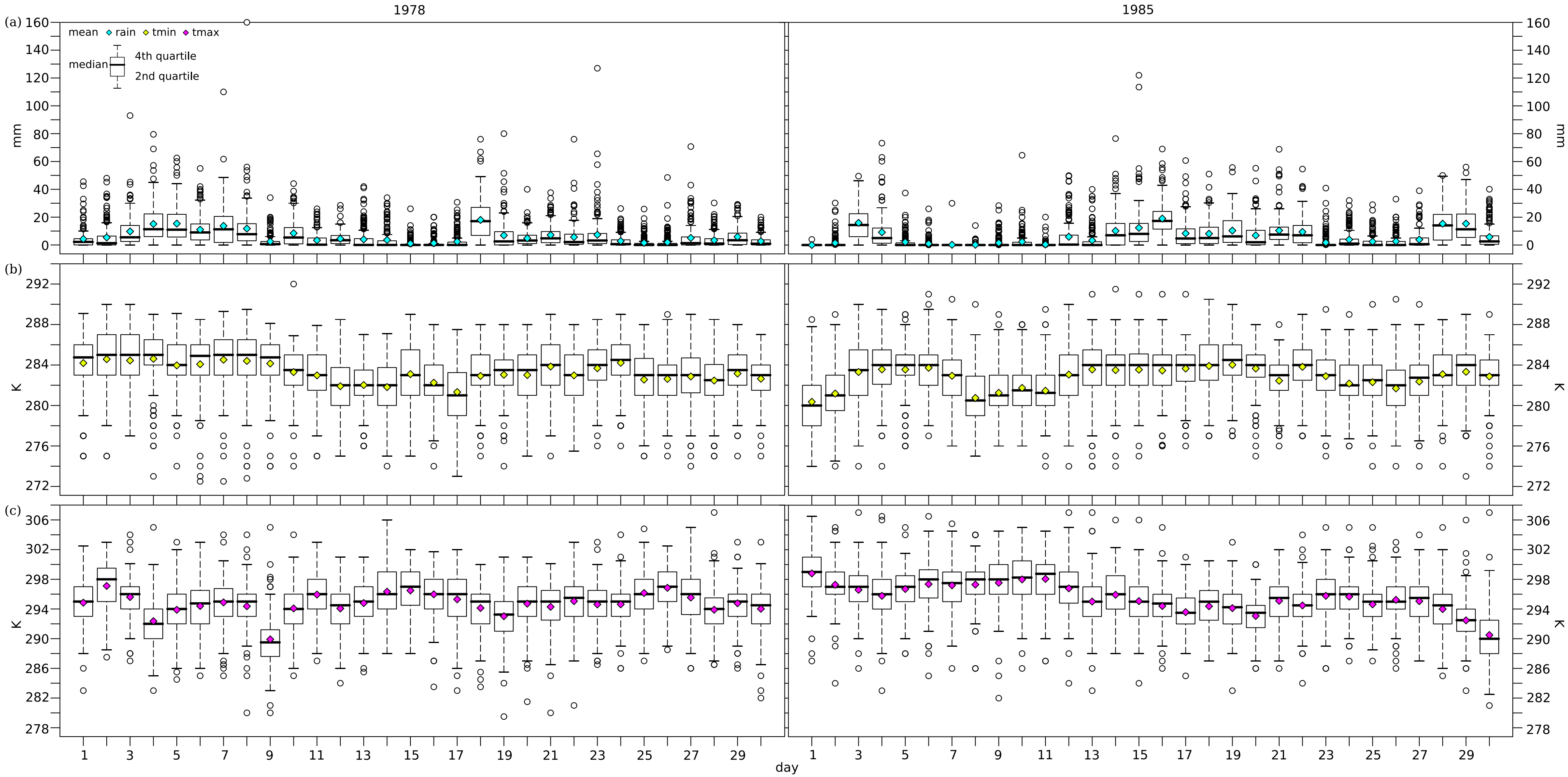

Fig. 4. Whisker plots of daily climatological variables for June 1978 and 1985: (a) rainfall, (b) minimum temperature and (c) maximum temperature

Before applying the Kriging algorithm, a semivariogram model needs to be fitted to the experimental semivariogram values in order to infer the model that best represents the spatial variation of the climatological variables. To this end, different models were used and fitted manually to the experimental semivariogram of the daily data; during this stage it was observed that the Bessel model was the one that best represented all three of the climatological variables. As a result, the Bessel semivariogram was used in the automatic fitting procedure for daily data, which is expressed as (Pebesma and Wesseling, 1998):

$$\gamma(\mathbf{h}) = 1 - \frac{\mathbf{h}}{a} K_1 \left(\frac{\mathbf{h}}{a} \right) \tag{13}$$

where K_1 is the first order modified Bessel function of the second kind, a the range which represents the distance at which the semivariogram reaches its maximum value (sill) and $\mathbf{h}$ represents distance.

In order to undertake the daily interpolations, the automated procedure provided by GSTAT (Pebesma, 2004) which uses weighted least squares was used to fit the Bessel semivariogram. The semivariograms fitted in this way are shown in Fig. 5, which also shows the semivariograms of the residuals. The semivariograms of the residuals show the effect of removing the trend due to elevation on the experimental semivariogram. It can be noticed from this figure that when the climatological variable presents higher correlation with elevation (Fig. 3) the semivariogram of the residuals has a lower sill than the original values; this effect is quite obvious for maximum temperature in June 23, 1978 (Fig. 5(c) which has a correlation of 0.60. For this variable on this particular date, the fitted semivariogram of the residuals shows a range of nearly 5 km and a sill of approximately 3 K^2, while the semivariogram fitted to the temperature data shows an asymptotic range. The difference between these semivariograms indicates that temperature has a drift caused by elevation, which would not be the case if the semivariograms were identical. The effect of removing the drift caused by elevation for maximum temperature had the same effect on June 15, 1985. It can be noted that the semivariograms fitted to both rainfall and its residuals are almost identical (Fig. 5(a)). The difference between the semivariogram of the variable and of its residual shows that the climatological variables exhibit a trend due to elevation; this is obvious for both minimum and maximum temperature (Fig. 5(b) and 5(c)) but not for rainfall (Fig. 5(a)). From this same figure, it can be seen that the automatic fitting procedure yielded an interesting theoretical semivariogram for rainfall in June 15, 1985 and for maximum temperature in June 3, 1985 as closer values show larger semivariances which is the opposite behavior of the theoretical semivariogram chosen. This situation might be improved if a smaller range is chosen for both the experimental and the theoretical semivariogram; however, the range was not modified as an automatic procedure is needed to undertake the large amount of data used in this work.

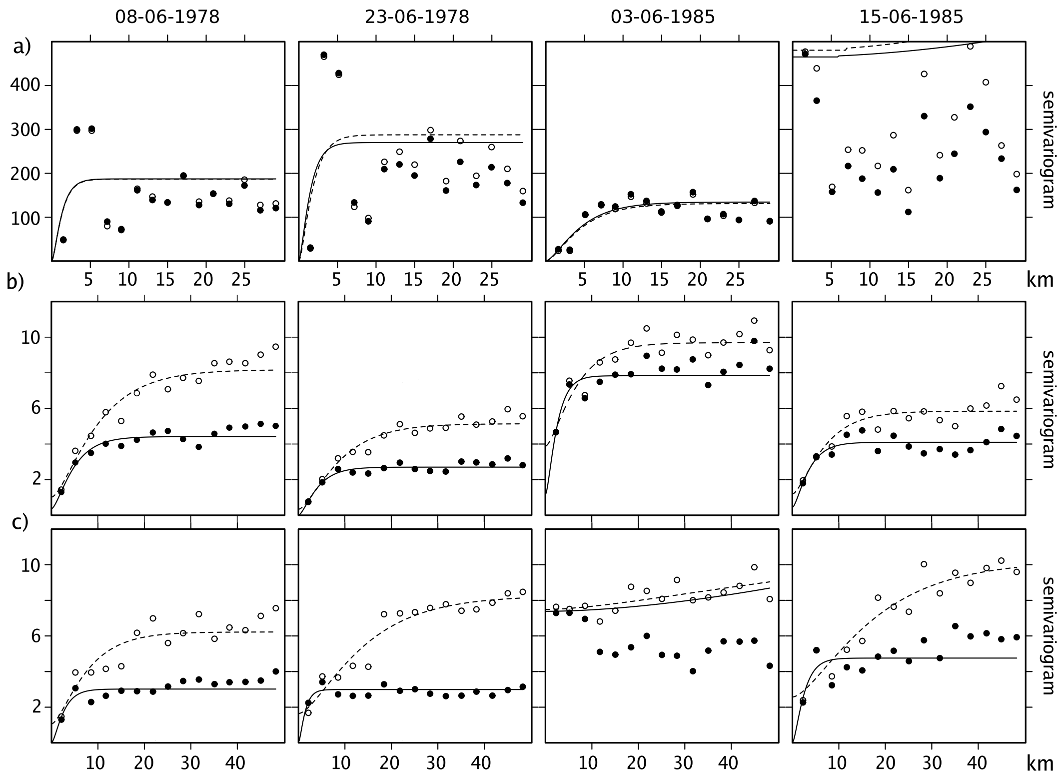

Fig. 5. Semivariograms of daily climatological variables: (a) rainfall, (b) minimum temperature and (c) maximum temperature. Circles represent the experimental semivariogram of the variables while the dashed line is the fitted semivariogram to these values. Black dots represent the semivariograms of the residuals and the solid line its fitted semivariogram

The spatial distribution of the climatological variables was developed using: Ordinary Kriging (OK) in a global neighborhood (Fig. 6), Kriging with External Drift (KED) in a global neighborhood (Fig. 7), Block Kriging with External Drift (BKED) using 30 x 30 km^2 blocks (Fig. 8), Ordinary Kriging on a local neighborhood (OK_l) using a 20 point neighborhood (Fig. 9) and

Kriging with External Drift on a local neighborhood, using 20 points as well (Fig. 10). Although daily interpolations were undertaken for both June 1978 and June 1985, the aforementioned figures show the spatial distribution of the climatological variables for only four different dates: two days in June 1978 and two in June 1985. For June 1978, days 8 and 23 were selected because they exhibit the largest precipitation values, with 160 and 127 mm respectively (Fig. 4); this was the same criterion used to select June 23, 1985 (122 mm), while June 3, 1985 was selected based on the fact that the maximum rainfall depth for that day shows the smallest difference with its mean (Fig. 4). It should be noted that negative values were produced by the interpolation algorithms at some points, which were considered to have a value equal to zero.

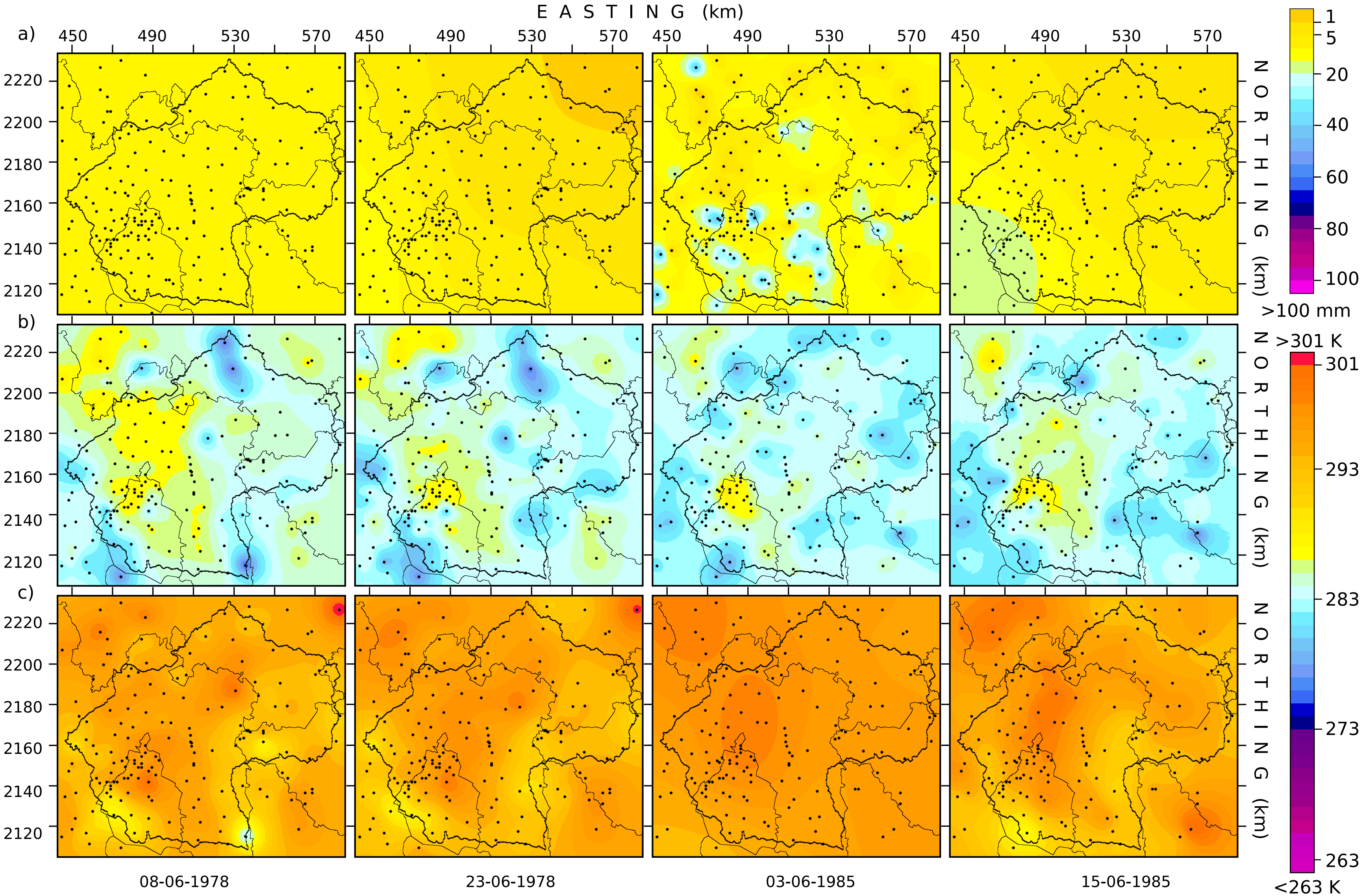

Fig. 6. Interpolation of daily climatological variables using Ordinary Kriging: (a) rainfall, (b) minimum temperature and (c) maximum temperature

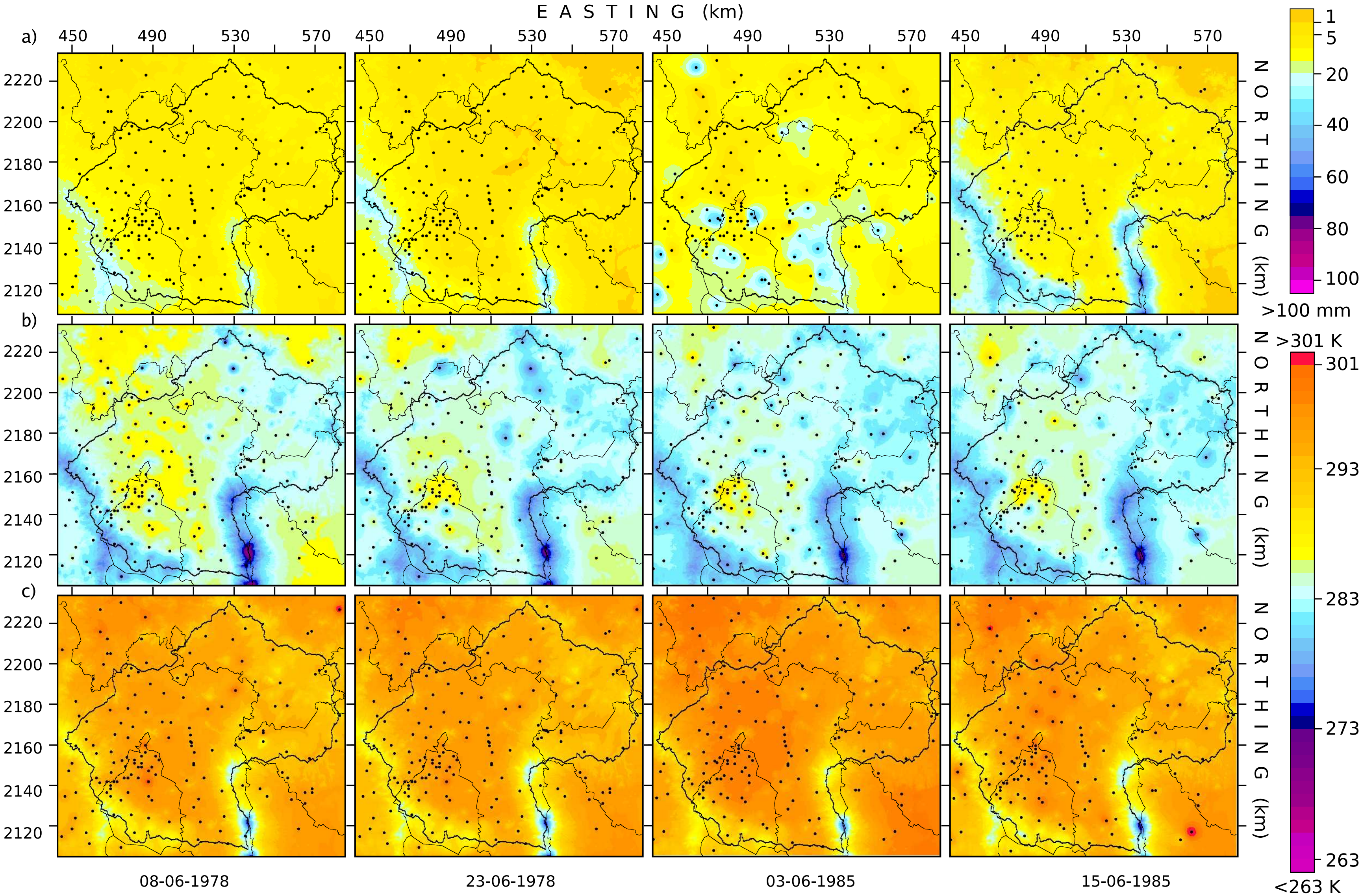

Fig. 7. Interpolation of daily climatological variables using Kriging with External Drift: (a) rainfall, (b) minimum temperature and (c) maximum temperature

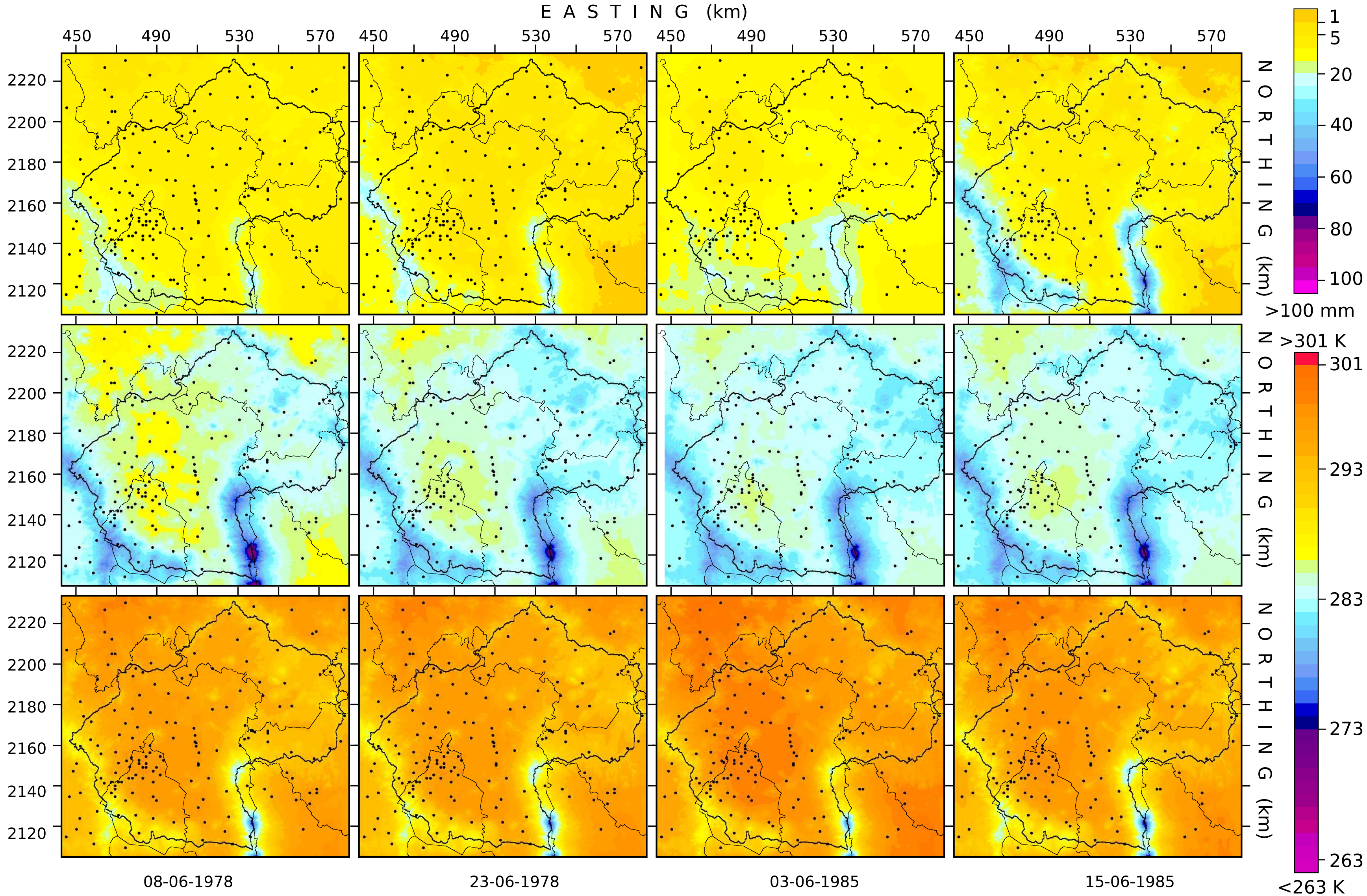

Fig. 8. Interpolation of daily climatological variables using Block Kriging with External Drift: (a) rainfall, (b) minimum temperature and (c) maximum temperature

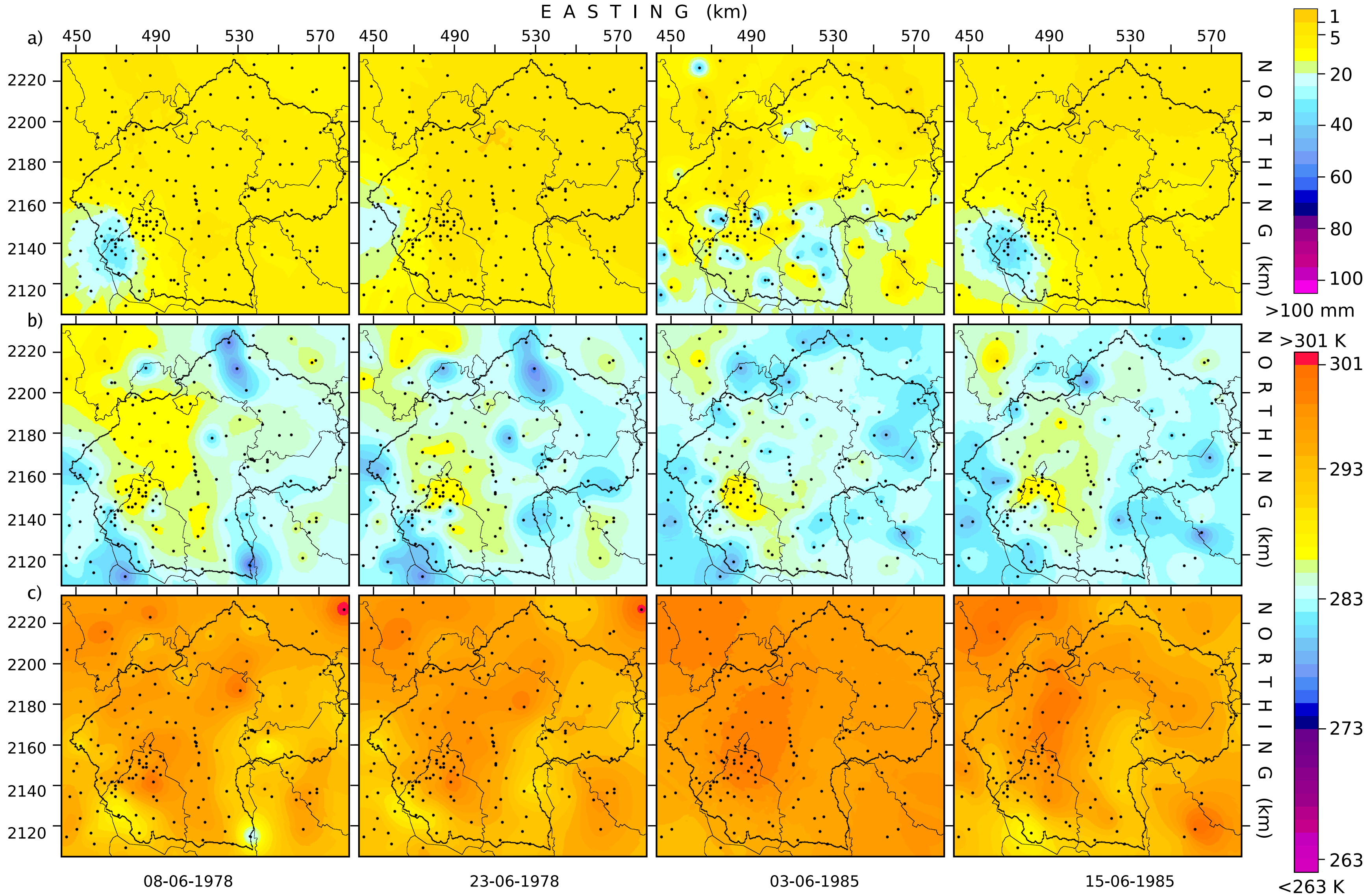

Fig. 9. Interpolation of daily climatological variables using Ordinary Kriging on a local neighborhood: (a) rainfall, (b) minimum temperature and (c) maximum temperature

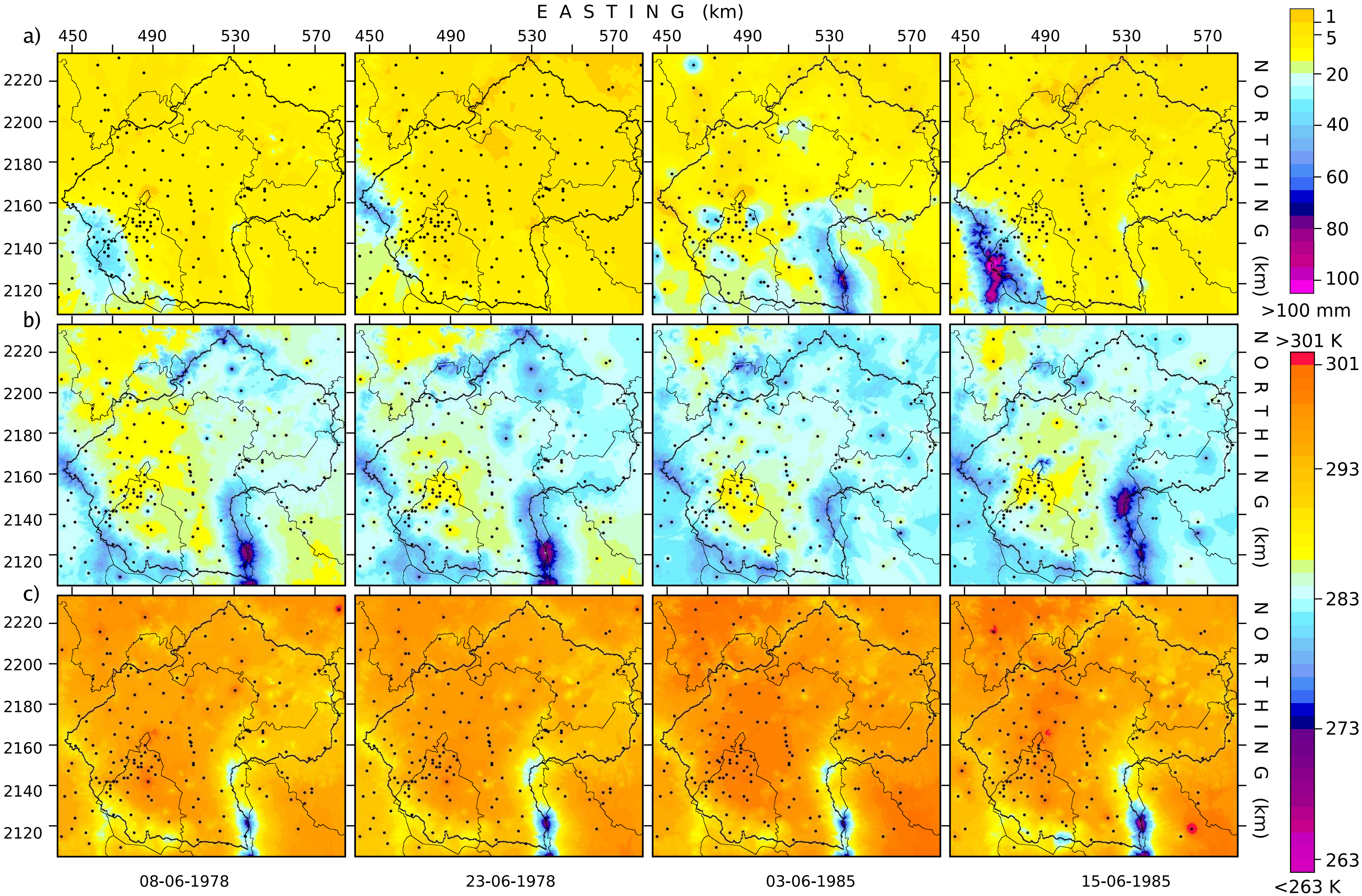

Fig. 10. Interpolation of daily climatological variables using Kriging with External Drift on a local neighborhood: (a) rainfall, (b) minimum temperature and (c) maximum temperature

4 Discussion

4.1 Visual inspection of spatial interpolation

4.1.1 Rainfall

The spatial variation of rainfall obtained by Ordinary Kriging (OK) (Fig. 6(a)) does not resemble the spatial pattern in the original data, which is expected due to their large variation and by their non-stationarity throughout the study area. For June 8, 1978 OK yields a uniform spatial distribution of rainfall, while its distribution in June 23, 1978 and June 15, 1985 shows an increasing trend towards the southwestern region of the Basin. The rain distribution for June 23, 1978 shows some scattered points where precipitation values are above 20 mm. The interpolated maps are quite different when Kriging with External Drift is used (Fig. 7(a)), as precipitation values are larger in the mountainous regions located southwards except for June 3, 1985 for which the use of elevation as an external variable produced a more continous surface for precipitation values between 20 and 25 mm without increasing them over the mountainous areas.

The rainfall maps produced with Block Kriging with External Drift (BKED) (Fig. 8(a)) are very similar to those produced with KED, except for June 3, 1985 on which the larger precipitation spots located towards the Basin's centre are not reproduced and larger precipitation is shown at higher elevation areas (e.g. Sierra Nevada). When using Ordinary Kriging on a local neighborhood (Fig. 9(a)), the resulting distribution of rainfall is quite different than the one obtained by using the previous three methods, in particular for June 8, 1978 and June 15, 1985. These results should be expected by considering the large differences between maximum and mean rainfall for those dates (Fig. 4) as OK_l uses the mean of the 20 climatological stations closest to the point being interpolated while OK considers the mean of the entire domain. For a day in which the difference between the mean and maximum value is smaller the outcome of OK in a local neighborhood is similar to that of OK on a global neighborhood. This is the case for June 3, 1985 which is the day on which the mean and maximum value of rainfall show the smallest difference.

The local distribution of rainfall (i.e. heterogeneity) is more noticeable when Kriging with External Drift on a local neighborhood is applied (Fig. 10a), producing precipitation values above 100 mm in the Southwestern area of the Basin for June 15, 1985 and over 70 mm in the south. The spatial distribution of rainfall obtained with this last method shows larger precipitation events on more elevated areas, where the effect of elevation of rainfall is accounted for locally, as is evident for June 15, 1985. If the relationship between precipitation and elevation is accounted for globally, other mountainous areas would also show larger precipitation events, as is the case when using KED on a global neighborhood (Fig. 7a). By visual inspection, the maps developed with KED on a local neighborhood show a more realistic distribution of rainfall; however, in order to obtain a more quantitative criterion, cross validation was used (Isaaks and Srivastava, 1989). With this procedure, one point is removed at a time and interpolation is undertaken; this is done for all points in the dataset and the root mean square error (RMSE) between the interpolated and the observed values is computed. This method was applied for the different Kriging methods as shown in Table 1 except for BKED which was omitted as it estimates the average value over a block (Goovaerts, 1997).

The RMSE values of the Kriging methods used (Table 1) do not show a drastic difference and although KED_l has the lowest RMSE, these values do not provide a solid basis on which to decide the better interpolation method. In order to provide further insight into the performance of each method, maps of monthly accumulated precipitation for June 1978 and June 1985 were developed by adding the daily interpolation maps (Figures 11 and 12, respectively).

Table 1

Cross validation for Ordinary Kriging (OK), Kriging with External Drift (KED),Ordinary Kriging on a local neighborhood (OK_l), Kriging with External Drift on a local neighborhood (KED_l)

	8-6-1978			23-6-1978			3-6-1985			15-6-1985		
	rain	tmin	tmax	rain	tmin	tmax	rain	tmin	tmax	rain	tmin	tmax
OK	14.56	2.50	2.48	12.09	1.99	1.97	9.82	2.99	2.72	14.37	2.20	2.79
KED	14.34	2.18	2.00	11.70	1.79	1.77	9.87	2.89	2.36	13.94	2.06	2.41
OK_l	13.31	2.48	2.48	11.41	1.99	1.99	9.40	3.02	2.67	12.72	2.23	2.77
KED_l	13.65	2.18	2.09	11.49	1.76	1.86	9.47	3.01	2.42	11.76	2.13	2.36

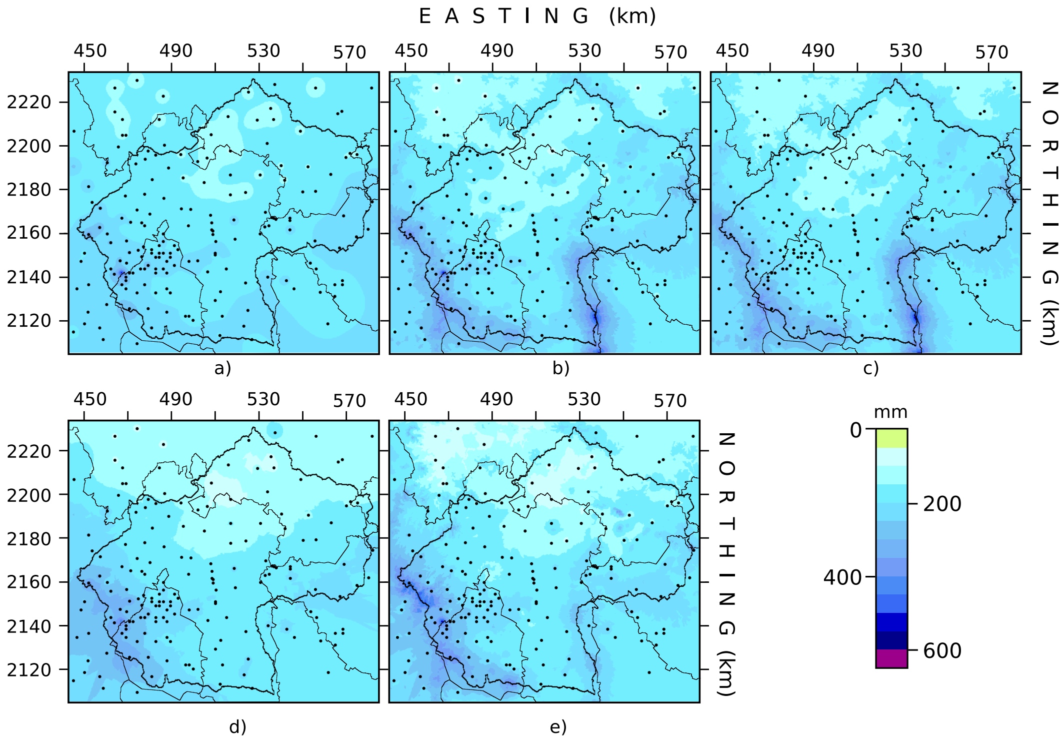

Fig. 11. Accumulated rainfall for June 1978 derived from daily interpolated maps using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Local Kriging and (e) Local Kriging with External Drift

The spatial distribution of rainfall obtained for June 1978 (Fig. 11) shows that for four out of the five interpolation methods, the effect of mountainous terrain can be noticed, even the map derived from OK_l, which does not consider elevation as an auxiliary variable shows larger precipitation values at the Sierras Chichinautzin and Las Cruces (Fig. 11d). The effect of topography on each monthly map is different according to each Kriging technique, which was also the case for the daily distributions due to the fact that OK and KED use both a global mean and global relationship between rainfall and elevation (e.g. they use all the stations to determine these values). The monthly rainfall pattern obtained with KED shows larger values as elevation increases, an effect that is observed on the southern sierras that enclose the Basin. The maps produced by KED_l (Fig. 11(e)) show larger precipitation values only at the Sierras de las Cruces and Chichinautzin which is also the case for the monthly map derived from the OK_l interpolations. The difference between these two maps is the maximum value of precipitation at these Sierras, as KED_l extrapolates precipitation using the local trend caused by elevation. In fact, the map derived from the KED_l daily interpolations shows larger precipitation values at

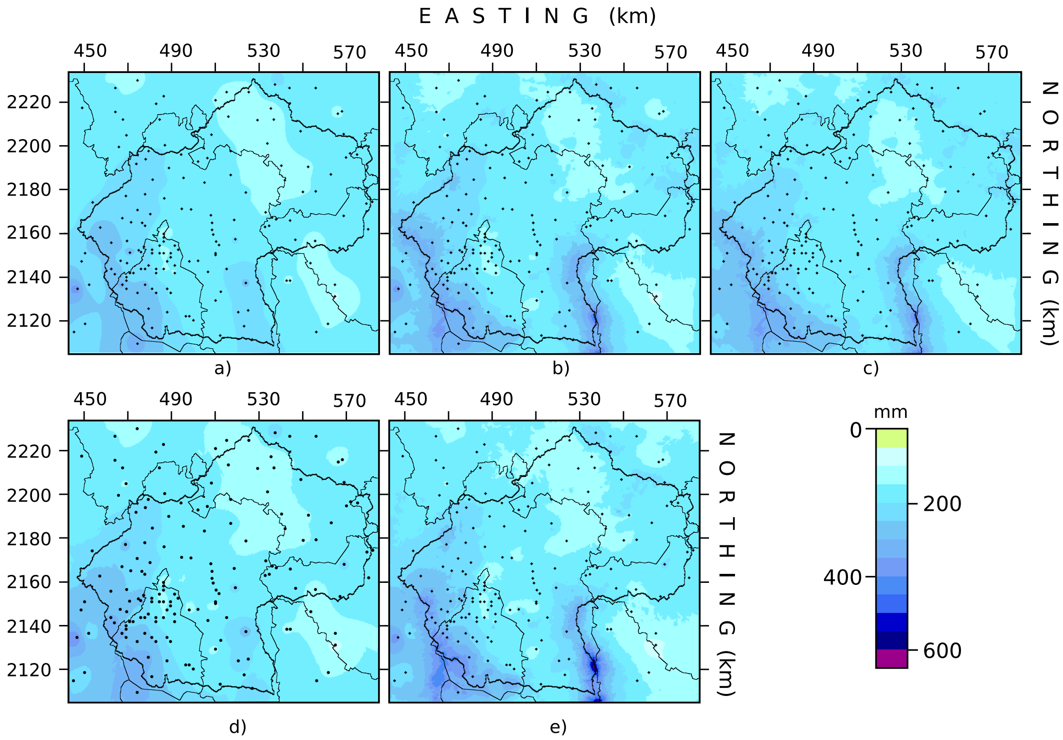

Fig. 12. Accumulated rainfall for June 1985 derived from daily interpolated maps using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift (d) Local Kriging and (e) Local Kriging with External Drift

the Sierra Nevada, although smaller than those at the Sierra Las Cruces. The monthly maps for 1985 (Fig. 12) show the same effect: precipitation is extrapolated, except that the Sierra Nevada shows larger rainfall values in the map derived from the KED_l interpolations. From these two figures, it can be concluded that the southern Sierras are the ones that cause the orographic effect on precipitation and that each of the Kriging methods produce maps with different rainfall distribution; however it is hard to choose a better method based on the monthly maps and on the cross validation values, as these values are not significantly different from each other. In order to provide further insight into this problem, the difference between the interpolated and observed values will be analyzed by their location and elevation after analyzing the spatial distribution of temperature.

4.1.2 Temperature

The temperature maps obtained from the different Kriging methods differ from the precipitation maps in that the only maps that show the effect of topography are those that use elevation as an auxiliary variable. It is interesting to note the effect that the presence of one particular station has on the spatial distribution of temperature, which is the climatological station *Atlauta*. This station is the one with the highest elevation, located at approximately 3700 masl in the Sierra Nevada near the southeasternmost point of the watershed limit and with temperature records for only one of the four selected days. The effect of the records from this station can be easily identified on the spatial distribution of temperature using OK (Fig. 6) for June 8th, 1978, which shows how both minimum and maximum temperatures decrease in the sierras that surround the Basin, particularly on the Sierra Chichinautzin and Sierra Nevada. However, for the remaining three days on which no data from the Atlauta station are availabe, the southern part of the Sierra Nevada has higher temperatures than the Sierra Chichinautzin. This difference of temperature is not feasible, as in this area the highest peaks of the region are located with an elevation above 5,000 masl and are snow-capped all year round. This anomaly is due to the fact that OK does not consider the relationship between elevation and the climatological variable; furthermore, the density of the climatological stations in this area is not appropriate to capture the effects due to the drastic change in topography.

The maps developed from the use of KED (Fig. 7) for both temperatures show lower minimum temperatures in the Sierra Nevada than those observed in the Sierra Chichinautzin and Las Cruces, reaching 273 K even in those days without data from the Atlauta climatological station. This is the result of using the relationship between temperature and elevation, as it is used to extrapolate temperature on elevations above that on which the highest climatological station is located. As in the case of the maps derived by using OK, the ones developed by KED show an increase in minimum temperature in the area in which Mexico City is located, an effect known as the urban heat island effect. The spatial extent of this effect varies according to the Kriging technique used; the maps developed from BKED (8) show a smaller area with higher temperatures than those obtained from the other Kriging techniques as can be easily seen for June 3, 1985. The minimum temperature maps produced by KED_l (10)show lower temperatures in the Sierra Nevada, which should be expected; however due to the small density of the climatological network in that area, the maps show a drastic change in temperature instead of a smooth and continuous pattern. This is caused by the effect of using a 20 point neighborhood and the areas that show this drastic change use a station that has data not related to the estimation point (i.e. the datum used is too far from the estimation point). To overcome this problem both smaller and larger neighborhoods were used without improving the results; when a smaller neighborhood was used, the spatial distribution of minimum temperature showed more drastic changes, due to the small density of climatological stations in some areas. When a larger neighborhood was used, the spatial distribution of temperature was almost identical to the one obtained from KED on a global neighborhood.

4.2 Error analysis

4.2.1 Rainfall

In order to verify which Kriging method yields better results, the monthly accumulated maps derived from daily interpolations (Figures 11 and 12) were sampled at the location of each climatological station and the absolute difference between them and the accumulated point values was computed. The statistics of these values are grouped in Table 2 which shows that for June 1978, KED provides the smallest mean value of errors, but that it is almost equal to the mean error value obtained with KED_l (22.12 and 22.16) while the correlation of the computed and observed values is also better for KED, although they are also quite similar (0.83 and 0.81 respectively). KED_l also yields a larger maximum difference value (145.41 mm) than KED (109.42 mm) and Block Kriging is the method that yields both the largest mean and maximum difference value as well as the least correlation between observed and computed monthly accumulated rainfall ($\rho = 0.45$). For june 1985, the correlation between these monthly values is the same for OK, KED, OK_l and KED_l; while KED_l yields the smallest mean difference value, it is worth noting that again, this value is not drastically different from the one obtained by OK_l (18.65 and 19.43 respectively).

Table 2
Statistics for the difference between meassured monthly accumulated values and monthly accumulated values derived from daily interpolations for Ordinary Kriging (OK), Kriging with External Drift (KED), Block Kriging (BK), Ordinary Kriging in a local neighborhood (OK_l) and Kriging with External Drift on a local neighborhood (KED_l)

		June 1978					June 1985				
var	method	ρ	mean	std. dev.	min	max	ρ	mean	std. dev.	min	max
rain	OK	0.84	24.58	20.45	0.32	128.27	0.85	20.20	19.70	0.06	131.17
	KED	0.83	22.12	19.50	0.12	109.42	0.85	20.61	18.86	0.44	137.82
	BKED	0.45	41.78	32.49	0.18	159.27	0.56	32.20	31.83	0.08	245.09
	OK_l	0.77	25.23	22.90	0.25	147.14	0.85	19.43	18.53	0.50	123.13
	KED_l	0.81	22.16	21.77	0.09	145.41	0.86	18.65	17.33	0.16	138.44
tmin	OK	0.96	10.86	9.51	0.01	51.17	0.84	20.21	20.67	0.11	115.06
	KED	0.98	7.72	6.88	0.06	40.45	0.83	20.21	21.56	0.01	134.64
	BKED	0.60	29.41	30.47	0.27	189.59	0.52	31.30	32.70	0.29	205.66
	OK_l	0.96	10.66	9.61	0.01	50.99	0.84	20.16	20.66	0.09	112.75
	KED_l	0.98	7.41	6.78	0.06	39.06	0.82	19.78	20.36	0.35	125.32
tmax	OK	0.92	19.05	15.65	0.04	89.75	0.76	30.14	31.05	0.43	174.97
	KED	0.94	18.48	16.01	0.50	88.50	0.86	23.39	22.26	0.34	151.49
	BKED	0.72	30.80	27.14	0.51	149.93	0.57	37.66	36.09	0.28	215.18
	OK_l	0.92	18.86	15.54	0.08	88.54	0.76	29.83	30.81	0.39	174.70
	KED_l	0.98	9.55	8.72	0.15	49.40	0.86	23.40	21.37	0.02	146.32

The values of measured and derived monthly accumulated rainfall for June 1978 and 1985 were plotted along with their regression line as illustrated in Figs. 13 and 14 which show how for both months the regression line fitted to the observed and derived values from the KED_l interpolations is closer to the 45° line (Figs. 13e, and 14e) and that the points show less dispersion; the 45° line is shown in order to have a reference on where a perfect fit between the observed and derived accumulated value should be located. The scattergrams of BK show more dispersion, which should be expected by considering the values of Table 2. In order to analyze whether or not the maximum errors were located at higher or lower elevations, the absolute difference between observed and derived monthly values are shown in Fig. 15 for June 1978 and Fig. 16 for June 1985.

For 1978 (Fig. 15) it can be noted that errors are larger when BKED is used and that when elevation is used as a secondary variable the errors for low and high elevations are in general smaller than when elevation is not used; furthermore, when KED_l is used (Fig 15(e)) the errors are in general lower. The opposite is observed in the case of BKED as the errors are distributed throughout the elevation range for both years. For 1985 (Fig. 16) it can be seen that by using elevation as a secondary variable, the errors observed are smaller except for BKED as previously mentioned. The difference between using a local neighborhood instead of a global one is evident by comparing the distribution of errors with elevation obtained from KED_l (Fig. 16(e)) and KED (Fig. 16(b)) as most of the errors derived from the use of KED_l are below 50 [mm]. An interesting pattern can also be observed from these figures, as the largest errors for four out of the five Kriging techniques are observed at the same elevations (near 2600 masl) which indicates that some type of error can be associated to the climatological stations located at that elevation and that probably the same stations are the ones which exhibit these large errors. In order to verify whether or not the errors are associated with a particular climatological station, the spatial distribution of errors will also be analyzed.

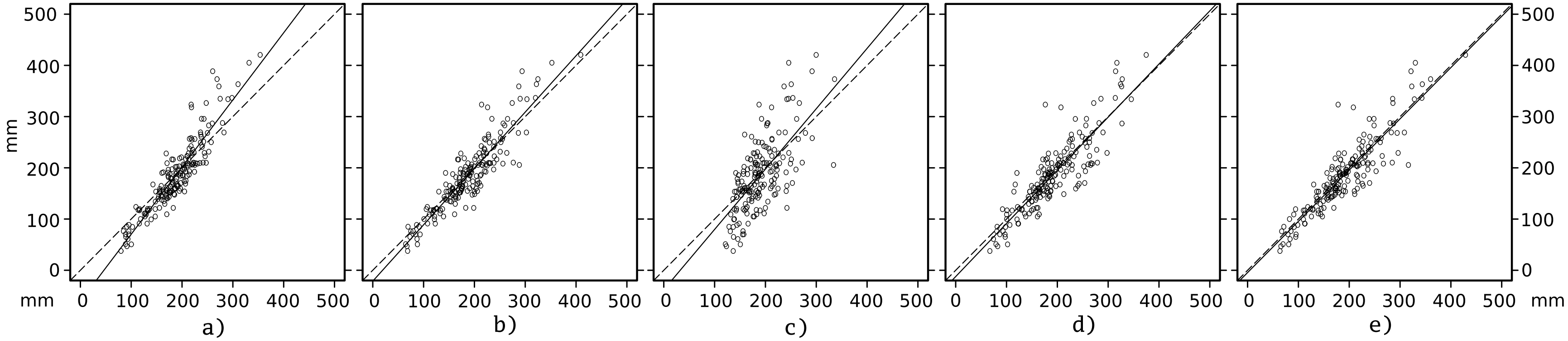

Fig. 13. Comparison between measured accumulated monthly rainfall and monthly accumulated rainfall derived from daily interpolated rainfall for June 1978 using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

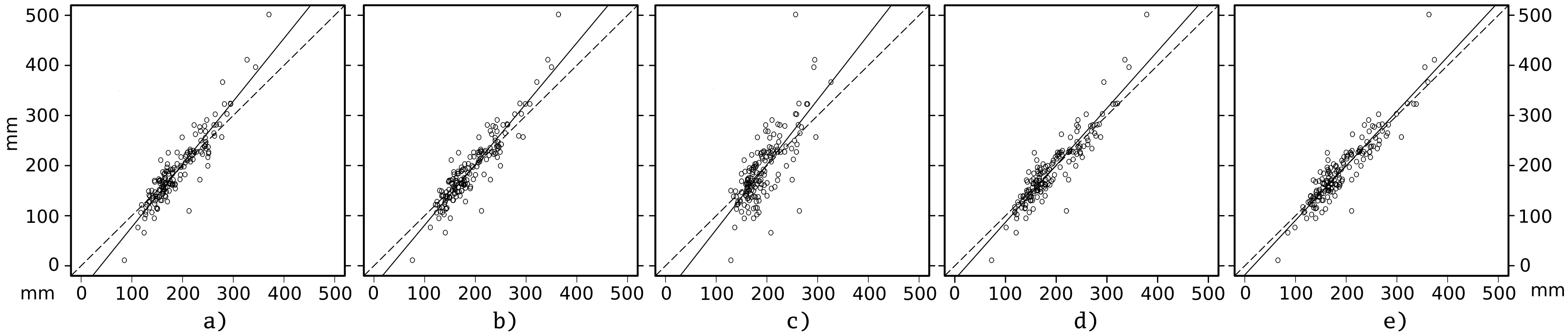

Fig. 14. Comparison between measured accumulated monthly rainfall and monthly accumulated rainfall derived from daily interpolated rainfall for June 1985 using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

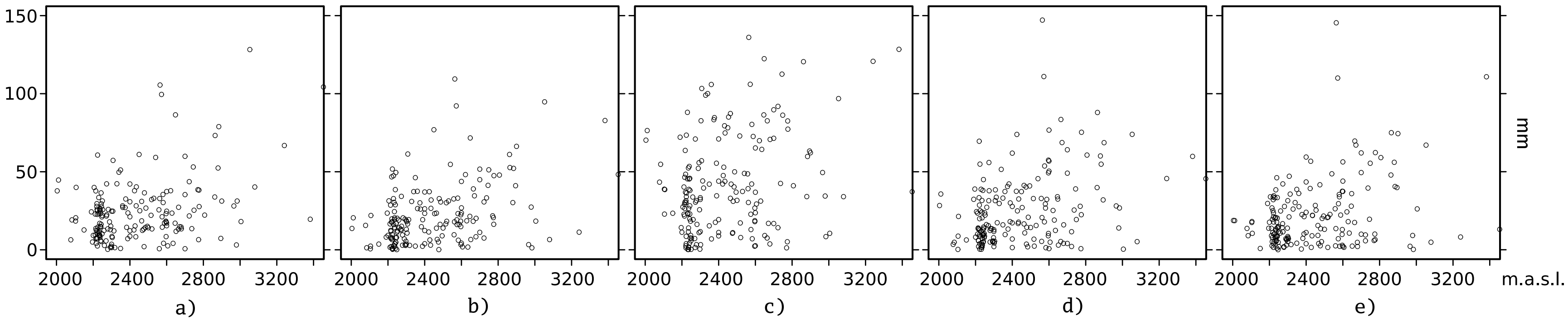

Fig. 15. Elevation of climatological stations and absolute difference between measured accumulated rainfall for June 1978 and accumulated monthly rainfall from daily interpolations using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

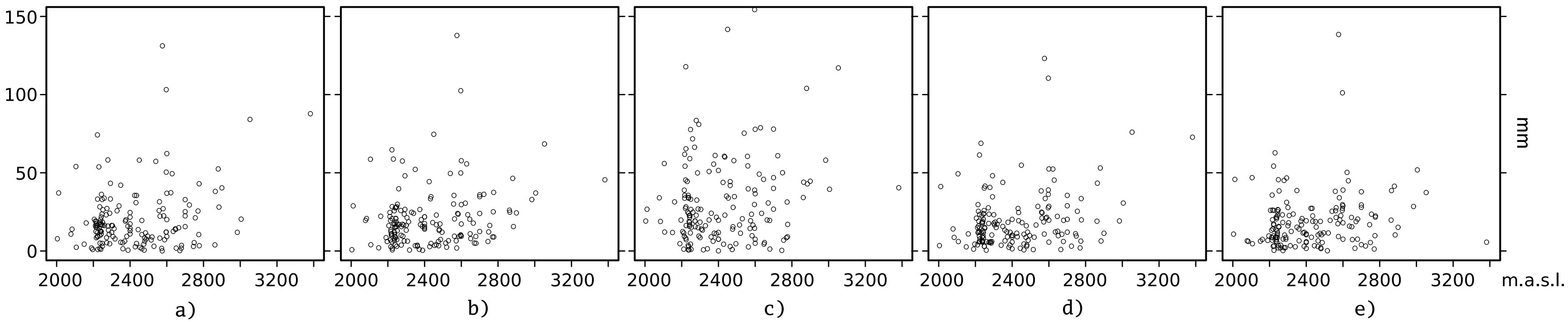

Fig. 16. Elevation of climatological stations and absolute difference between measured accumulated rainfall for June 1985 and accumulated monthly rainfall from daily interpolations using: (a) Ordinary Kriging,(b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

4.2.2 Temperature

This analysis will only compare the observed and the interpolated monthly accumulated value for both minimum and maximum temperature, as done in the previous section for rainfall. These values were also plotted on a scattergram and a regression line was fitted to them using June 1978 (Figs. 17 and 19) and June 1985 (Figs. 18 and 20). The statistics of the errors are shown in Table 2 in which it can be observed that the method that yields the largest mean value of errors for both temperatures and for both months as well as the minimum correlation value is BKED while the remaining four methods yield similar values. For minimum temperature in June 1978, KED and KED_l exhibit the highest correlation value (ρ =0.98) between the accumulated observed temperature and the accumulated interpolated temperature, and a similar maximum, minimum and mean value of errors which is also the case for OK and OK_l. For June 1985 the methods which yield the largest correlation value are OK_l and OK, although it is similar to the correlation obtained by KED and KED_l; again, the difference between the mean error value is not drastic for the four Kriging methods. Based on the values shown on Table 2, the best results are provided by KED_l although the statistics obtained from KED are almost identical to the previous method. However if the spatial pattern produced by each Kriging method is considered (Figs. 6, 7, 8, 9 and 10) then KED would be the preferred option due to the spatial continuity of minimum temperature throughout the study area, as discussed in section 4.1.2.

In the case of maximum temperature, the methods that yield the largest correlation values between the observed and interpolated accumulated values for 1978 are KED and KED_l (with ρ =0.94 and ρ =0.98 respectively) as well as for 1985 ($\rho = 0.86$ for both cases). These two methods provide the smallest mean value of error, which is the same for both techniques in 1985, while it doubles from 9.78 K for KED_l to 18.48 for KDE in 1985. As these values are similar for both techniques and both years, the spatial distribution of the interpoalted variable should be used as previously done for minimum temperature. Again, the method that should be used is KED as the maps produced with KED_l shows discontinuities in some areas due to the low density of climatological stations in some areas.

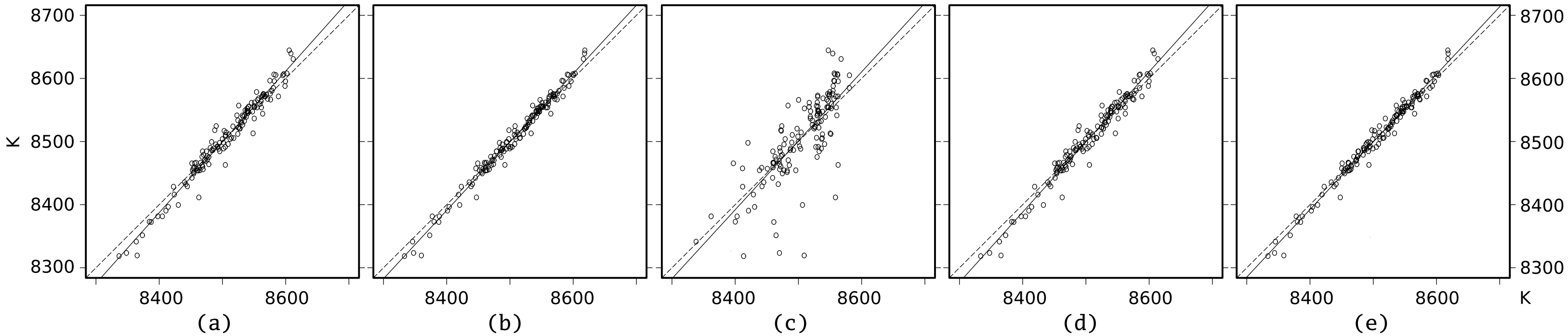

Fig. 17. Comparison between monthly accumulated minimum temperature and monthly accumulated minimum temperature derived from daily interpolated minimum temperature for June 1978 using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

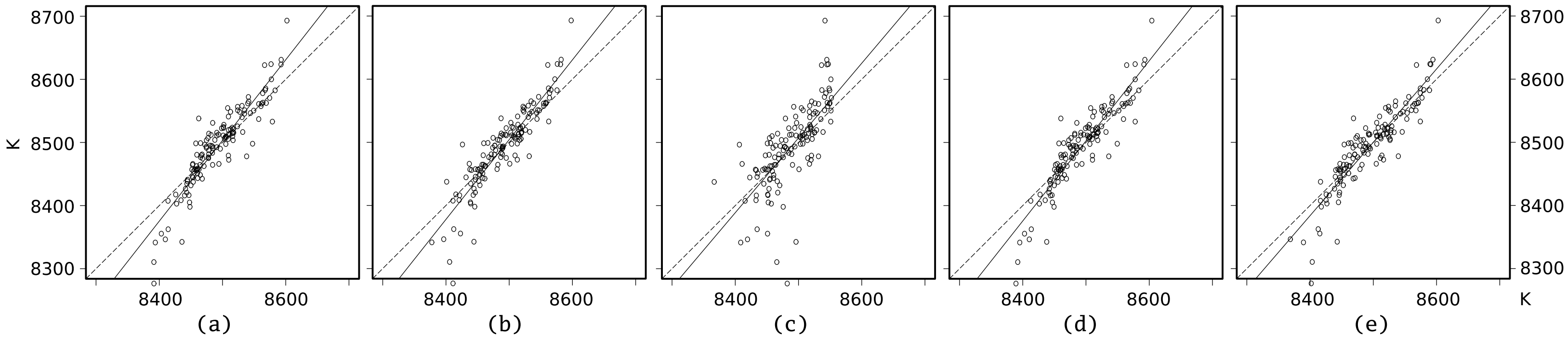

Fig. 18. Comparison between monthly accumulated minimum temperature and monthly accumulated minimum temperature derived from daily interpolated minimum temperature for June 1985 using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

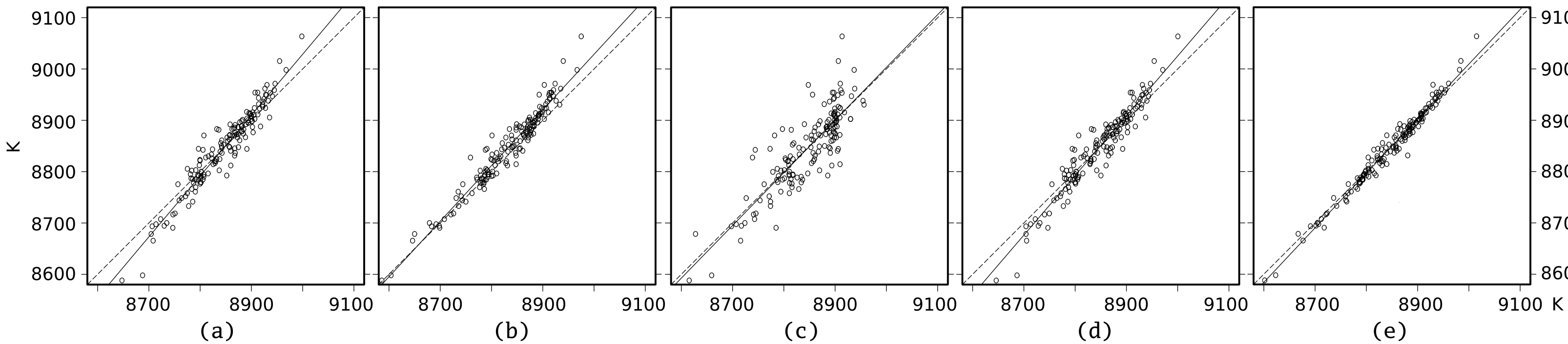

Fig. 19. Comparison between monthly accumulated maximum temperature and monthly accumulated maximum temperature derived from daily interpolated maximum temperature for June 1978 using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

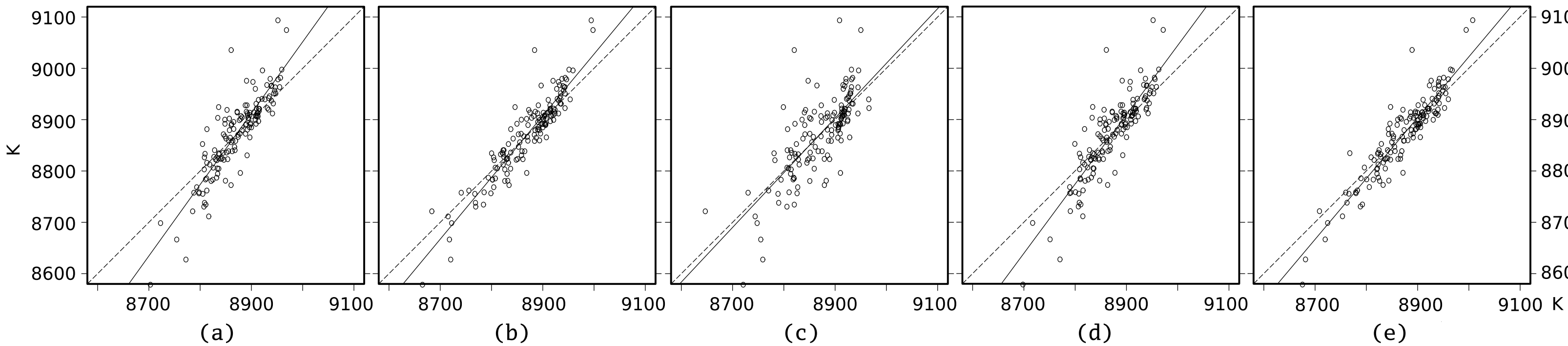

Fig. 20. Comparison between monthly accumulated maximum temperature and monthly accumulated maximum temperature derived from daily interpolated maximum temperature for June 1985 using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

4.3 *Spatial distribution of errors*

The previous section, which analyzed the relationship between errors and elevation raised the question of whether or not two climatological stations yielded large interpolation errors, which might be caused by the use of erroneous coordinates. In order to provide further insight into this question, the spatial location of the differences obtained with each Kriging method is analyzed in this section, as shown in Fig. 21 for June 1978 and Fig. 22 for June 1985. For June 1978 all Kriging variants yield a large error (>100 mm) in particular for two locations (as expected from the previous analysis): one in the Sierra Nevada and another one just north of it (Fig. 15 and Fig. 21) in fact all Kriging variants produced a value that was smaller than the measured value (light colored circles represent larger observed values).

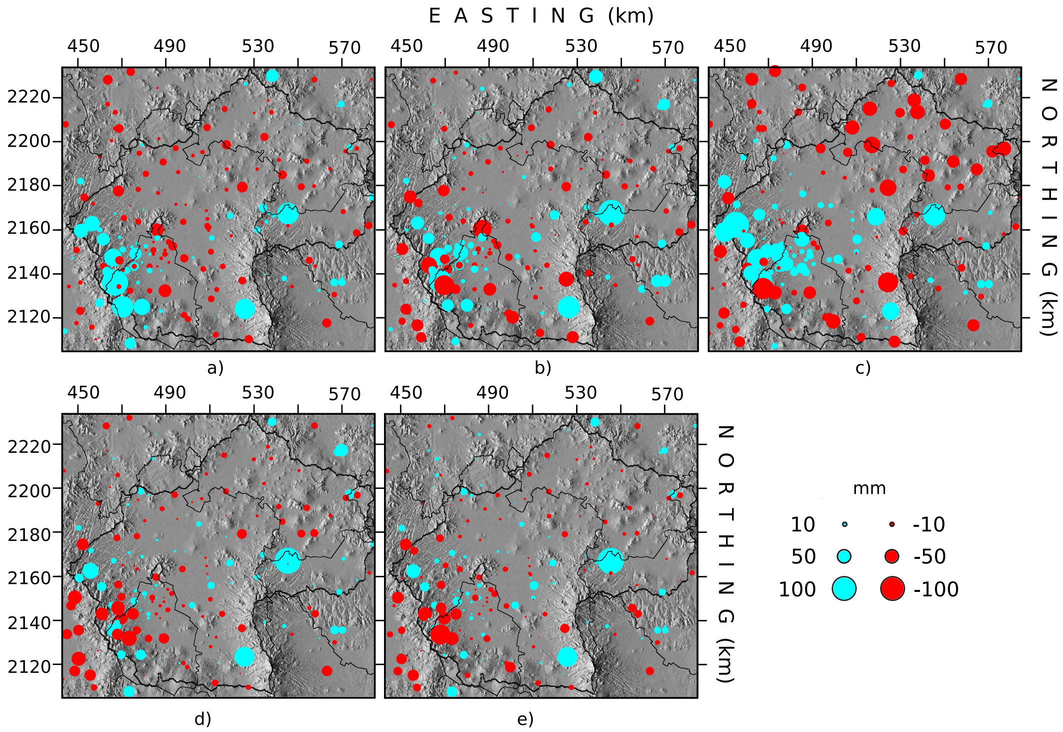

Fig. 21. Spatial distribution of difference between observed accumulated monthly rainfall and accumulated monthly rainfall derived from daily interpolations for June 1978 using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

The differences are smaller when elevation is used, except in BKED, which yielded larger errors throughout the study area (Fig 21c). It is interesting to

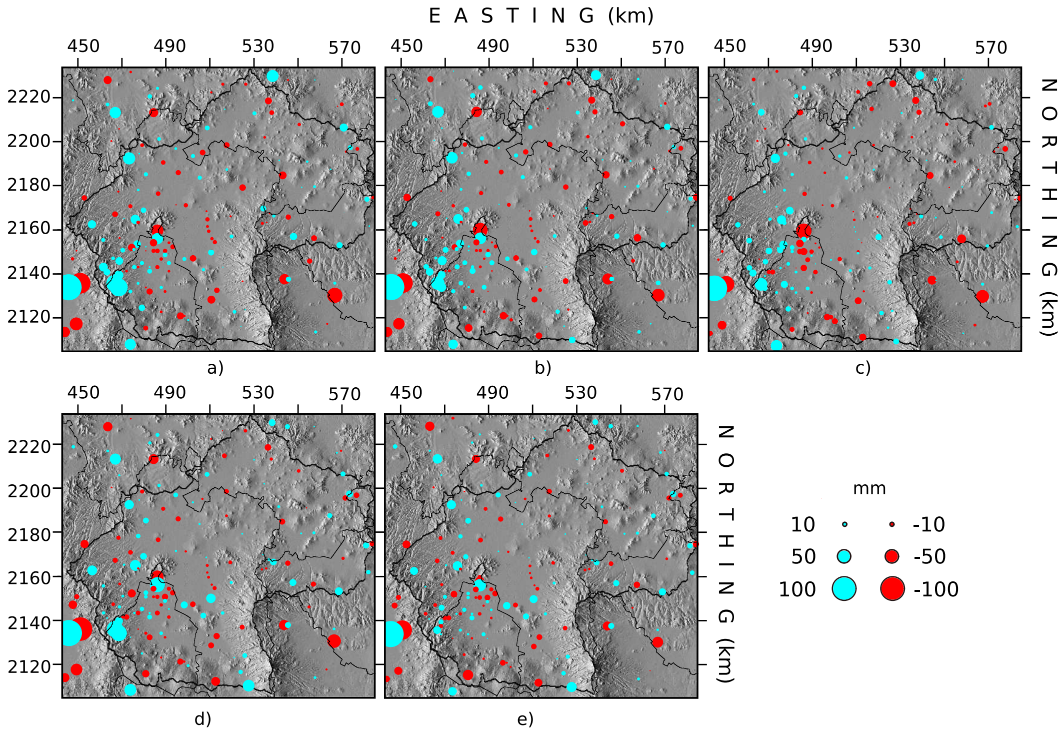

Fig. 22. Spatial distribution of difference between observed accumulated monthly rainfall and accumulated monthly rainfall derived from daily interpolations for June 1985 using: (a) Ordinary Kriging, (b) Kriging with External Drift, (c) Block Kriging using External Drift, (d) Ordinary Kriging in a local neighborhood and (e) Kriging with External Drift on a local neighborhood.

note that OK underpredicts rainfall in the Sierra las Cruces as well as KED; on the other hand, OK_l and KED_l overpredict monthly rainfall in this area. Analyzing the two points previously mentioned for June 1978, it can be noted that for the northernmost point, its measured monthly accumulated rainfall is 323 mm while the value recorded at its closest climatological station is 196 mm; for the remaining point, its measured rainfall depth was 318 mm while its nearest station recorded 151 mm. This large heterogeneity in such a small distance causes all Kriging variants to underestimate the rainfall value at those points as the fitted semivariogram can not represent this heterogeneity. Repeating this analysis for June 1985 the same phenomenom is found, as the rainfall depth at the eastern most point is 109 mm while at the remaining point this depth is equal to 501 mm; accordingly all Kriging variants overestimate precipitation at the first point and underestimate it at the second station. This difference in precipitation values is very large, while the distance between them is 1.3 km for June 1978 and 5.5 km for June 1985; furthermore, this variation is not accounted for on the semivariogram model.

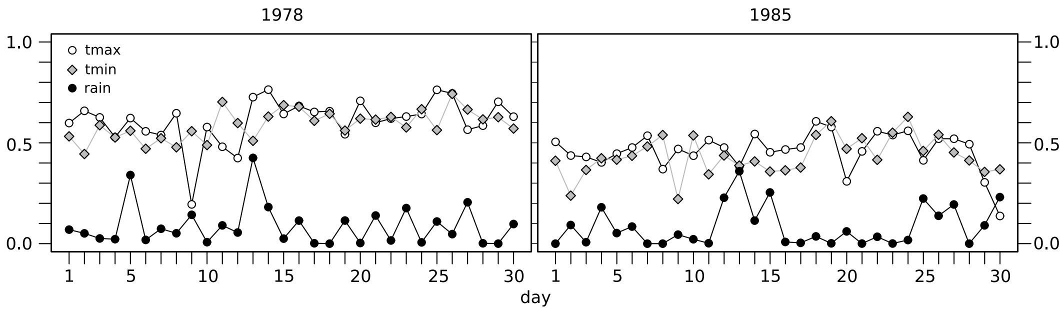

Fig. 23. Correlation of daily climatological variables with elevation in the southwestern area of the Basin for (a)June 1978 and (b) June 1985

4.4 Local neighborhoods

The previous sections discussed the effect of the different Kriging method and the difference between using a global or local neighborhood will be further discussed in this section by using the statistics of the global and local neighborhood, in the southwestern area of the Basin indicated by the dashed rectangle of Fig. 1.

The correlation between rainfall and elevation varies greately on some days, e.g. June 5, 1978 changed from a global correlation of 0.04 (Fig. 3(a)) to a local one of 0.34 (Fig. 23(a)) while the correlation for June 13, 1985 changed from 0.25 (Fig. 3(b)) to 0.43 (Fig. 23(a)). For the four different days discussed in the previous sections of this work, the difference between correlations was not as drastic as for June 8, 1978 the correlation was the same (0.06 for global and 0.05 for local) while for June 23 1978 it changed from 0.13 to 0.18. For 1985, the correlation observed on June 3, decreased from 0.03 to 0.01 although it increased in June 15 from 0.13 to 0.25. The effect of these values is noticeable in the interpolated maps for rainfall as the day which exhibited the maximum local correlation in the southwestern area shows a large difference in the rainfall values when elevation is used (Figs. 6(a) and 10(a)).

The importance of using a local neighborhood is made evident by comparing the statistics between the global (Fig.4) and local (Fig. 24) neighborhoods For June 8, 1978 the maximum value is located in the southwestern area of the Basin and the mean value of rainfall for this day changes from around 11 mm to near 22 mm when using a global or a local neighborhood respectively. The same pattern is observed for June 23 1978 and June 15 1985 as the maximum value of rainfall is also located in this area and its mean value is also increased when a local neighborhood is used. For June 3, 1985 the maximum value of rainfall is not located in the southwestern subarea. When a global neighborhood is used this value is 50 mm while on the southwestern area the mean value is near 46 mm, an effect that is observed in the interpolated distribution of rainfall for this day by KED_l (Fig. 10).

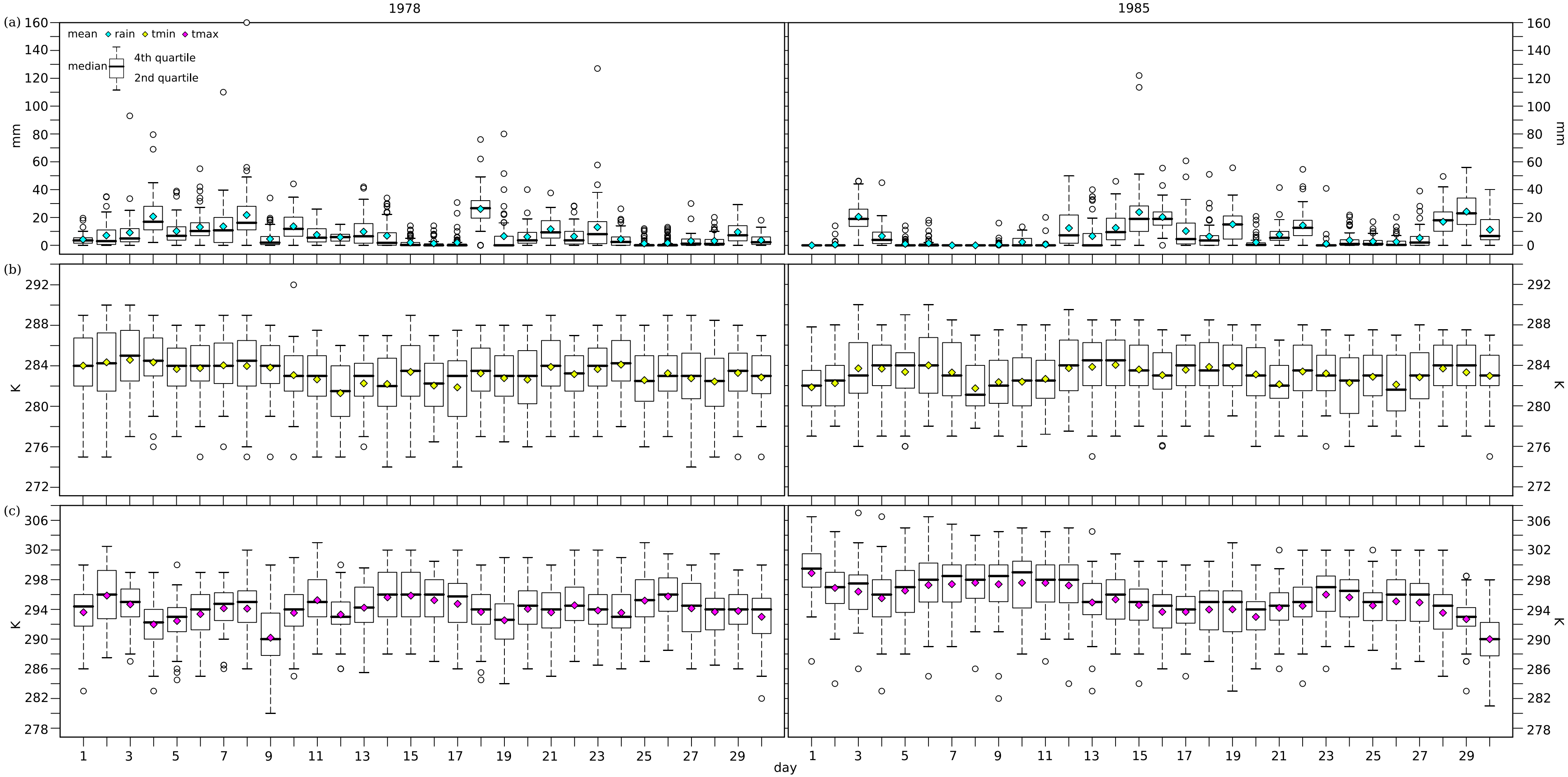

Fig. 24. Whisker plots of daily climatological variables for June 1978 and 1985 in the southwestern area of the Basin: (a) rainfall, (b) minimum temperature and (c) maximum temperature

In the case of temperature, the relationship between elevation and both temperatures increases when only the southwestern area is considered for the four days considered in this work, as well as for the remainder days of both months (Fig. 4) and local (Fig.24). This is to be expected as the soutwestern subarea is predominantely mountainous (Fig. 1). The mean, maximum and minimum values for both temperature fields do not change as drastically as those for rainfall which also should be expected as temperature is a continous field through the study area, instead of a localized one, as rainfall is. Accordingly, the spatial distribution obtained with KED and KED_l is similar (as well as the errors' statistics) except in those aereas which have a low density of climatological stations. As explained in section 2, Kriging uses the mean of a selected neighborhood thus in order to represent a heterogeneous field such a rainfall, Kriging with External Drift in a local neighborhood is the method that provides the best results as it accounts for the local relationship between rainfall and elevation, as well as the way in which the mean changes with a local neighborhood. Based on the errors' statistics, along with the spatial distribution obtained by each interpolation method, for this study KED_l is selected to interpolate daily rainfall, while KED is selected to interpolate both minimum and maximum temperature.

5 Conclusions

The effect of considering elevation as a secondary variable to interpolate daily rainfall, minimum and maximum temperature was analyzed in this study using approximately 200 climatological stations for an area of 16,800 km^2. In this study, the use of elevation as a secondary variable improved the spatial variation of all climatological fields even when they exhibited low correlation with elevation. According to the analyses presented, the use of Kriging with External Drift on a local neighborhood (KED_l) is recommended to undertake the spatial interpolation of rainfall, (a highly heterogeneous variable) while Kriging with External Drift (KED) is recommended to undertake the interpolation of a more continuous field such as minimum and maximum temperature. In the case of a more dense climatological network, KED_l would probably improve the spatial pattern of minimum and maximum temperature. The selection of a local neighborhood is an interactive process, in order to provide continuous fields which capture the local behaviour of the variable of interest. To summarize:

(1) Anisotropy does not need to be considered when using a local neighborhood, as observed in the interpolated maps.
(2) The RMSE values should not be used alone in order to decide whether an interpolation method yields the best interpolation. Other issues need to be considered, such as the density and location of measurement points.
(3) The ability to undertake spatial interpolation is an add-on capacity of many Geographic Information Systems; however, the capacity of the end-user to undertake this type of interpolations is presented as a black-box exercise. Knowledge about the principles behind each interpolation method is required from the end-users of interpolation software in order to produce reliable spatial patterns of the variable under study.

Acknowledgements

Financial support by the Mexican Council for Science and Technology (CONACyT) and the Natural Sciences and Engineering Research Council of Canada (NSERC) are acknowledged. We would also like to thank Edzer Pebesma for reading a draft version of this paper and providing valuable feedback.

References

Allard, D. (1998). Geostatistical classification and class Kriging. *J. of Geographic Information and Decision Analysis*, 2:87–101.

Atkinson, P. M. and Lloyd, C. D. (1998). Mapping precipitation in Switzerland with Ordinary and Indicator Kriging. *J. of Geographic Information and Decision Analysis*, 2:65–76.

Bivand, R. S. (2000). Using the R statistical data analysis language on GRASS 5.0 GIS database files. *Computers & Geosciences*, 26:1043–1052.

Carrera-Hernández, J. J. and Gaskin, S. J. (2005). The valley of mexico hidrogeological database (VMHDB): Implementation and basic queries. *Hydrogeology journal*. In review.

Demyanov, V., Kanevski, M., Chernov, S., Savelieva, E., and Timonin, V. (1998). Neural network residual kriging application for climatic data. *J. of Geographic Information and Decision Analysis*, 2:34–58.

Deutsch, C. V. and Journel, A. G. (1998). *GSLIB: Geostatistical software library and user's guide*. Applied Geostatistics. Oxford University Press, New York, 2nd edition.

Dingman, L. S., Seely-Reynolds, D. M., and Reynolds, R. C. (1988). Application of kriging to estimate mean annual precipitation in a region of orographic influence. *Wat. Res. Bulletin*, 24:329–339.

Dubois, G. (1998). Spatial interpolation comparison 97: Foreword and introduction. *J. of Geographic Information and Decision Analysis*, 2:1–10.

Goovaerts, P. (1997). *Geostatistics for Natural Resources evaluation*. Oxford University Press, New York.

Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. of Hydrology*, 228:113–129.

Hofierka, J., Parajka, J., Mitasova, H., and Mitas, L. (2002). Multivariate interpolation of precipitation using regularized spline wiht tension. *Transactions in GIS*, 6(2):135–150.

Hudson, G. and Wackernagel, H. (1994). Mapping temperature using kriging with external drift: Theory and and example from Scotland. *Int. Journal of Climatology*, 14:77–91.

Hutchinson, M. F. and Galland, J. C. (1999). Representation of terrain. In Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W., editors, *Geographical Information Systems*, pages 105–124. John Wiley and Sons.

Isaaks, E. H. and Srivastava, R. M. (1989). *An Introduction to Applied Geostatistics*. Oxford University Press.

Jarvis, C. H. and Stuart, N. (2001). A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part II: The interaction between number of guiding variables and the type of interpolation method. *J. of Applied Meteorology*, 40:1075–1084.

Kyriakidis, P. C., Kim, J., and Miller, N. L. (2001). Geostatistical mapping of precipitation from rain gauge data using atmospheric and terrain characteristics. *J. of Applied Meteorology*, 40:1855–1877.

Lloyd, C. D. (2005). Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. *J. of Hydrology*, 308:128–150.

Martinez-Cob, A. (1996). Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. *J. of Hydrology*, 174:19–35.

Mitasova, H. and Mitas, L. (1993). Interpolation by regularized spline with tension: I. theory and implementation. *Math. Geol.*, 25:641–655.

Pebesma, E. J. (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, 30:683–691.

Pebesma, E. J. and Wesseling, C. G. (1998). GSTAT: a program for geostatistical modelling, prediction and simulation. *Computers & Geosciences*, 24(1):17–31.

Phillips, D. L., Dolph, J., and Marks, D. (1992). A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. *Agricultural and Forest Meteorology*, 58:119–141.

R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Saveliev, A. A., Mucharamova, S. S., and Piliugin, G. A. (1998). Modeling of the daily rainfall values using surface under tension and Kriging. *J. of Geographic Information and Decision Analysis*, 2:62–71.

Tabios, G. Q. and Salas, J. D. (1985). A comparative analysis of techniques for spatial interpolation of precipitation. *Water Resources Bulletin*, 21(3):365–380.

Thieken, A. H. (1998). Estimating daily regional rainfall fields by multiqadric functions: Accuracy of interpolation and decision making. *J. of Geographic Infomration and Decision Analysis*, 2:86–99.

Tomczak, M. (1998). Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW) - cross validation/jacknife approach. *J. of Geographic Information and Decision Analysis*, 2:18–30.

Wackernagel, H. (2003). *Multivariate Geostatistics: An introduction with applications*. Springer.