A COMPARISON OF SPATIAL RAINFALL ESTIMATION TECHNIQUES: A CASE STUDY OF NYANDO RIVER BASIN KENYA

***Re-published**

F. Mutua and D. Kuria
Jomo Kenyatta University of Agriculture and Technology
E-mail: felix.mutua@gmail.com

Abstract

Many hydrological models for watershed management and planning require rainfall as an input in a continuous format. This study analyzed four different rainfall interpolation techniques in Nyando river basin, Kenya. Interpolation was done for a period of 30 days using 19 rainfall stations. Two geostatistical interpolation techniques (kriging and cokriging) were evaluated against inverse distance weighted (IDW) and global polynomial interpolation (GPI). Of the four spatial interpolators, kriging and cokriging produced results with the least root mean square error (RMSE). A digital elevation model (DEM) was introduced into the cokriging method and this improved the results considerably. The results demonstrate that for low-resolution rain gauge networks, geostatistical interpolation methods perform better than other techniques that ignore spatial dependence patterns. The use of secondary information improved the prediction results, as demonstrated by the inclusion of the DEM in this study.

Key words: DEM, GPI, IDW, kriging, rainfall

1.0 Introduction

Watershed models are powerful tools for water resources planning and management. Models may be used to predict how conditions are expected to change over time, to understand the nature and scope of a problem and to evaluate alternative management options. Precipitation is the primary input to hydrological models and is characterized by spatial variations. This is brought about by differences in the type and scale of development of precipitation producing processes, and strongly influenced by local or regional factors, such as topography and wind direction at the time of precipitation (Sumner, 1998). Rainfall data is traditionally presented as point data. However, hydrological modelling requires spatial representation of rainfall and thus the gauge measurements need to be transformed into areal coverages.

Several methodologies exist for spatial interpolation of climate and weather parameters. These include; inverse distance weighting (Englund and Weber, 1994), Thiessen polygon method, isohyetal method, and more sophisticated statistical methods such as kriging , and it's various extensions, also known as geostatistical methods (Symeonakis, 2008). Recent advances in the fields of geographic information systems (GIS) and remote sensing have made satellite-based rainfall estimates readily available in coverage forms. However, the scales, temporal and spatial resolutions are usually coarse for application at a basin scale.

Several attempts have been made to compare these methods with most of the studies suggesting that geostatistical methods provide the most accurate estimates. Goovaerts (2000) used geostatistical algorithms to include elevation into the interpolation procedure at the South of Portugal. Johansson and Chen (2003) developed a regression model that included the wind variable for Sweden. Marquínez et al. (2003) used a regression model performed with topographic variables for Cantabria (Spain). Vicente-Serrano et al. (2003) concluded that the best results were achieved by geostatistical methods and a regression model formed by four geographic variables for the Ebro Valley (Spain). Subyani (2004) used geostatistical methods in the study of annual and seasonal rainfall patterns in south west Arabia; whilst Eulogio (1998) used geostatistical methods in estimating areal climatological rainfall mean using data and precipitation in southern Spain.

Majority of these studies focused on interpolating precipitation for small to regional scale applications emphasizing the need for similar research over larger areas to support the respective work on hydrological processes, such as surface runoff and soil erosion (Symeonakis, 2008). This study attempts to evaluate four geostatistcial interpolation techniques and compare their performance in generating spatial distributions of rainfall in Nyando river basin. Section 2) provides a brief discussion of spatial interpolation methods, section 3) discusses methods, data, and tools while results are presented in section 4.

2.0 Interpolation methods

2.1 Inverse Distance Weighting (IDW)

Inverse Distance Weighting (IDW) is an interpolation technique in which interpolated estimates are made based on values at nearby locations, weighted only by distance from the interpolation location. IDW explicitly implements the assumption that things that are close to one another are more alike than those that are farther apart. In the IDW approach, the values to be interpolated Z^*_{IDW} are estimated as a linear combination of several surrounding observations, with the weights being inversely proportional to the distance between observations and location *u* to the power of *p*.

$$Z^*_{IDW}(u) = \frac{\sum_{\alpha=1}^{n(u)} \lambda_\alpha(u) Z(u_\alpha)}{\sum_{\alpha=1}^{n(u)} \lambda_\alpha(u)}, \text{ with } \lambda_\alpha(u) = \frac{1}{|u - u_\alpha|^p} \quad \text{.....(1)}$$

Where *n(u)* is the number of the points at location *u* considered for the estimation and $\lambda_\alpha(u)$ is the weight.

2.2 Global polynomial interpolation (GPI)

Global Polynomial interpolation (GPI) is a quick and smooth deterministic interpolator. A first-order global polynomial (GP) fits a single plane through the data; a second-order fits a surface with a bend in it, allowing the calculation of surfaces representing valleys; a third-order allows for 2 bends; and so forth. However, when a surface has a different shape, as in a landscape that slopes, levels out, and then slopes again, a single GP will not fit well (Johnston et al., 2001).

2.3 Kriging

Kriging is a moderately quick interpolator that can be exact or smoothed depending on the measurement error model. Kriging uses statistical models that allow a variety of map outputs including predictions, standard errors and probability. Kriging assigns weights according to a (moderately) data-driven weighting function, rather than an arbitrary function, but it is still an interpolation algorithm and will give very similar results to others in many cases (Isaaks and Srivastava, 1992). Kriging estimators are variants of the basic linear regression estimator:

$$Z^*(u) - m(u) = \sum_{\alpha=1}^{n(u)} \lambda_\alpha(u)\left[Z(u_\alpha) - m(u_\alpha)\right] \quad \text{.....(2)}$$

Where *n(u)* is the number of the points at location *u* and λ_α *(u)* is the weight assigned to the datum $z(u_\alpha)$ interpreted as a realization of the random variable $\mathrm{Z}(\mathrm{u}_\alpha)$. The values $m(u)$ and $m(u_\alpha)$ are the expected values of the random variables *Z(u)* and $\mathrm{Z}(\mathrm{u}_\alpha)$ respectively.

Kriging estimators are required to be unbiased and to minimize error variance, i.e. $\sigma_{\mathrm{E}}^2(\mathrm{u}) = \mathrm{Var}\left\{\mathrm{Z}^*(\mathrm{u}) - \mathrm{Z}\ (\mathrm{u})\right\}$ under the constraint that the expected error is zero: $E\left\{Z^*(u) - Z(u)\right\} = 0$. Each random function is usually decomposed into a residual component and a trend component: $Z(u) = R(u) + m(u)$. The residual component is modeled as a stationary random function with zero mean and covariance function.

$\mathrm{C}_{\mathrm{R}}(\mathrm{h}):\ E\{R(u)\} = 0$

$\mathrm{cov}\{R(u), R(u+h)\} = E\{R(u).R(u+h)\} = C_R(h)$

The expected value of the random variables *Z* at a certain location *u* is the value of the trend component at that location $E\{Z(u)\} = m(u)$. Three Kriging variants can be distinguished according to the model considered for the trend: Simple, ordinary and universal. Simple kriging (SK) considers the mean $m(u)$ to be known and constant through the study area. The SK estimator is expressed mathematically by equation 3.

$$\mathrm{Z}^*_{\mathrm{SK}}(\mathrm{u}) = \sum_{\alpha=1}^{\mathrm{n(u)}} \lambda_\alpha^{\mathrm{SK}}(\mathrm{u})\mathrm{Z}(\mathrm{u}_\alpha) + \lambda_{\mathrm{m}}^{\mathrm{SK}}(\mathrm{u})\mathrm{m}\ ,\ \text{with}\ \lambda_\alpha^{\mathrm{SK}}(\mathrm{u}) = 1 - \sum_{\alpha=1}^{\mathrm{n(u)}} \lambda_\alpha^{\mathrm{SK}}(\mathrm{u}) \quad [3]$$

The SK weights are determined such as to minimize the error variance, while ensuring the unbiasedness of the estimator. More detailed discussion on kriging methods can be found in Goovaerts (2000).

2.4 Cokriging

The addition of cross-related information reduces the variance of estimation error by using the cokriging method. In this case, the primary data $\{\mathrm{z}_1(\mathrm{u}\alpha_1), \alpha_1 = 1,2.....\mathrm{n}_1\}$ is supplemented by secondary data related to $(N_v - 1)$ continuous attributes Z_{i}, $\{z_1(u\alpha_1), \alpha_1 = 1,2.....n_i, i = 2,...., N_v\}$ at N possible different locations. The linear kriging estimator (equation 2) is extended to incorporate such additional information.

$$\mathrm{Z}_1^*(\mathrm{u}) - \mathrm{m}_1(\mathrm{u}) = \sum_{\alpha_1=1}^{\mathrm{n_1(u)}} \lambda_{\alpha_1}(\mathrm{u})\left[\mathrm{z}_1(\mathrm{u}_{\alpha_1}) - \mathrm{m}_1(\mathrm{u}_{\alpha_1})\right] + \sum_{\mathrm{i}=2}^{\mathrm{N_v}} \sum_{\alpha_\mathrm{i}=1}^{\mathrm{n_i(u)}} \lambda_{\alpha_{\mathrm{i}1}}(\mathrm{u})\left[\mathrm{z}_{\mathrm{i}}(\mathrm{u}_{\alpha_\mathrm{i}}) - \mathrm{m}_{\mathrm{i}}(\mathrm{u}_{\alpha_\mathrm{i}})\right] \quad [4]$$

Where $\lambda_{\alpha_1}(u)$ is the weight assigned to the primary datum $z_1(u_{\alpha_1})$ and $\lambda_{\alpha_1}(u)$, i>1 is the weight assigned to the secondary datum $z_i(u_{\alpha_i})$. The terms $m_1(u_{\alpha_1})$ and $m_i(u_{\alpha_i})$ are the expected values for the random variables $z_1(u_{\alpha_1})$ and $z_i(u_{\alpha_i})$ respectively.

Cokriging estimators are required to be unbiased and to minimize error variance $\sigma_E^2(u) = \mathrm{Var}\{z_1^*(u) - z_1(u)\}$ under the constraint that the expected error is zero: $E\{z_1^*(u) - z_1(u)\} = 0$. Each random function can be written as the sum of a residual component plus a trend component: $Z_i(u) = R_i(u) + m_i(u), i = 1,....., N_v$.

The residual component is modeled as a stationary random function with zero mean and covariance function $C_i^R(h) : E\{R_i(u)\} = 0$ and $\mathrm{cov}\{R_i(u), R_i(u+h)\} = E\{R_i(u).R_i(u+h)\} = C_i^R(h)$. The cross variance is given by $C_{ij}^R(h) = \mathrm{cov}\{R_i(u), R_j(u+h)\}$. According to the trend model $m_i(u)$, three cokriging models can be distinguished: Simple, ordinary and universal. The Simple cokriging (SCK) method considers each local mean, known and constant within the study area. The SCK estimator is expressed mathematically by equation 5.

$$\sum_{j=1}^{N_v} \sum_{\beta j i=1}^{n_j(u)} \lambda_{\beta_j}^{sck}(u) C_{ij}(u_{\alpha_i} - u_{\beta_j}) = C_{i1}(u_{\alpha_i} - u), ... \alpha_i = 1,...., n_i(u), i = 1,....., N_v \quad ....[5]$$

3.0 Materials and methods

3.1 Study area

Nyando basin (Figure 1) is located on the western region of Kenya and is a part of the greater Lake Victoria Basin. It's geographically located along the equator bounded by latitudes $0^0 7'$N and $0^0 24'$S; longitudes $34^0 25'$E and $35^0 43'$E. It covers an area of about 3500 km^2.

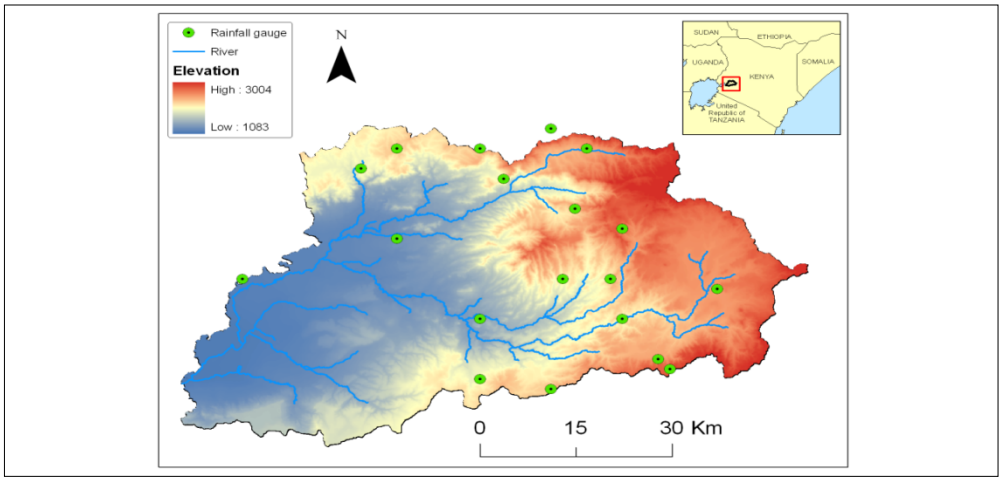


Figure 1: Study Area

Rainfall in the region is mainly influenced by the migration of the Inter tropical Convergence zone (ITCZ) and exhibits a bi-modal pattern with peaks in the long rains season March-April-May (MAM) and short rains October-November-December (OND). The mean annual rainfall varies from 1,000 mm near Lake Victoria to approximately 1,600 mm in the highlands. Land use and property rights vary across the basin. The upper part of the basin is comprised of gazetted forests, commercial tea production, and small-scale agriculture on steep hillsides that were de-gazetted as forests during the last 40 years. Mid-altitude land uses are composed of a mixture of smallholder farms (with maize, beans and some coffee, bananas, sweet potatoes and dairy activities) and large-scale commercial farms (mostly sugar cane).

3.2 Data

Table 1 shows a summary of the datasets used in this study

Table 1: Datasets

Dataset	Type	Source	Specifications
Rainfall data	Rain gauge data, tabular	Kenya Meteorological Department	Primary data(mm/day)
DEM	Elevation(Raster)	USGS	Secondary data (90 m resolution)
Nyando Basin boundary	GIS File (shapefile)	Department of Biomechanical and Engineering Department (BEED), Jomo Kenyatta University of Agriculture Technology (JKUAT)	Secondary data (Shapefile)

Others: Roads, river network, boundaries	GIS files(Shapefile)	Department of Geomatic Engineering and Geospatial Information Systems (GEGIS), Jomo Kenyatta University of Agriculture Technology	Auxiliary data (Shapefile)

Data preparation and analysis was carried out using Ms Office Excel® and ArcGIS 9.3®. ArcGIS 9.3 provided the GIS platform for visualization, manipulation of data production of maps. The ArcGIS Geostatistical Analyst® tool was used for interpolation, production of maps and error plots. The tool provides advanced statistical tools for surface generation, analysis and mapping of continuous datasets. It includes exploratory spatial data analysis tools providing insights about data distribution, global and local outliers, global trends, levels of spatial autocorrelation, and variation among multiple datasets (ESRI, 2007).

Rainfall data was obtained from the Kenya meteorological department. There are 25 rainfall stations near the Nyando river basin and only 19 stations are within the basin. This data was characterized by gaps (~40%) and a period of 30 days was chosen partly because of this challenge.

Table 2: Rainfall Stations in Nyando Basin

Station ID	Station Name (stn)	Latitude (N/S)	Longitude (E)	Year Opened	Height (M)
9035002	Londiani Forest Station	-0.150	35.600	1908	2316
9035020	Kipkelion Railway Station	-0.200	35.467	1904	1931
9035042	Equator Barguat Estate	-0.017	35.400	1932	2012
9035068	Kipkelion Morau Company Ltd.	-0.133	35.450	1938	1920
9035075	Kaisugu House, Kericho	-0.317	35.367	1939	2134
9035102	S.Kalya's Farm, Kedowa	-0.267	35.517	1946	2286
9035148	Koru Bible School	-0.200	35.267	1960	1707
9035150	Tinderet Estate	-0.133	35.383	1959	2134
9035199	Ainamoi Chiefs Camp, Kericho	-0.300	35.267	1960	1981
9035240	Keresoi Forest Station, Londiani	-0.283	35.533	1961	2682

9035256	Maragat Forest Station	-0.050	35.467	1965	2134
8935001	Kabagendui Kibet Farm	0.033	35.300	1920	1890
8935013	Nandi,Koisagat Tea Estate	0.083	35.267	1921	2073
8935033	Nandi Hills, Savani Estate	0.050	35.100	1929	1829
8935148	Kipkurere Forest Station	0.083	35.417	1959	2256
8935159	Cerengoni Forest Station	0.117	35.367	1964	2438
8935161	Nandi Hills,Kibweri Tea Estate	0.083	35.150	1958	2103
9035046	Chemelil Plantation	-0.067	35.150	1932	1229

The 30 days were selected randomly across the years considering days when rainfall was recorded at all the stations. This eliminated the skewedness that would be introduced by stations with no rainfall.

3.3 IDW and GPI interpolation

In these two methods, few decisions were made with regards to interpolation parameters. In both cases, the power *p* (equation 1) for the weighting function was varied whilst for IDW; the neighborhood search was also specified. Some tests were made and three different models for each interpolation procedure chosen as below:

IDW with *p=1* ; GPI with *p=1*
IDW with *p=2* ; GPI with *p=2*
IDW with *p=3* ; GPI with *p=3*

3.4 Kriging and cokriging interpolation

Kriging methods involve many decisions and specification of several parameters. The flow of activities is summarized below:

3.4.1 Model fitting

Kriging, like most interpolation techniques, is built on the basis that things that are close to one another are more alike than those farther away (quantified here as spatial autocorrelation). The empirical semivariogram is a means to explore this relationship. Pairs that are close in distance should have a smaller difference than those farther away from one another. The extent that this assumption is true can be examined in the empirical semivariogram. Empirical Semivariograms were computed for each of the 30 days. Depending on the shape of the variograms,

appropriate models were selected. Fitting a model was done by defining a line that provided the best fit through the points.

3.4.2 Neighborhood search

Kriging equations are defined by matrices and vectors that depend on the spatial autocorrelation among the measured sample locations and prediction location. The autocorrelation values come from the semivariogram model. The matrices and vectors determine the kriging weights that are assigned to each measured value in the neighborhood search. All stations were included in the neighborhood search for interpolation at any point in the basin.

3.4.3 Making a prediction

From the kriging weights and the measured values, predictions were computed for locations with unknown values. Maps were produced and cross validation parameters used to assess the accuracy of the predicted values.

3.5 Interpolator evaluation

For IDW and GPI, only two measures were possible i.e. Mean and RMSE. For the kriging methods, additional measures were used and a brief discussion of these is provided below:

Standardized mean (MS): is a measure of biasness, the mean prediction error should be near zero.
Prediction error = measured – predicted
Root-mean-squared prediction error(RMSE)

Average standard error (ASE): measure of variability in the predictions. If ASE >RMSE, then the system is overestimating variability and vice versa

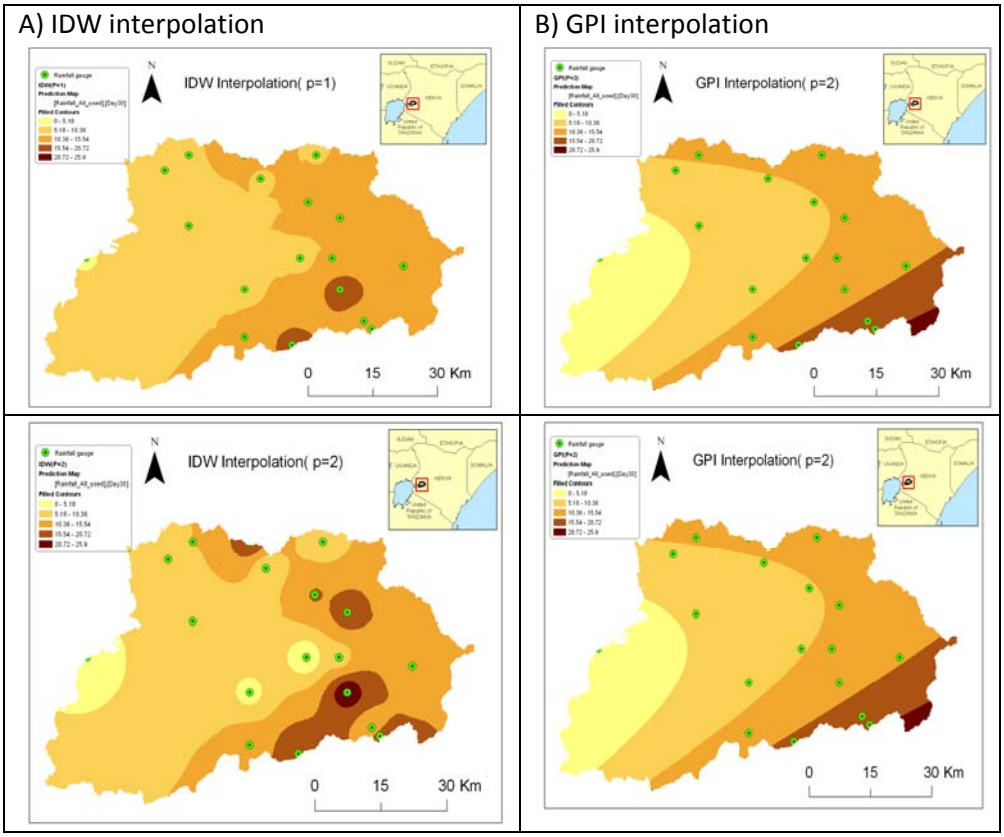
Standardized root-mean-squared prediction (RMSS): If the root-mean-squared standardised errors are greater than 1, then there's underestimating of variability in the predictions; if the root-mean-squared standardized errors are less than 1, then the system is overestimating variability in the predictions.

4.0 Results

Due to lack of adequate data, cross validation was used in which each of the stations was omitted from the interpolation, new values for the same station were derived and then compared to the true (observed) value. This was done for all the stations iteratively. Results presented as map and charts are for 30th April 1968 (day 30) while tables are used to present the remaining days.

4.1 IDW and GPI interpolations

The IDW method required two levels of decision to implement, the power *(p)* and neighborhood search. Figure 2 shows the IDW (A) and GPI (B) interpolated surfaces with *(p=1, 2, 3)* respectively. The south west section of the basin has only one station and this explains the uniform gradual decrease in rainfall around this region for all the interpolators. In contrast to the *p=1* interpolation, IDW interpolation with *p=2* produces smoother surfaces with gradual transition of rainfall. A "hull effect" (circle around the data) forms in the interpolated surfaces for the IDW. It was clearly demonstrated that the larger the value of *p* the larger the RMSE. Interpolation for the rest of the study period was implemented with *p=1* for IDW.



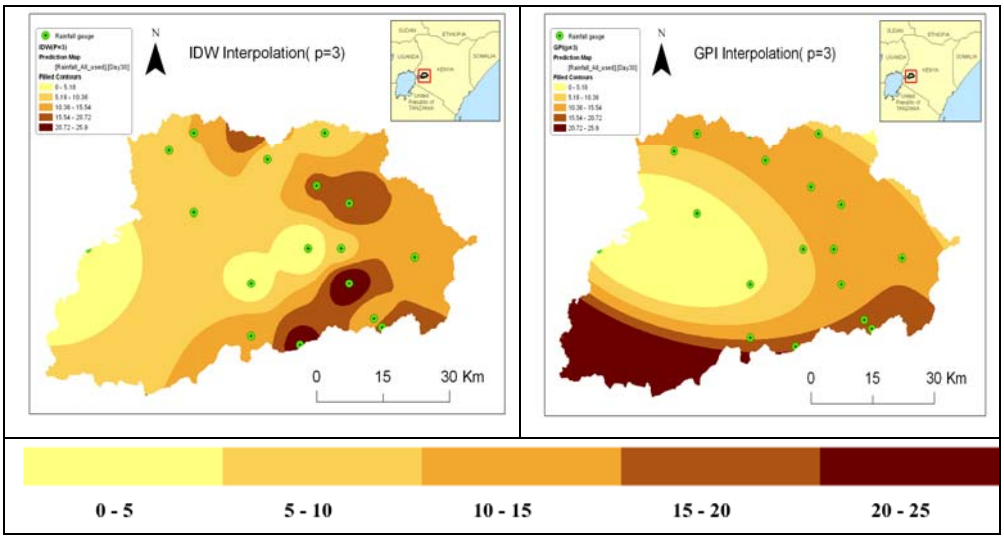


Figure 2: IDW (A) and GPI (B) interpolation surfaces generated using different values of power p. From top to bottom, p = 1, 2, 3

The GPI method required the least of decisions to implement; only power *(p)*. The interpolation method produced surfaces characterized by stripping and devoid of the hull effect. However, the transition from one value to the other is drastic and quite unrealistic. Unlike when *p=1*, the GPI interpolator with *p=2* tends to smoothen the surface and captures the trend in the rainfall observations, with the low amounts recorded to the south west location of the basin. Transition is extremely uniform and quite unrealistic. GPI with *p=3* produced the poorest results with the largest error (RMSE=11.37). The method over-estimates rainfall over the southern parts of the basin where the stations recorded the least amounts of rainfall. For the remaining interpolations, GPI was implemented using *p=2.*

4.2 Kriging

Figure 3 shows a scatter plot (prediction versus observations) for the kriging method on day 30.

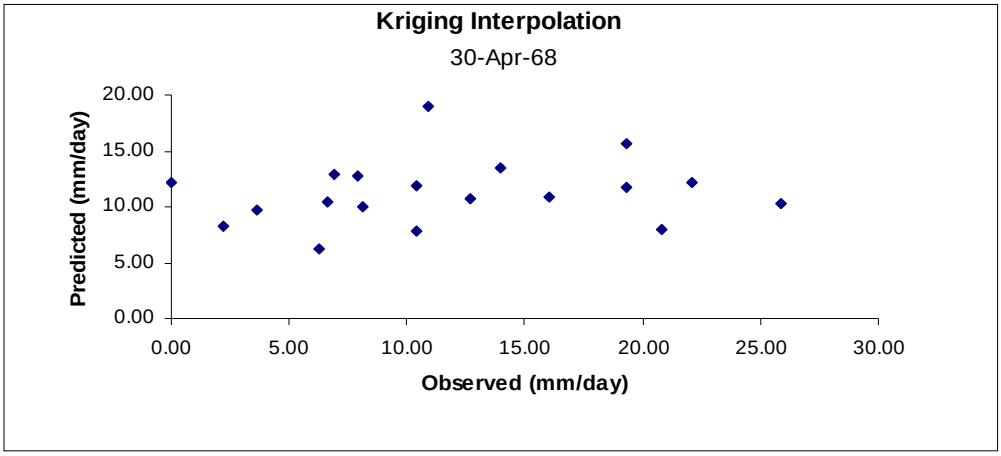


Figure 3: Kriging interpolation: predicted versus observed for 30th April 1968

The RMSE for the selected day 30 is 7.016, ASE=6.856 and Mean=0.009. A comparison was also done between the Gaussian and Spherical models (figure 4).

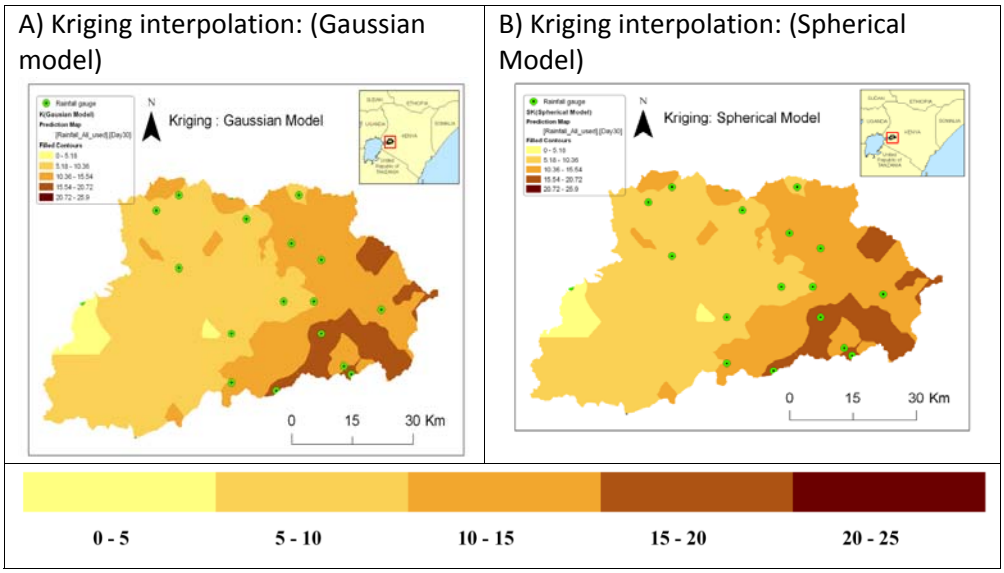


Figure 4: Kriging interpolation (A) using Gaussian model and (B) using spherical model

The difference between the two is minimal and in both cases the RMSE was the same. The Gaussian model was then used for the rest of the interpolations.

4.4 Cokriging

Cokriging works by coupling primary data and secondary data.

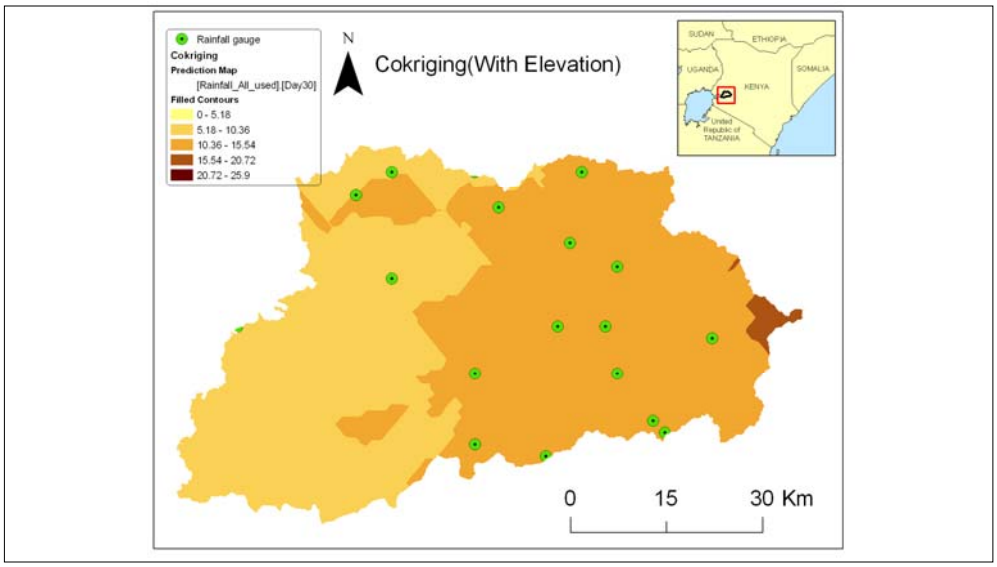


Figure 5: Cokriging interpolation (Day 30)

Figure 5 shows the output from the cokriging interpolator using gauge data coupled with 90m elevation data (DEM). The method produces similar results to kriging though it has a rugged spatial representation and gives the interpolation a more natural appearance. The interpolator does not reproduce high rainfall amounts but produces moderately averaged estimates.

4.5 Comparison between the Interpolation Techniques

Generally, the best model is the one that has the standardized mean nearest to zero, the smallest root-mean-squared prediction error, the average standard error nearest the root-mean-squared prediction error, and the standardized root-mean-squared prediction error nearest to one.

4.5.1 IDW versus GPI

The results produced by GPI are simple, and somehow unrealistic. The method does not reproduce the natural pattern that a rainfall event would exhibit. IDW on the other hand produces slightly better results but is highly affected by the hull effect around the stations. This is because more weight is given to the closest station. Comparatively, IDW produced better results than the GPI interpolator.

4.5.2 Kriging versus Cokriging

Comparison between the two kriging methods was done using error statistics. This is because the spatial surfaces produced were similar and visual interpretation would be insufficient. Figure 6 summarizes error parameters.

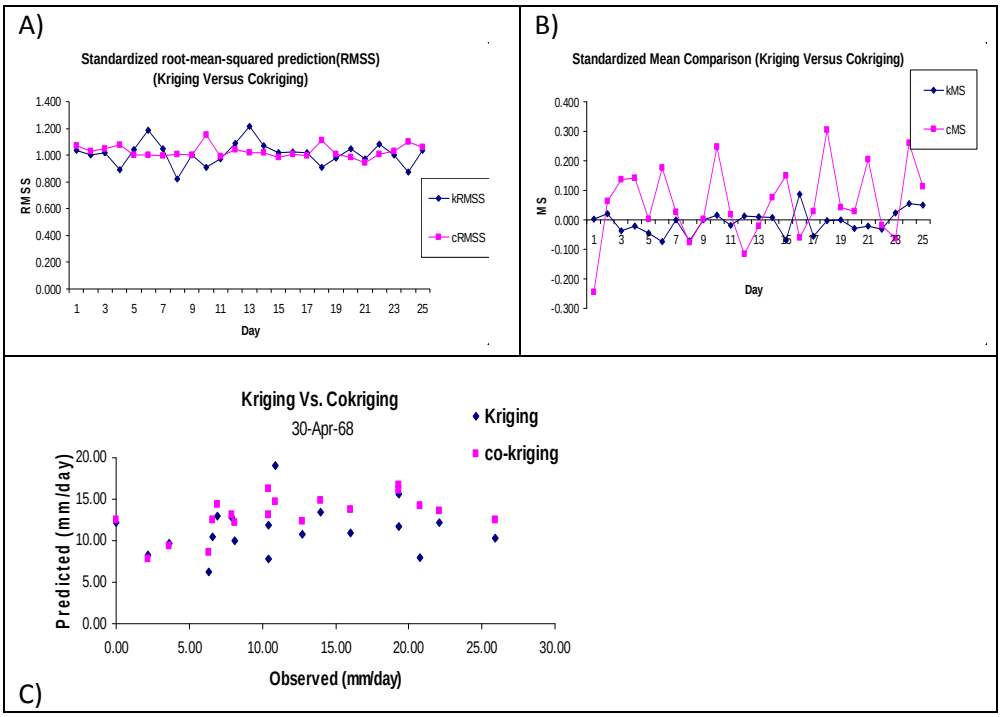


Figure 6: Comparison of RMSS (A) and RMSE (B) for kriging and cokriging; and (C) predicted versus measured rainfall for day 30

Cokriging produced better RMSE (cRMSE) (figure 6A) compared to kriging (kRMSE) for most of the days under investigation. This improvement is attributed to the inclusion of elevation data as secondary source of information. Kriging interpolation represents rainfall well with most of the standardized mean (Figure 6(B)) values close to 0 compared to cokriging whose values are above 0.1. In terms of variability, cokriging scores better than kriging because all its RMSS values are close to 1 as shown in figure 6(A). Figures 6 (C) shows a scatter plot for the two geostatistical interpolators for day 30.

4.5.3 Overall evaluation

Evaluation for all the interpolation techniques was only possible by comparing the common measures of reliability. Table 3 summarizes the RMSE for all the interpolators.

Table 3: RMSE values for all the Interpolators

Date	IDW	Kriging	Cokriging	GPI
5-Jan-88	6.88	7.31	6.00	45.00
14-Jun-91	10.71	9.62	8.70	51.24
13-Jun-91	12.88	13.20	11.13	53.09
1-Mar-62	5.82	5.23	5.23	53.14

18-Apr-73	7.44	5.82	5.24	25.08
25-Jun-91	3.42	3.77	2.58	50.23
21-Aug-70	5.60	5.18	4.86	25.13
19-Feb-62	5.18	4.34	5.43	44.46
24-Jun-91	6.69	5.71	5.07	50.08
17-Aug-70	8.15	7.84	7.77	26.13
13-Apr-73	1.23	1.14	1.09	23.69
6-Jan-88	5.39	5.14	4.57	44.25
26-Jan-62	0.95	0.99	0.84	52.84
9-May-73	9.76	9.21	8.64	37.56
14-Aug-70	6.33	7.69	6.75	27.49
15-Dec-81	9.75	9.72	8.11	35.48
21-Apr-73	8.45	6.41	5.83	25.16
23-Aug-70	5.36	5.00	5.26	26.08
22-Apr-73	8.59	7.51	7.22	25.85
28-Feb-62	2.28	2.05	1.80	43.91
19-Aug-70	10.54	9.71	8.10	27.51
25-Aug-70	8.58	7.60	7.06	26.09
4-Jan-88	2.31	2.06	1.93	43.54
24-Aug-70	7.31	6.53	6.66	26.28
Average	6.65	6.19	5.66	37.05

GPI produced the largest values of RMSE for all the days under investigation. IDW, Kriging and cokriging produced comparable results with cokriging producing the best results (least RMSE).

4.6 Discussion and conclusions

This study aimed at interpolating rain gauge observations into spatial surfaces using IDW, GPI, Kriging, and cokriging interpolators. Thirty days of study were selected and spatial surfaces generated. The performance of the interpolators was evaluated by visual inspection of the generated surfaces as well as analysis of reliability measures. GPI produced the poorest results with surfaces characterized by massive stripping and large values of RMSE. IDW on the other hand produced better results compared to GPI but was limited due to the “hull effect”. For this purpose, they are quite unsuitable for rainfall interpolation especially if the gauge stations are sparsely located. Though IDW produced better results than GPI, it placed more weight to the nearest stations. The cokriging interpolator was used by combining observations with a 90m DEM as secondary information. While it was impossible to differentiate between kriging and cokriging from the interpolated surfaces visually, evaluation of measures of reliability showed that cokriging produced the best results. It produced the least values of RMSE and the best values of the standardized root-mean-squared prediction error nearest to one, implying that cokriging represented variability well in the interpolation.

While these findings suggest better performance of co-kriging using elevation, further research is recommended to incorporate more days in the study and investigate the inherent relationship between rainfall and elevation. Research on the incorporation of other sources of secondary spatial data into cokriging (e.g radar) to improve the interpolations would be useful.

Acknowledgement
The authors would like to thank the department of Biomechanical and Environmental Engineering (BEED), Jomo Kenyatta University of Agriculture (JKUAT) for the provision of data. We appreciate the department of Geospatial Engineering and Geospatial Information Systems (GEGIS) for the provision of GIS tools. We are grateful to the editor and anonymous reviewers for their constructive and valuable comments which greatly improved the quality of this paper.

References

Ali M. Subyani (2005). Geostatistical study of annual and seasonal mean rainfall patterns in southwest Saudi Arabia. *Hydrological Sciences–Journal–des Sciences Hydrologiques,* **49**(5).

Borga M. and Vizzaccaro A. (1997). On the interpolation of hydrological variables: formal equivalence of multiquadratic. *Journal of Hydrology,* **195**, pp. 160–171.

Chen. D. and Johansson. B. (2003). The influence of wind and topography on precipitation distribution in Sweden: statistical analysis and modelling. *International Journal of Climatology*, **23**, pp. 1523-1535.

Englund E. J. and Weber D. D. (1994). Evaluation and comparison of spatial Interpolators. *Mathematical Geology*, pp. 589 -603.

ESRI. (2007). *ArcGIS 9.2 help*. Retrieved December 16, 2008, from ESRI: http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?

Eulogio Pardo-igu´ Zquiza (1998). Comparisons of geostatistical methods for estimating areal average climatological rainfall mean using data on precipitation and topography. *International journal of climatology, Int. J. Climatol.,* **18**, pp. 1031–1047

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

Isaaks E. and Srivastava R. (1992). An Introduction to Applied Geostatistics. New York: Oxford University Press.

Johnston K., Ver Hoef J., Krivoruchko. K. and Lucas. N. (2001). Using ArcGIS Geostatistical Analyst. Redlands CA: ESRI.

Marquínez J., Lastra J. and García, P. (2003). Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. *Journal of Hydrology.* **270**, pp. 1-11.

Nalder I. and Wein R. (1998). Spatial interpolation of climatic normals: test of a new method in the Canadian boreal. *Agricultural and Forest Meteorology,* **92**, pp. 211–225.

Vicente-serrano S.M., Saz-sánchez. M.A. and Cuadrat. J.M. (2003). Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Climate research.,* **24***,* pp. 161-180.
Sumner G. (1998). Precipitation: Process and Analysis*.* New York: John Wiley.

Swallow B., Onyango. L. and Meinzen-Dick. R. (2005). *Catchment Property Rights and the Case of Kenya's Nyando Basin.* Nairobi: World Agroforestry Agro forestry Centre.

Symeonakis E. (2008). A comparison of Rainfall estimation techniques for sub saharan Africa. *International Journal of Applied Earth observation Geoinformatics*.