Theor Appl Climatol (2016) 123:335–348
DOI 10.1007/s00704-014-1359-9

ORIGINAL PAPER

Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil

Pablo de Amorim Borges · Johannes Franke · Yumiko Marina Tanaka da Anunciação · Holger Weiss · Christian Bernhofer

Received: 12 May 2014 / Accepted: 20 December 2014 / Published online: 9 January 2015

Abstract Available climatological information of Distrito Federal does not satisfy the requirements for detailed climate diagnosis, as they do not provide the necessary spatial resolution for water resources management purposes. Annual and seasonal climatology (1971–2000) of precipitation from 6 meteorological stations and 54 rain gauges from Central Brazil were used to test eight different spatial interpolation methods. Geographical factors (i.e., altitude, longitude and latitude) explain a large portion of precipitation in the region, and therefore, multivariate models were included. The performance of estimations was assessed through independent validation using mean square error, correlation coefficient and Nash–Sutcliffe efficiency criterion. Inverse distance weighting (IDW), ordinary kriging (OK) and the multivariate regression with interpolation of residuals by IDW (MRegIDW) and OK (MRegOK) have performed the lowest errors and the highest correlation and Nash–Sutcliffe efficiency criterion. In general, interpolation methods provide similar spatial distributions of rainfall wherever observation network is dense. However, the inclusion of geographical variables to the interpolation method should improve estimates in areas where the observation network density is low. Nevertheless, the assessment of uncertainties using a geostatistical method provides supplementary and qualitative information which should be considered when interpreting the spatial distribution of rainfall.

P. d. A. Borges (✉) · J. Franke · C. Bernhofer
Dresden University of Technology—TUD, 01737 Tharandt, Germany
e-mail: pablo.amorim@mailbox.tu-dresden.de

P. d. A. Borges
e-mail: pablo.borges@ufz.de

Y. M. T. da Anunciação
Coordenação Geral de Desenvolvimento e Pesquisa—CDP, Instituto Nacional de Meteorologia—INMET, 70680-900 Brasília, DF, Brazil

H. Weiss
Department Groundwater Remediation, Helmholtz Centre for Environmental Research GmbH—UFZ, 04318 Leipzig, Germany

1 Introduction

Spatial distributed climate data is essential information to many questions addressing water resources (Dingman et al. 1988; Phillips et al. 1992; Martínez-Cob 1996; Lanza et al. 2001; Tveito et al. 2008; WMO 2008; Wagner et al. 2012). A challenging task for climatologists is to provide information about climate for any place at any time (Tveito et al. 2008). Most climatological parameters are, in the long-term, traditionally measured at point locations, such as meteorological stations and rain gauges. An accurate estimation of the spatial distribution of these parameters requires a very dense network of instruments as well as remote sensing methods (e.g. precipitation-radar) and process-oriented simulation, requiring large installation and operational costs (Frei and Schär 1998; Goovaerts 2000; Mair and Fares 2010). Spatial distributed data from radar and satellites have a great potential of application in hydrological studies. However, both require validation and corrections. The first using rain gauges (Anagnostou et al. 1999), while the second using radar data as reference (Habib and Krajewski 2002). Nevertheless, the available period of record frequently does not meet the needs of end users. For instance, radar data in Central Brazil is available from 2012 (CPTEC-INPE 2014) while the Tropical Rainfall Measuring Mission (TRMM, Simpson et al. 1988) data starts in 1997.

As an alternative, spatial distribution of climate variables can be estimated by applying interpolation methods from surrounding point stations. Spatial interpolation is a method or mathematical function that estimates the values at locations where no measured values are available (Lanza et al. 2001; Tveito et al. 2008; Di Piazza et al. 2011). There are many methods for interpolation of climatological information; however, the choice of the technique depends on the aim of the study, the climatological variable in concern, time scale, the spatial resolution wanted, and territorial context of the model region, such as network density, topography, etc. (Wackernagel 2003; Renard and Comby 2006; Tveito et al. 2008). Traditional methods are based on distance criteria, such as the Thiessen polygons, which correspond to defined homogeneous areas where the climate variable is assumed to be constant (Thiessen 1911). More sophisticated methods, such as artificial neural networks, have been applied to estimate the non-linear spatial variability of climatic variables (Demyanov et al. 1998; Di Piazza et al. 2011). Nowadays, geographical information systems (GIS) gives a variety of possibilities for the integration, analysis and visualization of climatological data (Dobesch et al. 2007). Geostatistical tools of GIS enhanced the capacity of deriving detailed spatial representations of climatological data. Basically, the spatial correlation between neighbouring observations is taken into account in order to estimate values at unsampled locations (Tabios and Salas 1985; Phillips et al. 1992). In general, geostatistical methods perform better than traditional methods, such as Thiessen polygons and inverse distance weighting (IDW) (Chua and Bras 1982; Tabios and Salas 1985; Pardo-Igúzquiza 1998; Goovaerts 2000; Apaydin et al. 2004; Coulibaly and Becker 2007; Haberlandt 2007; Gan et al. 2010; Wagner et al. 2012). Moreover, the inclusion of explanatory variables, such as elevation, can potentially increase the reliability of rainfall estimates, especially in areas with low network density (Tabios and Salas 1985; Phillips et al. 1992). Several authors (e.g., Goovaerts 2000; Haberlandt 2007; Portalés et al. 2010; Di Piazza et al. 2011; Wagner et al. 2012; Delbari et al. 2013) demonstrate the outperformance of multivariate methods compared to univariate techniques.

Software packages have been specifically developed for climatology, such as PRISM (Parameter Regression on Independent Slopes Model, Daly et al. 1994) and MISH (meteorological interpolation based on surface homogenized data basis, Szentimrey et al. 2005). The first method includes explanatory variables such as terrain orientations, shape and distance to the coast. It seems to be a potential alternative in regions where the station network is unrepresentative for the variation in topography (Tveito et al. 2008). The MISH package combines deterministic and stochastic models by using a set of terrain characteristics together with several meteorological information such as remote sensing data. More recent studies have also incorporated information from radar and TRMM data to interpolation methods (e.g. Haberlandt 2007; Wagner et al. 2012); however, those methods require a massive amount of data which is often not available for users.

Essentially, precipitation in Central Brazil is influenced by large-scale atmospheric circulation systems from the Amazon region (Alves 2009). Intensive rainfall, which occurs between November and March, is carried from the Equatorial Atlantic region by low-level east winds when deviated from northwest towards southeast due to the Andes barrier (Virji 1981; Gan et al. 2004; Vera et al. 2006). The inclusion of variables representing the drivers of precipitation has a potential to increase the estimation power of interpolation schemes. Several studies demonstrate the correlation of precipitation with geographical factors, such as height above sea level, longitude and latitude (e.g. Price et al. 2000; Apaydin et al. 2004; Lhotellier 2005; Dobesch et al. 2007; Tveito et al. 2008; Portalés et al. 2010; Gan et al. 2010). Longitude and latitude may represent the influence of the atmospheric circulation systems coming from the west and northwest (i.e. Amazon basin). Although the study area is characterized by non-complex terrain, the elevation ranges from ca. 300 to 1500 m above sea level (Fig. 1). Under these topographic conditions, orographic precipitation may also contribute to the spatial distribution of rainfall (Alter 1919; Barrows 1933; Spreen 1947; Smith 1979; Phillips et al. 1992; Daly et al. 1994; Frei and Schär 1998).

Facing the urgency to take actions that will guarantee the water supply of Distrito Federal (DF), the project called IWAS/Água-DF (International Water Research Alliance Saxony) intends to develop an Integrated Water Resources Management (IWRM) system wherein climate is comprised as natural boundary condition (Lorz et al. 2012). Available climate studies in Central Brazil are restricted and do not fulfil the requirements for detailed climate diagnosis (Borges et al. 2014). Ramos (2009) developed a Brazilian climatology atlas (i.e. 1961–1990) for several climatological variables. However, the number of stations used is limited to those under the auspices of the National Institute of Meteorology (INMET), and the spatial resolution provided does not meet the requirements for water resource management in Distrito Federal. This study aims to identify the interpolation method(s) that performs the most realistic spatial distribution of seasonal and annual precipitations for the climatological period 1971–2000 in Distrito Federal. For that, we evaluate the performance of deterministic, probabilistic, and combined methods through independent validation. Visual analysis and statistical measures are applied as criterion to identify the finest estimates.

2 Database

The data used in this study comprises daily observations provided by the Brazilian Hydrological Information System (HIDROWEB, http://hidroweb.ana.gov.br/) and INMET. Additional datasets were obtained from Brazilian Enterprise

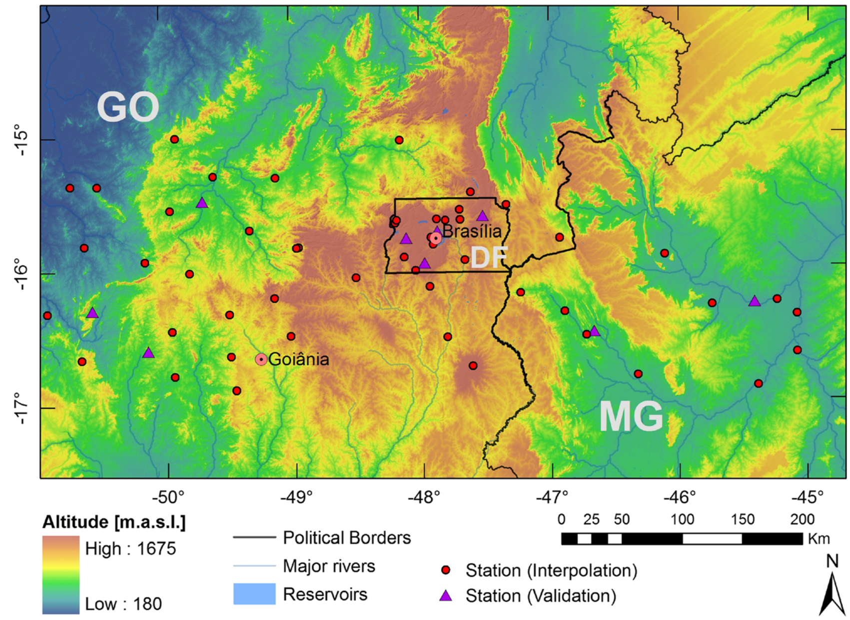

Fig. 1 Distribution of the station network and topography of the model area of Central Brazil which includes parts of the federal state of Goiás (GO) and Minas Gerais (MG), and the entire Distrito Federal (DF)

for Agricultural Research (EMBRAPA) and the regional water supplier, Environmental Sanitation Company of the Distrito Federal (CAESB). In order to avoid incorrect estimates along the political border of Distrito Federal, the databank also includes surrounding stations. This study considers time series of 6 weather stations and 54 rain gauges located in the Central Brazil area (longitude: −44° to −51°; latitude: −14° to −18°) (Fig. 1). As demonstrated in a previous study (Borges et al. 2014), verification of suspicious values, errors and outliers were performed after (Dixon 1950). Homogeneity was tested using graphical (i.e. Craddock test, double sum analysis, quotient criteria and difference in limits) and numerical techniques (i.e. Abbe, Buishand and Alexandersson tests). Topographic data comprises the 30× 30 m resolution ASTER Global Digital Elevation Model (DEM) V001 (http://asterweb.jpl.nasa.gov/gdem.asp).

3 Methods

Climatological normals are essential information to classify a region's climate (Nalder and Wein 1998; Perčec Tadić 2010). Normals are widely applied as an indicator of the climate conditions likely to be experienced in a given region. Additionally, they can be used as a benchmark to be compared against certain condition (e.g., current conditions or projections) (WMO 1996). WMO (1989) establishes general procedures to calculate monthly and annual 30-year standard normals for climate data. The standard period suggested is 1961–1990; however, most of the observations in Central Brazil were only initiated towards the end of sixties. In order to consider a reasonable observation network density, we use the 1971–2000's period for the calculation of climatological normals. A large variety of interpolation methods are available for climatology (Dobesch et al. 2007; Tveito et al. 2008). Before selecting the proper one, it is necessary to consider the purpose of the interpolation, the characteristics of the phenomena and the assumptions and limitations of the technique (Nalder and Wein 1998; Goovaerts 2000; Tveito et al. 2008). The complexity of the appropriated spatial interpolation model is a function of the time scale and spatial resolution wanted. As time aggregation and station density decreases, uncertainties associated with predictions increase considerably, demanding therefore more complex models (Tveito et al. 2008). The network density for the whole study area is ca. one station per 5000 km^2, while one station for ca. 400 km^2 in Distrito Federal. This study focuses on the spatial representation of seasonal and annual long-term means (1971–2000) of accumulated rainfall. At this level of time aggregation, the spatial variance of precipitation is likely to decrease (Nalder and Wein 1998; Tveito et al. 2008; WMO 2008). In order to guarantee a plausible model complexity and to satisfy the needs for water resources studies, we apply the DEM of 1-km grid resolution (Borga and Vizzaccaro 1997; Agnew and Palutikof 2000; Goovaerts 2000; Jarvis and Stuart 2001; Hong et al. 2005; Perčec Tadić 2010; Bárdossy and Pegram 2013). Several categories of interpolation methods are available, and they are classified according to their fundamental mathematics. The most frequent applications for establishing spatial representations of precipitation rely on the principle of ordinary kriging (OK) or inverse distance weighting—IDW (Phillips et al. 1992; Li and Heap 2011). Current studies (e.g., Phillips et al. 1992; Nalder and Wein 1998; Goovaerts 2000; Ninyerola et al. 2000; Apaydin et al. 2004; Hong et al.

2005; Ninyerola et al. 2007; Gan et al. 2010; Di Piazza et al. 2011; Bostan et al. 2012) make use of terrain characteristics, such as elevation, geographic position, land use and water bodies, to describe spatial representations of precipitation. Several authors (e.g. Martínez-Cob 1996; Nalder and Wein 1998; Goovaerts 2000; Ninyerola et al. 2000; Lhotellier 2005; Portalés et al. 2010; Di Piazza et al. 2011) demonstrated the reliability of regression models combined with residual interpolation, also referred as "detrended" interpolation. In order to identify the most appropriated interpolation method in estimating precipitation at season and annual scale in Distrito Federal, eight different methods are evaluated.

3.1 Interpolation methods

3.1.1 IDW

A type of deterministic method widely applied in spatial modelling. The estimation is based upon weighted averages, which are proportional to the inverse of the distance between the interpolated and measured points (Shepard 1968). The general formula is expressed as

$$\widehat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i) \tag{1}$$

where $\hat{Z}(s_0)$ is the estimate value of Z in s_0, $Z(s_i)$ is the measured value located in s_i, λ_i is the weight of $Z(s_i)$ and n is the number of measurements used for the estimate. And the weights are calculated as follows:

$$\lambda_i = \left(\frac{1}{d_i^P}\right) \Big/ \left(\Sigma_{i=1}^{n} \frac{1}{d_i^P}\right) \tag{2}$$

where d_i is the distance between s_0 and s_i, P is a power parameter and n represents the number of sampled points used for the estimation. The main factor affecting the accuracy of IDW is the value of the power parameter. As the distance increases, weights diminish, especially when the value of the power parameter rises. Nearby stations have a heavier weight and, therefore, more influence on the estimation (Isaaks 1989; Nalder and Wein 1998; Johnston 2004). The optimal power value (i.e. 1.5) was determined by minimizing the prediction error calculated from the cross-validation procedure available in the Geostatistical Analyst tool of ArcGIS®9.

3.1.2 Spline tension (Spline)

Spline is classified as a deterministic method and belongs to the radial basis functions (RBF) group. RBF encloses a series of exact interpolation techniques, where the spatial predictions must go through each measured sample value (Hutchinson and Gessler 1994; Apaydin et al. 2004; Johnston 2004; Di Piazza et al. 2011). The predictor is a linear combination of functions,

$$\widehat{Z}(s_0) = \sum_{i=1}^{n} \omega_i \varphi(\|s_i - s_0\|) + \omega_{n+1} \tag{3}$$

where $\varphi(r)$ is a radial basis function (see formula 4), $r=||s_i-s_0||$ is Euclidean distance between the prediction location s_0 and each data location s_i, and $\{\omega_i: i=1, 2, \ldots, n+1\}$ are weights to be estimated. As opposed to IDW, Spline methods can predict values above the maximum and below the minimum measured values. Spline is normally used for calculating smooth spatial distributions from a large number of data points. The technique may not be appropriate when there are large changes in observation values within a short horizontal distance (Hutchinson and Gessler 1994; Johnston 2004). In a previous task, all spline methods available in the Geostatistical Analyst tool of ArcGIS®9 were tested. From that, spline tension with smoothing parameter 2.0×10^{-4} revealed a slightly better performance than the other spline methods, and its function is described as

$$\varphi(r) = \ln\left(\sigma \cdot r \big/ 2\right) + K_0(\sigma \cdot r) + C_{E'} \tag{4}$$

where σ is the smoothing parameter, $K_0(x)$ is the modified Bessel function, and $C_{E'}$ is the Euler constant.

3.1.3 OK

The kriging family, also known as geoestatistical methods, is based on the idea that values measured at near locations tend to be statistically more related than values measured at other locations (Tveito et al. 2008). Similar to IDW, kriging also uses a weighting; however, as a probabilistic method, kriging depends on spatial and statistical relationships to predict unmeasured points. The empirical semivariogram provide means for assessing the spatial autocorrelation of datasets (Tabios and Salas 1985; Phillips et al. 1992; Goovaerts 2000; Dobesch et al. 2007). Pairs that are closer are expected to differ less than those distant from one another. To model the spatial autocorrelation of the data, statistical functions are tested to best fit the semivariogram. A kriging equation is based on spatial optimal linear prediction, where the unknown mean of the random process is estimated through the best linear unbiased estimator (B.L.U.E.). The estimator is "unbiased" because the mean of error is 0; "linear" since the estimated values are weighted linear combinations of the available data;

and "best" because the estimator aims to minimize the variance of the errors (Goldberger 1962; Cressie 1990). The ordinary kriging performed the lowest errors among the other kriging methods (e.g., simple and disjunctive kriging) in a previous cross-validation procedure. The prediction of ordinary kriging is a linear combination of measured values where the constant mean (μ) is assumed to be unknown (Isaaks 1989; Cressie 1990; Dobesch et al. 2007). The ordinary kriging representation for spatial stochastic process Z(s) is

$$Z(s) = \mu + \delta(s) \tag{5}$$

where μ is unknown expected value of random process, $\delta(s)$ is a zero mean intrinsically stationary random process with existing semivariogram $\gamma(H)$. The estimator $\hat{Z}(s_0)$ can be expressed as the formula *(1)*. Linear coefficients λ_i, $i=1,\ldots,n$ are calculated under the condition for uniformly unbiased predictor as

$$E\left(\hat{Z}(s_0)\right) = E(Z(s_0)), \quad \sum_{i=1}^{n} \lambda_i = 1 \tag{6}$$

and under the restriction of minimal prediction error variance $\sigma^2(s_0)$ at location s_0 as

$$\sigma^2(s_0) = E\left(Z(s_0) - \widehat{Z}(s_0)\right)^2 \tag{7}$$

The Geostatistical Analyst tool of ArcGIS®9 offers a range of functions to fit the empirical semivariogram. Among them, the spherical function shows the lowest estimation errors and a linear behaviour at the origin which is plausible to the semivariogram (Fig. 2). The spherical

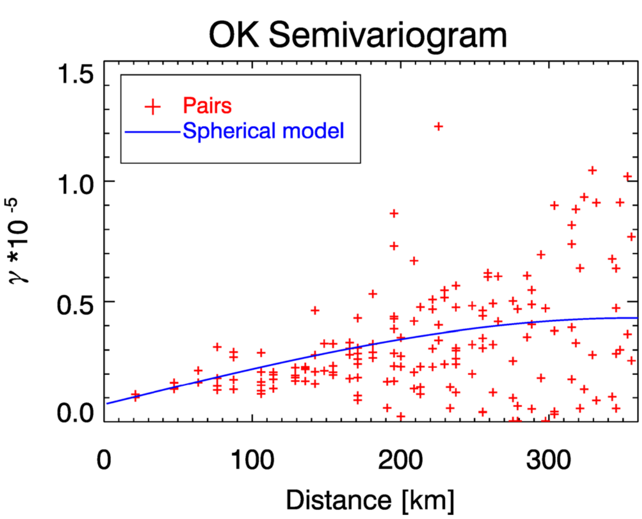

Fig. 2 The semivariogram and optimal fitted model (i.e. spherical) of the ordinary kriging function for 30-year (1971–2000) average of annual precipitation in Central Brazil

model is one of the most commonly used models (Johnston 2004) and it is written as

$$\gamma(H) = \theta_0 + \theta_s \cdot \left[1.5 \cdot \frac{h}{\theta_r} - 0.5 \cdot \left(\frac{h}{\theta_r}\right)^3\right] \text{ for } 0 \leq h \leq \theta_r \tag{8}$$

$$\gamma(H) = \theta_s \text{ for } \theta_r < h \tag{9}$$

where H is the lag vector, h is the length of H (distance between two points), θ_0 is the nugget, $\theta_s \geq 0$ is the partial sill value and θ_r is the range of the model. As suggested by Johnston (2004), the lag size is the average distance between neighbouring samples (i.e. 30 km) and the number of lags is set as 12. Moreover, at an infinitesimally distance, the semivariogram exhibits a nugget effect. The nugget effect can be attributed to measurement errors at distances smaller than the data interval (Goovaerts 2000). The spherical model presents a nugget effect of ca. 0.07×10^{-5} and partial sill (sill minus nugget) of 0.36×10^{-5}.

3.1.4 CoOK

CoKriging (CoOK) is an extension of kriging where more than one auxiliary variable can be added to the prediction scheme (Isaaks 1989; Stein et al. 1991; Phillips et al. 1992; Nalder and Wein 1998). In a conventional kriging model, a response is assumed to be a spatial random process with stationary covariance function, which implies that the smoothness of a response is fairly uniform in each region of the domain area (Goovaerts 1997). However, cases are common where the level of smoothness of a response could change considerably due to geographical characteristics. In such situations, CoKriging is a regularly used technique wherewith interpolations are improved by adding secondary attributes that may drive the spatial distribution of the variable concerned, for instance longitude, latitude and elevation (Stein et al. 1991). CoKriging is most effective when the covariates are highly correlated (Nalder and Wein 1998). Here, we apply elevation as the auxiliary variable. According to Cressie (1993), supposing the data as $k \times 1$ vectors (variables) measured on n locations, the multivariate process can be written with the $n \times k$ matrix:

$$Z_j = \left(Z_j(s_1), \ldots, Z_j(s_n)\right)' \tag{10}$$

With (i,j)th element $Z_j(s_i)$ of $j=1,\ldots,k$ variables in $i=1,\ldots,n$ locations. The aim is to predict the $Z_1(s_0)$ based not only on $Z_1=(Z_1(s_1),\ldots,Z_1(s_n))'$ but also on the covariables (Eq. 10 with $j \neq 1$). The same assumptions as by ordinary kriging are expected for each of the k variables. The kriging predictor of

$Z_1(s_0)$ is a linear combination of all available data values of all k variables:

$$\widehat{Z}_1(s_0) = \sum_{i=1}^{n} \sum_{j=1}^{k} \lambda_{ij} Z_j(s_i) \tag{11}$$

Assuming a uniformly unbiased predictor with the conditions:

$$\sum_{i=1}^{n} \lambda_{1i} = 1 \ , \quad \sum_{i=1}^{n} \lambda_{ji} = 0 \quad j = 2 \ , \ \ldots \ , \ \mathrm{k} \tag{12}$$

Therefore, the best linear unbiased estimator is obtained by minimizing

$$E\left(Z_1(s_0) - \sum_{i=1}^{n} \sum_{j=1}^{k} \lambda_{ij} Z_j(s_i) \right)^2 \tag{13}$$

The spherical function (see Eq. 8) was applied for all variable pairs of interest (autocorrelations and cross-correlation). The model parameters of the semivariogram of precipitation pairs are the same as OK. The spherical function fitting the semivariogram of altitude pairs does not present a nugget effect and the partial sill is 1.00×10^{-5}. The model of the semivariogram between precipitation and altitude pairs does not present a nugget effect and the partial sill is ca. 0.32×10^{-5}. The lag size and number of lags are the same as for OK. The Fig. 3 illustrates the assumptions concerning to the CoOK model. The empirical semivariogram of altitude pairs is well fitted by a spherical model (Fig. 3a). Figure 3b shows the empirical cross-covariance for all pairs of locations between precipitation and altitude as well as the fitted function. In general, the spatial covariance between precipitation and altitude is low at the origin and tend to increase with the distance.

3.1.5 DUK

The detrended universal kriging (DUK) is a generalized case of kriging where the trend is modelled as a function of the coordinates. The method can add substantial value to estimates in case of a spatial trend, which might be true for rainfall in Central Brazil (see Fig. 5c, d). The deterministic component of the plane is represented by a mathematical formula while the random process is estimated by the function fitting the semivariogram of the "detrended" data (Tveito et al. 2008). In fact, the deterministic component uses the spatial coordinates as the explanatory variables whereas the residuals are modelled assuming an autocorrelation function (Johnston 2004). The DUK assumes the model

$$Z(s) = \mu(s) + \varepsilon(s) \tag{14}$$

where μ(s) is a deterministic function and $\varepsilon(s)$ are the errors which are assumed to be random. In this case study, a first-order polynomial trend was detected and, therefore, assumed as the deterministic function. The empirical semivariogram and fitted function (i.e. spherical) of the "detrended" data is illustrated in Fig. 4. The lag size and number of lags are the same as for OK. The nugget effect is 0.05×10^{-5} and partial sill 0.11×10^{-5}.

3.1.6 MReg

Multiple linear regression (MReg) is a deterministic method which expresses the relation between a predicted variable and explanatory variables (Tveito et al. 2008). In a linear manner, we assume that the spatial distribution of rainfall in Distrito Federal is dependent on location (i.e. longitude and latitude) and terrain elevation. Figure 5 illustrates the spatial

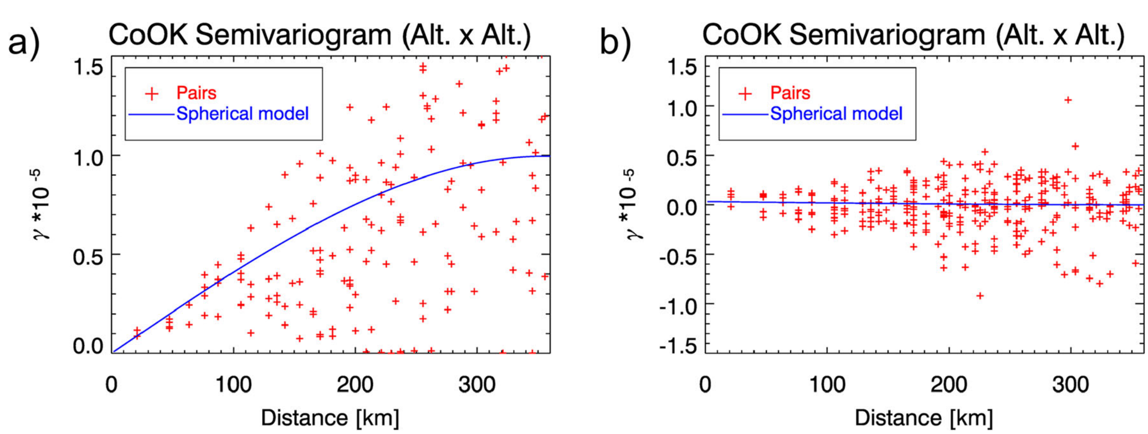

Fig. 3 The CoOK assumptions. **a** The empirical semivariogram and fitted spherical function of altitude pairs and **b** the empirical cross-covariance and fitted spherical function between 30-year average of annual precipitation and altitude

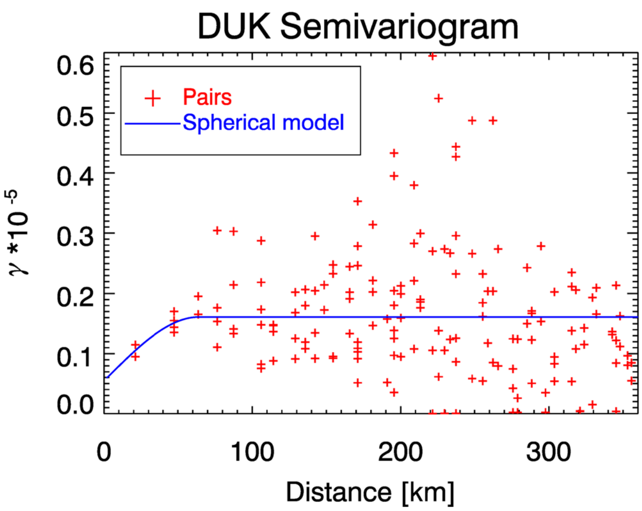

Fig. 4 The semivariogram and optimal fitted model (i.e. spherical) of the detrended universal kriging function for 30-year (1971–2000) average of annual precipitation in Central Brazil

distribution of the explanatory variables and their linear relationship (i.e. scatterplot) with annual precipitation. A limitation of the technique is the risk of the estimation turning into extrapolation due to its dependence on the fit of the regression model and distribution of the input datasets (Bostan et al. 2012; Bárdossy and Pegram 2013). On the other hand, the simplicity of the multiple regression method can produce reasonable estimates, and its analysis may significantly improve the GIS techniques for elaborating an objective mapping (Ninyerola et al. 2000; Naoum and Tsanis 2004; Mair and Fares 2010). In its simplest form, it is used to fit a straight line through points scattered in a plane. The mathematical background is described as

$$\widehat{Z}(s) = \beta_0 + (\beta_1 \cdot x_1(s)) + (\beta_2 \cdot x_2(s)) + \ldots + \left(\beta_p \cdot x_p(s)\right) \quad (15)$$

where $\hat{Z}$ is the predicted variable at location s; $x_1, x_2, \ldots, x_p$ are explanatory variables at the point of interest; and β_0 is the intercept and $\beta_1, \beta_2, \ldots, \beta_p$ are coefficients of linear combination.

3.1.7 Residual interpolation (MRegIDW and MRegOK)

The residual interpolation can be classified as a combined method (Tveito et al. 2008). The first-order trend component (i.e. multiple linear regression model) is removed from the observations before a spatial interpolation technique is applied. The resulting gridded estimates are then added to the gridded regression model, as demonstrated by many authors (e.g. Martínez-Cob 1996; Nalder and Wein 1998; Goovaerts 2000; Ninyerola et al. 2000; Lhotellier 2005; Ninyerola et al. 2007; Portalés et al. 2010; Di Piazza et al. 2011; Bostan et al. 2012). The most preferred methods for the interpolation of the residual field are geostatistical methods, inverse distance weighting and various spline techniques (Dobesch et al. 2007). We opted for applying both a deterministic method (i.e. IDW) and a geostatistical method (i.e. OK), also referred

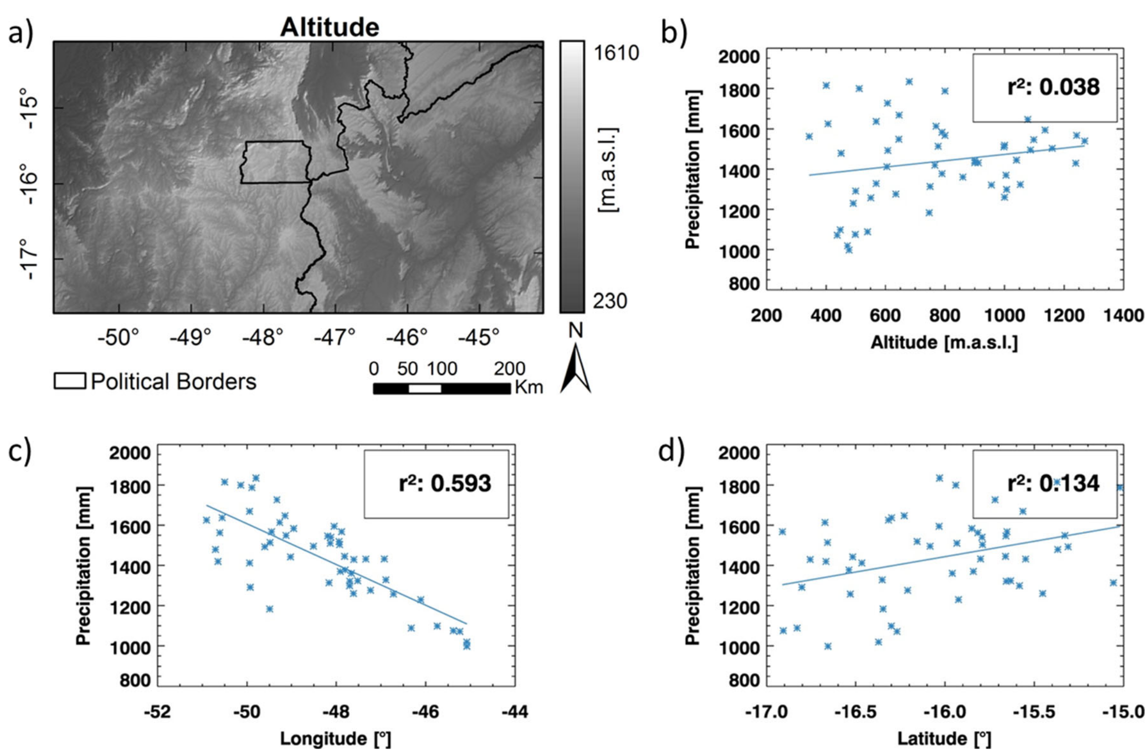

Fig. 5 The **a** spatial distribution of altitude and the linear relationship (scatter plot) of annual precipitation with **b** altitude, **c** longitude and **d** latitude

as kriging with an external drift or residual kriging. The basic formula is defined as

$$Z(s) = \widehat{Z}(s) + \varepsilon_{IDW/OK}(s) \tag{16}$$

where $\hat{Z}$ is the predicted variable at location s as demonstrated in formula (Eq. 15). The $\varepsilon_{\mathrm{IDW}}$ is the result of the interpolated residuals at location s using IDW with power value equal to 2 and $\varepsilon_{\mathrm{OK}}$ uses OK as residual interpolator. The semivariogram of the residual distribution is illustrated in Fig. 6. The fitting model is spherical and the lag size and number of lags are the same as for OK. The nugget effect is 0.03×10^{-5}, and partial sill is equal to 0.10×10^{-5}.

3.2 Calculation and validation

All calculations were performed using the Geostatistical Analyst tool of ArcGIS®9. Multiple linear regression equations were derived with a standard statistical programme and computed by using map calculator functions in ArcGIS®9. Before producing spatial distribution maps, the accuracy of estimations must be assessed. Cross-validation and validation may assist the decision to which model provides the best predictions (Tveito et al. 2008). The calculated statistics provides diagnostics to whether the model and/or its associated parameter values are reasonable (Johnston 2004). Cross-validation is probably the most widely applied method within climatology. In a cross-validation technique, one data point is left out of the data sample at a time while all the other data points are used to estimate the value for the point of interest left. This procedure is repeated until a value is estimated for all original data points (Isaaks 1989). One objection to using cross-validation is that the whole data sample is often used to define the interpolation model, and that the validation therefore might be considered to not be totally independent (Tveito et al. 2008). In order to avoid the limitation of using cross-validation, we opted for using an independent validation procedure that consists of splitting the data sample into two parts. Hence, 51 stations were used for interpolation purposes, while the remaining 15 % (9 stations) were used for independent validation (Fig. 1).

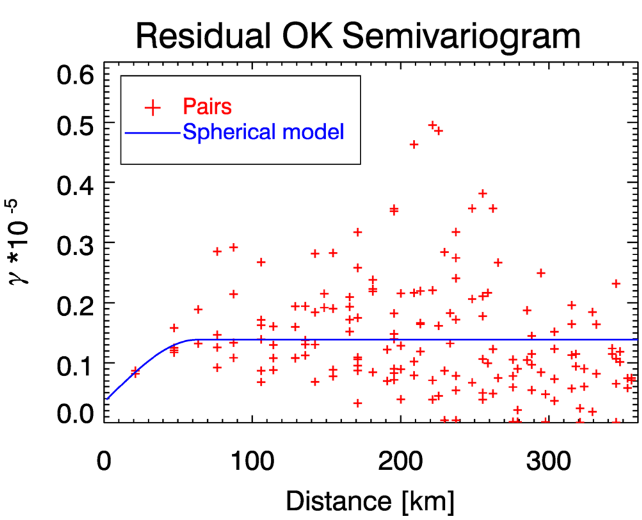

Fig. 6 The semivariogram and optimal fitted model (i.e. spherical) of the ordinary kriging function for the residual derived from the multivariate regression model. The data is a 30-year (1971–2000) average of annual precipitation in Central Brazil

Four measures were used for the comparison:

(1) Visual analysis according to the physical plausibility and consistency of the spatial estimates. It is expected that methods should capture local variations, especially those influenced by terrain characteristics. Nevertheless, methods are subject of inconsistent estimates due to the non-uniformity of the spatial distribution of the observation network. Deterministic methods may produce isolated "islands", while geostatistical models can create discontinue edges.

(2) Mean square error (MSE) measuring the difference between observed (z) and modelled ($\hat{z}$) in its average of the squares. This error measure is usually used as a criterion to compare the results of interpolation methods and should be as low as possible (Tveito et al. 2008);

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n} \left(\widehat{z}(x_i) - z(x_i) \right)^2}{n} \tag{17}$$

(3) Correlation coefficients (r), also referred to as Pearson's coefficient, is a measure of the linear dependence between two variables, in this case, the observed (z) and modelled ($\hat{z}$) (Tveito et al. 2008);

$$r = \frac{\sum_{i=1}^{n} \left(z_i - \overline{z} \right) \left(\widehat{z}_i - \overline{\widehat{z}} \right)}{\sqrt{\sum_{i=1}^{n} \left(z_i - \overline{z} \right)^2} \sqrt{\sum_{i=1}^{n} \left(\widehat{z}_i - \overline{\widehat{z}} \right)^2}} \tag{18}$$

(4) Nash–Sutcliffe efficiency criterion (NSE; Nash and Sutcliffe 1970): although widely applied for validating hydrological models, the NSE criterion can also be applied to assess the performance of interpolation methods in climatology (Zoccatelli et al. 2010; Wagner et al.

2012). This approach gives more weight to the real quantities than the correlation coefficient and therefore might be appropriated for the validation of rainfall.

$$\mathrm{NS} = 1 - \frac{\sum_{i=1}^{n}\left(z_i - \widehat{z}_i\right)^2}{\sum_{i=1}^{n}\left(z_i - \overline{z}\right)^2} \tag{19}$$

3.3 Assessment of uncertainties

No matter which interpolation technique is used, the interpolation values are only estimates of what the real values should be at a particular location. For any analysis of interpolated observed data, the level of uncertainty must be considered. The spatial distribution of uncertainty is essential to inform the reliability of the estimates to potential end users (Martínez-Cob 1996; Chiles 1999; Bárdossy and Pegram 2013). Geostatistical methods use the statistical properties of the measured data to estimate surfaces. The kriging assumes that the spatial variation can be modelled by random processes dependent on certain spatial autocorrelation. The statistical background allows not only the estimates of surfaces but also the quantification of uncertainties. In this study, a general assessment of the uncertainties is expressed in terms of the standard error of the prediction derived from the ordinary kriging method.

4 Results

Following the visual analysis, all methods have shown similar pattern on the spatial distribution of rainfall. Figure 6 illustrates the heterogenic distribution of annual rainfall where higher amounts are observed over the northwest and western parts of Distrito Federal. However, IDW, spline, OK, CoOK and DUK (Fig. 7a–e) do not capture local variations the way the multiple regression models do (Fig. 7f–h). Since MReg, MRegIDW and MRegOK take geographical factors into account, the distribution of rainfall over space is directly related to the spatial variability of the explanatory variables, especially topography. Moreover, spatial and temporal distribution of precipitation can be explained by its dependence on the Amazonian atmospheric circulation system, located to the northwest of the study region. The upper-tropospheric anticyclone, named the Bolivian High, causes strong convective heating of the atmosphere in the Amazon during the austral summer (Virji 1981). Low-level winds from the North Tropical Atlantic converge to the southeast due to Andes Cordillera bringing increasing temperature and humidity from the Amazon basin to Central Brazil (Gan et al. 2004). During the same period, a band of cloudiness and moisture is observed which extends over Amazon region to the Subtropical Atlantic Ocean, with the orientation northwest-southeast, denominated the South Atlantic Convergence Zone—SACZ (Carvalho and Jones 2009). Topography also plays a significant role on rainfall mechanisms by forcing the air lift in higher areas, as in the west of the Distrito Federal, creating greater convective activity compared to areas downwind of these natural heights (Vera et al. 2006). Following the literature, the multiple regression model (MReg), using geographical position and altitude as explanatory variables, has given reliable results showing fairly the same distribution patterns as the other methods. Figure 8 shows that spatial variance of precipitation is highly explained by longitude at annual scale, as well as for summer (December, January and February—DJF) and autumn (March, April and May—MAM). Latitude does not influence the precipitation to the same high degree as longitude, while altitude can add explanatory information to the model for winter (June, July and August—JJA) and spring (September, October and November—SON). The results support the inclusion of geographical variables in the spatial interpolation approach. The multiple regression model (MReg) was able to explain 67 % of the annual precipitation variance, and 60 % for summer (DJF), 77 % for autumn (MAM), 49 % for winter (JJA) and 60 % for spring (SON). Additionally, any kind of residual interpolation (i.e. MRegIDW and MRegOK) is very likely to increase the predictive performance of the regression model. The results achieved are in close agreement with several studies. For instance, Goovaerts (2000) and Haberlandt (2007) highlight the importance of topography in estimating spatial distribution of rainfall. Agnew and Palutikof (2000) and Portalés et al. (2010) demonstrated the added value of geographic information (e.g. longitude and latitude) when incorporated to interpolation methods.

Tables 1 and 2 show the mean square error and correlation coefficient between measured and estimated values, respectively. Spatial interpolation for summer (DJF) rainfall has the lowest error and highest correlations when MRegOK method is applied while MReg has the highest error and lowest correlation. Except for spline, all methods demonstrated low errors/high correlation in autumn (MAM). The DUK method performs the lowest errors/highest correlation in winter (JJA), while MRegIDW gives the poorest results. However, the rainfall average in JJA is very low (less than 30 mm), and therefore, statistical validation measures might not significantly influence the choice of the proper interpolation method. For spring (SON), the OK, IDW, MRegOK

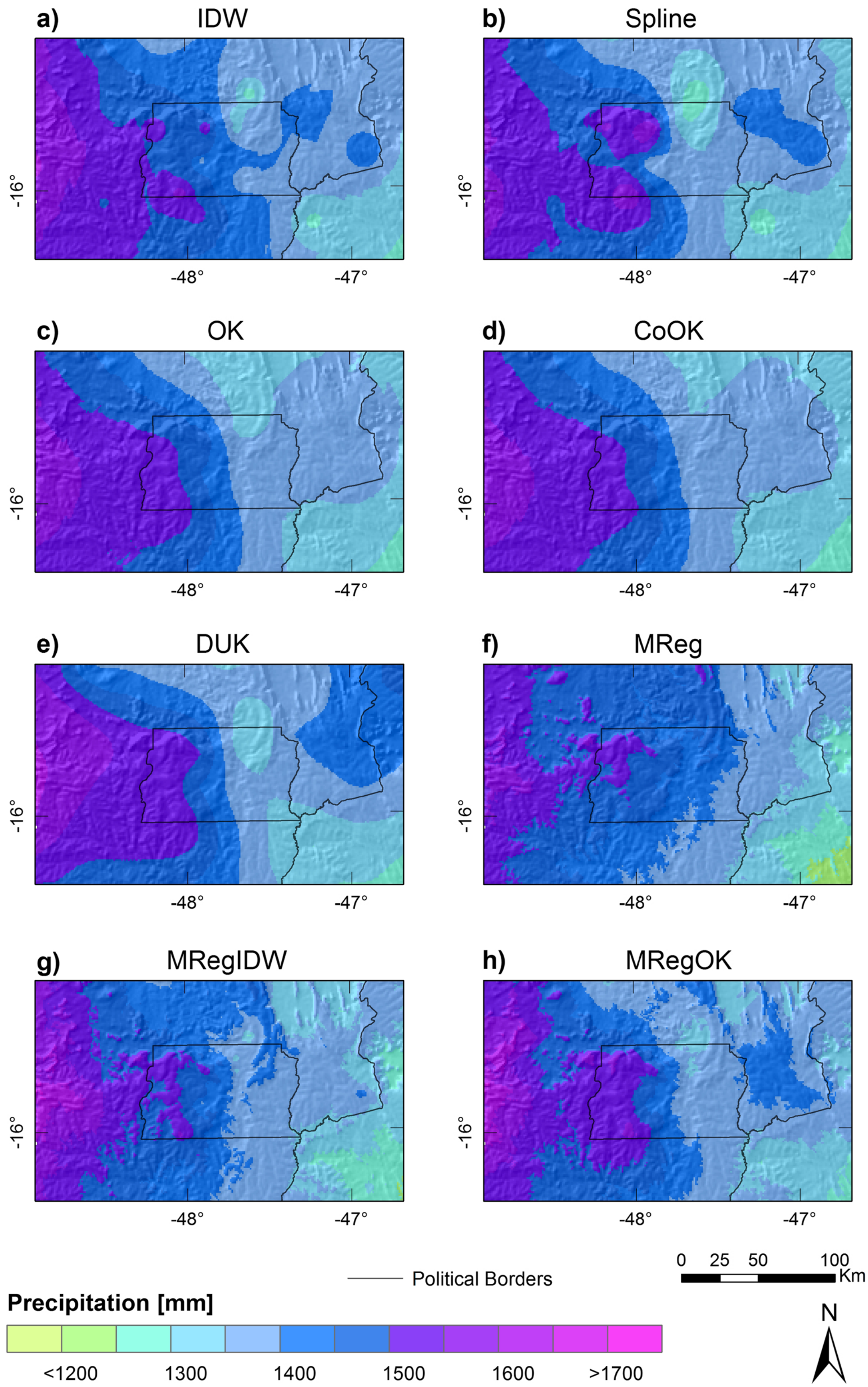

Fig. 7 Spatial estimates of 30-year (1971–2000) annual precipitation for Distrito Federal using eight different interpolation methods

and MRegIDW have the lowest error/highest correlation, while MReg performs the worst. The MRegIDW method has demonstrated the finest results in representing the distribution of annual rainfall, while spline gives the poorest results. Table 3 exhibits the results according to the NSE criterion where CoOK, followed by IDW and OK have performed the finest estimates while spline and MReg demonstrate inferior performance.

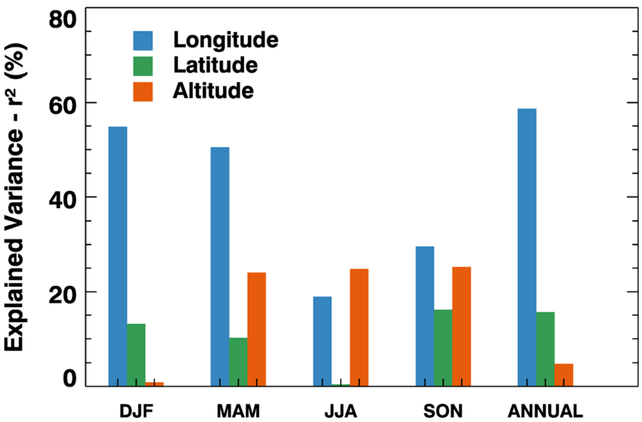

Fig. 8 Explained variance (i.e. square of the Pearson's correlation coefficient) of each potential explanatory variable in the linear regression model of seasonal and annual precipitation (1971–2000)

Table 2 Correlation coefficient between measured and predicted values

	DJF	MAM	JJA	SON	Annual
IDW	0.95[b]	0.93	0.81	0.96[b]	0.95
Spline	0.94	0.90[a]	0.82	0.94	0.92[a]
OK	0.94	0.92	0.89[b]	0.96[b]	0.94
CoOK	0.94	0.93	0.88	0.95	0.94
DUK	0.92	0.93	0.89[b]	0.95	0.92
MReg	0.92[a]	0.93	0.87	0.92[a]	0.95
MRegIDW	0.94	0.93	0.79[a]	0.96[b]	0.97[b]
MRegOK	0.95[b]	0.94[b]	0.81	0.96[b]	0.96

[a] Lowest correlation

[b] Highest correlation

In general, differences between methods are low and decrease as observation network increases. The raster statistics reveals that, over the Distrito Federal domain, the spatial average of annual precipitation varies from 1426 mm (spline) to 1450 mm (MReg). The standard deviation of the spatial distribution lies between 40.4 mm (MReg) and 80.5 mm (spline). The uncertainty, which is also a function of observation density, is addressed in this study. The representativeness of the observation network is a very challenging issue within climatology (Dobesch et al. 2007). Observation networks usually have an irregular spatial distribution and are mostly located in populated areas and lower altitudes (Dobesch et al. 2007; Tveito et al. 2008). The representativeness of observations can be described by standard error maps, such as from geostatistical methods. Figure 9 illustrates quite clearly the problems that arise with an irregular observation network. Low uncertainties are located mainly in the Distrito Federal domain area and populated areas in Goiás state, for instance, in the capital Goiânia and its surroundings. Remote areas in the north and northeast, where most of the observations are only available from the 1980s due to later settlement, are under high uncertainties. The map illustrates the spatial distribution of the standard error of the prediction in terms of percentage of the OK estimates. Assuming that the random process of the OK is normally distributed, there is a 95 % of confidence that the true value is within the interval of ±2 times the standard error (Johnston 2004). For instance, in a specific point in DF (longitude −47.9° and latitude −15.8°), the value estimated by the OK method is 1455 mm with a standard error of 6.7 %. That means that in 95 % of the cases in the random process of OK, the true prediction is a value between 1260 and 1650 mm. For a remote location in northeast of the Central Brazil area (longitude −45.0° and latitude −14.5°), the uncertainty of the OK estimate is very high. The standard error is 19.2 % and, hence, the true prediction is a value between 724 and 1626 mm. Analogous to other studies (e.g. Martínez-Cob 1996; Borga and Vizzaccaro 1997; Bárdossy and Pegram 2013), the performance of the interpolation methods is highly dependent on the density of the observation network. Some methods may be more appropriated for areas with sparse observation networks and specific topographic conditions. Given these findings, the inclusion of explanatory variables,

Table 1 Mean square error (mm) between measured and predicted values

	DJF	MAM	JJA	SON	Annual
IDW	1141.5	327.0[b]	10.3	197.5	3334.5
Spline	1251.3	478.3[a]	9.2	245.8	4950.0[a]
OK	1256.2	374.7	5.7	180.3[b]	3681.2
CoOK	1204.8	330.7	5.7	190.2	3393.4
DUK	1509.9	348.4	5.6[b]	202.8	4602.4
MReg	1786.0[a]	428.6	6.5	314.4[a]	3833.0
MRegIDW	1244.5	391.4	13.4[a]	223.6	2512.0[b]
MRegOK	1059.4[b]	352.1	12.1	213.1	2956.9

[a] Highest errors

[b] Lowest errors

Table 3 Nash–Sutcliffe criterion between measured and predicted values

	DJF	MAM	JJA	SON	Annual
IDW	0.88	0.87[b]	0.55	0.90	0.89
Spline	0.86	0.81[a]	0.60	0.88	0.84[a]
OK	0.86	0.85	0.75[b]	0.91[b]	0.88
CoOK	0.87	0.85	0.75[b]	0.91[b]	0.89
DUK	0.84	0.87[b]	0.75[b]	0.90	0.85
MReg	0.81[a]	0.83	0.72	0.85[a]	0.87
MRegIDW	0.87	0.85	0.42[a]	0.89	0.92[b]
MRegOK	0.89[b]	0.86	0.47	0.90	0.90

[a] Lowest value

[b] Highest value

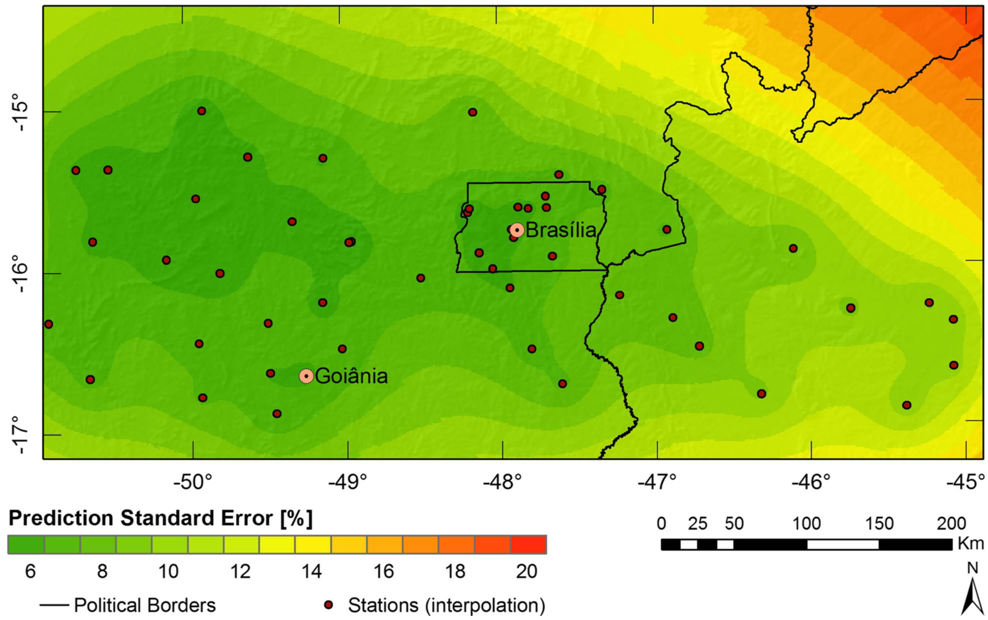

Fig. 9 Uncertainties of interpolation demonstrated by prediction standard error [%] map of ordinary kriging for annual accumulated precipitation in Central Brazil

such as location and topography, to the interpolation model should assist estimates of rainfall in areas with low observation density in Central Brazil.

5 Conclusions

Eight univariate and multivariate techniques are compared for the estimation of seasonal and annual precipitation climatology (1971–2000) in DF. According to existing literature, regional precipitation is very likely to be dependent on atmospheric circulation patterns originating from the Amazon region and might be increased due to orography. The dependence of long-term accumulated rainfall on regional geographical characteristics was confirmed. The explanatory variables used, particularly longitude, could explain a large portion of precipitation. In order to assess the performance of the methods tested, a validation approach based on visual analysis and basic statistics was applied. Visual examination reveals that the multiple linear regression models provide more detailed spatial variability than others. The MSE, correlation coefficient and NSE criterion confirmed the reliability of IDW, OK and residual interpolation (MRegIDW and MRegOK) in estimating rainfall in DF. In general, for the resolution wanted (i.e. 1 km), interpolation methods provide similar spatial distributions wherever observation network is dense. In remote areas, a general assessment illustrates the large uncertainties associated to the geostatistical methods. The inclusion of geographical factors to the interpolation method demonstrates a potential to improve estimates in areas where observation density is low. The use of MRegIDW and MRegOK is preferred; however, the deterministic component of these methods does not allow the assessment of prediction errors. It is, therefore, recommended to use the standard error maps of geostatistical methods as supplementary and qualitative information for further applications of the methods suggested.

Nevertheless, before applying this climate information to any water-related study, it is recommended to test the sensitivity of the target system to the spatial variability of rainfall derived from a range of plausible but distinct interpolation methods. The sensitivity analysis is a common approach applied in hydrological modelling studies. The main objective is to understand the behaviour of a system to alterations in its major driving forces (e.g. spatial distribution of rainfall). In the case of DF, water planners should consider more than one method to assess the uncertainties associated to the interpolation of point data, for instance the methods recommended in this study. If the system under concern is substantially sensitive to the interpolation method applied, the ensemble of these methods is likely to be the most appropriated estimate (Strauch et al. 2012). Additionally, several statistical downscaling methods are limited to single site projections. As impact modellers may be interested in the multi-site information, the investigations conducted in this study may support the use of the recommended methods in providing plausible distributions of future climate projections over DF.

Further research should investigate whether other environmental variables (e.g. direction of winds and slope orientation) allow increase in explanatory power of the regression models. The use of auxiliary rainfall information, such as radar and satellite data, is likely to increase the reliability of estimates. This might well be the case for lower time-aggregated data, such as monthly data. Furthermore, following this study, we recommend evaluating the performance of the residual interpolation methods using the same number of stations for periods where the observation network is higher, for instance for 1981–2010.

Acknowledgments The author wishes to thank the International Water Research Alliance Saxony—IWAS initiative (grant 02WM1028) and the International Postgraduate Studies in Water Technologies—IPSWaT (grant number IPS 10/15P) scholarship programme, both funded by the German Federal Ministry of Education and Research—BMBF, for the opportunity given.

References

Agnew M, Palutikof J (2000) GIS-based construction of baseline climatologies for the Mediterranean using terrain variables. Clim Res 14: 115–127. doi:10.3354/cr014115

Alter JC (1919) Normal precipitation in Utah. Mon Weather Rev 47:633–636. doi:10.1175/1520-0493(1919)47<633:NPIU>2.0.CO;2

Alves LM (2009) Clima da Região Centro-Oeste do Brasil. In: Cavalcanti IFA, Ferreira NJ, Da Silva MGAJ, Silva Dias MAF (eds) Tempo e Clima no Brasil. Oficina de Textos, São Paulo, pp 235–241

Anagnostou EN, Krajewski WF, Smith J (1999) Uncertainty quantification of mean-areal radar-rainfall estimates. J Atmos Ocean Technol 16: 206–215. doi:10.1175/1520-0426(1999)016<0206:UQOMAR>2.0.CO;2

Apaydin H, Sonmez FK, Yildirim YE (2004) Spatial interpolation techniques for climate data in the GAP region in Turkey. Clim Res 28: 31–40. doi:10.3354/cr028031

Bárdossy A, Pegram G (2013) Interpolation of precipitation under topographic influence at different time scales. Water Resour Res 49: 4545–4565. doi:10.1002/wrcr.20307

Barrows HK (1933) Precipitation and runoff and altitude relations for Connecticut River. Trans Am Geophys Union 14:396. doi:10.1029/TR014i001p00396

Borga M, Vizzaccaro A (1997) On the interpolation of hydrologic variables: formal equivalence of multiquadratic surface fitting and kriging. J Hydrol 195:160–171. doi:10.1016/S0022-1694(96)03250-7

Borges PA, Franke J, Santos Silva F, Weiss H, Bernhofer C (2014) Differences between two climatological periods (2001–2010 vs. 1971–2000) and trend analysis of temperature and precipitation in Central Brazil. Theor Appl Climatol 116:191–202. doi:10.1007/s00704-013-0947-4

Bostan PA, Heuvelink GBM, Akyurek SZ (2012) Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. Int J Appl Earth Obs Geoinform 19:115–126. doi:10.1016/j.jag.2012.04.010

Carvalho LMV, Jones C (2009) Zona de Convergência do Atlântico Sul. In: Cavalcanti IFA, Ferreira NJ, Da Silva MGAJ, Silva Dias MAF (eds) Tempo e Clima no Brasil. Oficina de Textos, São Paulo, pp 95–109

Chiles J-P (1999) Geostatistics modeling spatial uncertainty. Wiley, New York

Chua S-H, Bras RL (1982) Optimal estimators of mean areal precipitation in regions of orographic influence. J Hydrol 57:23–48. doi:10.1016/0022-1694(82)90101-9

Coulibaly M, Becker S (2007) Spatial interpolation of annual precipitation in South Africa—comparison and evaluation of methods. Water Int 32:494–502. doi:10.1080/02508060708692227

CPTEC-INPE (2014) Radar precipitation. Satellite division and environmental systems—DSA. Center for Weather Forecasting and Climate Research—CPTEC, National Institute for Space Research—INPE. http://sigma.cptec.inpe.br/radar/. Accessed 25 November 2014

Cressie N (1990) The origins of kriging. Math Geol 22:239–252. doi:10.1007/BF00889887

Cressie N (1993) Statistics for spatial data. Wiley, New York

Daly C, Neilson RP, Phillips DL (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J Appl Meteorol 33:140–158. doi:10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2

Delbari M, Afrasiab P, Jahani S (2013) Spatial interpolation of monthly and annual rainfall in northeast of Iran. Meteorol Atmos Phys 122: 103–113. doi:10.1007/s00703-013-0273-5

Demyanov V, Kanevsky M, Chernov S, Savelieva E, Timonin V (1998) Neural network residual kriging application for climatic data. J Geogr Inf Decis Anal 2(2):215–232

Di Piazza A, Conti FL, Noto L, Viola F, La Loggia G (2011) Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. Int J Appl Earth Obs Geoinform 13:396–408. doi:10.1016/j.jag.2011.01.005

Dingman SL, Seely-Reynolds DM, Reynolds RC (1988) Application of kriging to estimating mean annual precipitation in a region of orographic influence. JAWRA J Am Water Resour Assoc 24:329–339. doi:10.1111/j.1752-1688.1988.tb02991.x

Dixon WJ (1950) Analysis of extreme values. Ann Math Stat 21:488–506. doi:10.1214/aoms/1177729747

Dobesch H, Dumolard P, Dyras I (2007) Spatial interpolation for climate data: the use of GIS in climatology and meteorology. ISTE, London

Frei C, Schär C (1998) A precipitation climatology of the Alps from high-resolution rain-gauge observations. Int J Climatol 18:873–900. doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9

Gan M, Kousky V, Ropelewski C (2004) The South America monsoon circulation and its relationship to rainfall over west-central Brazil. J Clim 17:47–66. doi:10.1175/1520-0442(2004)017<0047:TSAMCA>2.0.CO;2

Gan W, Chen X, Cai X, Zhang J, Feng L, Xie X (2010) Spatial interpolation of precipitation considering geographic and topographic influences-a case study in the Poyang Lake Watershed, China. Geosci Remote Sens Symp 2010 I.E. Int 3972–3975. doi:10.1109/IGARSS.2010.5650156

Goldberger AS (1962) Best linear unbiased prediction in the generalized linear regression model. J Am Stat Assoc 57:369–375. doi:10.1080/01621459.1962.10480665

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 Hydrol 228:113–129. doi:10.1016/S0022-1694(00)00144-X

Haberlandt U (2007) Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. J Hydrol 332:144–157. doi:10.1016/j.jhydrol.2006.06.028

Habib E, Krajewski WF (2002) Uncertainty analysis of the TRMM ground-validation radar-rainfall products: application to the TEFLUN-B field campaign. J Appl Meteorol 41:558–572. doi:10.1175/1520-0450(2002)041<0558:UAOTTG>2.0.CO;2

Hong Y, Nix HA, Hutchinson MF, Booth TH (2005) Spatial interpolation of monthly mean climate data for China. Int J Climatol 25:1369–1379. doi:10.1002/joc.1187

Hutchinson MF, Gessler PE (1994) Splines—more than just a smooth interpolator. Geoderma 62:45–67. doi:10.1016/0016-7061(94)90027-2

Isaaks EH (1989) Applied geostatistics. Oxford University Press, New York

Jarvis CH, 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 Appl Meteorol 40:1075–1084. doi:10.1175/1520-0450(2001)040<1075:ACASFI>2.0.CO;2

Johnston K (2004) ArcGIS 9: using ArcGIS geostatistical analyst. Esri Press, Redlands

Lanza L, Ramírez J, Todini E, others (2001) Stochastic rainfall interpolation and downscaling. Hydrol Earth Syst Sci 5(2):139–143. doi:10.5194/hess-5-139-2001

Lhotellier R (2005) Spatialisation des températures en zone de montagne alpine. Dissertation, Université Joseph Fourier—Grenoble 1

Li J, Heap AD (2011) A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. Ecol Inform 6:228–241. doi:10.1016/j.ecoinf.2010.12.003

Lorz C, Abbt-Braun G, Bakker F, Borges P, Börnick H, Fortes L, Frimmel F, Gaffron A, Hebben N, Höfer R, Makeschin F, Neder K, Roig HL, Steiniger B, Strauch M, Walde DH, Weiß H, Worch E, Wummel J (2012) Challenges of an integrated water resource management for the Distrito Federal, Western Central Brazil: climate, land-use and water resources. Environ Earth Sci 65(5):1363–1366. doi:10.1007/s12665-011-1219-1

Mair A, Fares A (2010) Comparison of rainfall interpolation methods in a mountainous region of a tropical island. J Hydrol Eng 16:371–383. doi:10.1061/(ASCE)HE.1943-5584.0000330

Martínez-Cob A (1996) Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. J Hydrol 174:19–35. doi:10.1016/0022-1694(95)02755-6

Nalder IA, Wein RW (1998) Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. Agric For Meteorol 92:211–225. doi:10.1016/S0168-1923(98)00102-6

Naoum S, Tsanis I (2004) Ranking spatial interpolation techniques using a GIS-based DSS. Glob Nest 6:1–20

Nash JE, Sutcliffe J (1970) River flow forecasting through conceptual models part I—a discussion of principles. J Hydrol 10:282–290. doi:10.1016/0022-1694(70)90255-6

Ninyerola M, Pons X, Roure JM (2000) A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. Int J Climatol 20:1823–1841. doi:10.1002/1097-0088(20001130)20:14<1823::AID-JOC566>3.0.CO;2-B

Ninyerola M, Pons X, Roure JM (2007) Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a geographic information system. Theor Appl Climatol 89:195–209. doi:10.1007/s00704-006-0264-2

Pardo-Igúzquiza E (1998) Comparison of geostatistical methods for estimating the areal average climatological rainfall mean using data on precipitation and topography. Int J Climatol 18:1031–1047. doi:10.1002/(SICI)1097-0088(199807)18:9<1031::AID-JOC303>3.0.CO;2-U

Perčec Tadić M (2010) Gridded Croatian climatology for 1961–1990. Theor Appl Climatol 102:87–103. doi:10.1007/s00704-009-0237-3

Phillips DL, Dolph J, Marks D (1992) A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. Agric For Meteorol 58:119–141. doi:10.1016/0168-1923(92)90114-J

Portalés C, Boronat N, Pardo-Pascual JE, Balaguer-Beser A (2010) Seasonal precipitation interpolation at the Valencia region with multivariate methods using geographic and topographic information. Int J Climatol 30(10):1547–1563. doi:10.1002/joc.1988

Price DT, McKenney DW, Nalder IA, Hutchinson MF, Kesteven JL (2000) A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agric For Meteorol 101:81–94. doi:10.1016/S0168-1923(99)00169-0

Ramos AM (2009) Normais climatológicas do Brasil 1961–1990. INMET, Brasília

Renard F, Comby J (2006) Evaluation de techniques d'interpolation spatiale de la pluie en milieu urbain pour une meilleure gestion d'événements extrêmes: le cas du Grand Lyon. La Houille Blanche 73–78. doi:10.1051/lhb:2006104

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. ACM Press, pp 517–524

Simpson JS, Adler RF, North GR (1988) A proposed Tropical Rainfall Measuring Mission (TRMM) satellite. Bull Am Meteorol Soc 69:278–295. doi:10.1175/1520-0477(1988)069<0278:APTRMM>2.0.CO;2

Smith RB (1979) The influence of mountains on the atmosphere. Adv Geophys 21:87–230

Spreen WC (1947) A determination of the effect of topography upon precipitation. Trans Am Geophys Union 28:285. doi:10.1029/TR028i002p00285

Stein A, Van Eijnsbergen A, Barendregt L (1991) Cokriging nonstationary data. Math Geol 23:703–719. doi:10.1007/BF02082532

Strauch M, Bernhofer C, Koide S, Volk M, Lorz C, Makeschin F (2012) Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation. J Hydrol 414:413–424. doi:10.1016/j.jhydrol.2011.11.014

Szentimrey T, Bihari Z, Szalai S (2005) Meteorological interpolation based on surface homogenized data basis (MISH). European Geosciences Union, General Assembly 2005, Vienna, Austria

Tabios GQ, Salas JD (1985) A comparative analysis of techniques for spatial interpolation of precipitation. JAWRA J Am Water Resour Assoc 21:365–380. doi:10.1111/j.1752-1688.1985.tb00147.x

Thiessen AH (1911) Precipitation averages for large areas. Mon Weather Rev 39:1082–1089. doi:10.1175/1520-0493(1911)39<1082b:PAFLA>2.0.CO;2

Tveito OE, Wegehenkel M, van der Wel F, Dobesch H (2008) COST Action 719: the use of geographic information systems in climatology and meteorology. EUR-OP, Luxembourg

Vera C, Higgins W, Amador J, Ambrizzi T, Garreaud R, Gochis D, Gutzler D, Lettenmaier D, Marengo J, Mechoso CR, Nogues-Paegle J, Silva Dias PL, Zhang C (2006) Toward a unified view of the American monsoon systems. J Clim 19:4977–5000. doi:10.1175/JCLI3896.1

Virji H (1981) A preliminary study of summertime tropospheric circulation patterns over South America estimated from cloud winds. Mon Weather Rev 109:599–610. doi:10.1175/1520-0493(1981)109<0599:APSOST>2.0.CO;2

Wackernagel H (2003) Multivariate geostatistics: an introduction with applications. Springer, New York

Wagner PD, Fiener P, Wilken F, Kumar S, Schneider K (2012) Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. J Hydrol 464:388–400. doi:10.1016/j.jhydrol.2012.07.026

WMO (1989) Calculation of monthly and annual 30-year standard normals. World Meteorological Organization (WMO), WCDP-No.10 1–14

WMO (1996) Climatological normals (CLINO) for the period 1961–1990 = normales climatologiques (CLINO) pour la période 1961–1990. Secretariat of the World Meteorological Organization (WMO), Geneva

WMO (2008) Guide to hydrological practices. Hydrology—from measurement to hydrological information. World Meteorological Organization (WMO), Geneva

Zoccatelli D, Borga M, Zanon F, Antonescu B, Stancalie G (2010) Which rainfall spatial information for flash flood response modelling? A numerical investigation based on data from the Carpathian range, Romania. J Hydrol 394:148–161. doi:10.1016/j.jhydrol.2010.07.019