Estimation of Spatial Rainfall Variability in a River Basin From Point Rainfall Data Using GIS

S. M. Wandera[1], Prof. S. K. Makhanu[2] & G.O. Makokha[3]

[1]**Asst. Lecturer of Civil Engineering),** Kimathi University College of Technology Dept. of Civil engineering smdondo2002@yahoo.co.uk

[2]**Professor of civil Engineering,** Masinde Muliro University of Science & Technology, Box 190 Kakamega. ksmakhanu@wust.ac.ke

[3]**Geomatic Engineer,** Kimathi University College of Technology **-** Dept. of Civil Engineering, Box 567-10100 Nyeri, makokha.godfrey@gmail.com

Abstract—

Rainfall information constitutes an input for water resources management and development. The users are climatologists, hydrologists, civil Engineers, water managers and decision makers. This paper presents a study performed to establish a few selected rainfall stations from the network in the Nzoia river basin. An attempt to solve the problem of spatial estimation of rainfall at ungauged points from data recorded by applying spatial technique. Quantitative analysis using data from the records of Kenya Meteorological Department (KMD) collected between 1989 and 2002 for selected years and months was adopted. Fifty three stations were sampled for the study among hundreds of stations in the basin. The data was analyzed using spatial experimental and correlational methods. The Kriging technique was applied. Idrisi32 and Arc-ViewGIS programs were used in analysis and simulation. Variograms and stream flows were the outputs from these softwares. The following results were obtained: spatial correlation of the rainfall data could be described by the Gaussian Kriging model, forty two of the fifty three stations sampled were optimal to represent the basin for recording rainfall and simulation of stream discharges using GeoSFM model. In the results, the number of rain stations in the basin may be reduced by twenty four percent while still adequately recording a representative rainfall data and may subsequently be used to predict the stream flows. It recommends that the same technique be extended to other river basins.

Key words: River catchment; Kriging technique; spatial analysis; stream flow model; optimal number of stations; model simulation,

Simon Mdondor, Department of Civil Engineering, Kimathi University College of Technology (corresponding author to provide phone: +2540721556662; e-mail: smdondo2002@yahoo.co.uke).

Prof.Sibilike K. Makhanu. Masinde Muliro University of Science & Technology.

interpolation, GeoSFM-model. Idrisi 32, Arc-ViewGIS

I. Introduction

When monitoring a spatial environmental phenomenon using a limited number of recording stations, deciding where to place the station is a fundamental task (Carlos G., et al.2002).Spatial analysis are used by introducing spatial correlation, based on Regionalized Variable Theory (RVT). In this theory a local systematic structure exist in the vicinity of any point within the sample space (McCuen, 1986). The structure of the correlation when established, it enhance the use of spatial trend to estimate the records at locations with missing data or uninstalled monitoring stations. The tendency to be most similar to one's nearest neighbors is quantified through measures of spatial autocorrelation and continuity (http://www.geog.uu.nl/gstat/).

The concern is primarily with environmental monitoring data whose patterns are driven by atmospheric processes (Monestiez P., 1993). In this paper, a study of the rainfall recording stations in the Nzoia basin has been examined by applying kriging analysis technique and GeoSFM stream flow model. The rainfall information thus collected constitutes a fundamental input for sizing reservoirs and networks of irrigation, for the estimation of runoff and infiltration. From this view-point, it appears that users of rainfall networks are various: climatologists, hydrologists, agronomists, hydrologists and engineers of civil (Jedidi K., et al.1999). These different users of network have a common requirement, that the collected data be of quality, available at all the points in the basin where they are representative of the spatial variability of the rain. The number and location of the rain gauge stations are set in such a way that they capture a representative rainfall over the entire region of concern (Mani and Panigraphy, 2001). Weather forecasts have traditionally been issued with the agricultural community in mind and so has been the development of social amenities (Kibiiy J.K., et al., 2005). It therefore follows that a number rain gauges are located in established areas and institutions, while vast areas have fewer rain gauges. The qualitative forecasts issued are to some extent useful to the agricultural community but rarely to experts in water resources development who may require quantitative data (Kibiiy J.K., et al., 2005).

A number of techniques are employed in spatial rainfall estimation. The Thiessen polygon technique has been used for a number of years but due to its inflexibility, other techniques taking cognizance of spatial distribution are gaining acceptance. Kriging technique is one of this technique. Kriging combines both the separation distances and point data recorded at a rain gauge station to depict the spatial trend of rainfall record (Davis, 1996). Phenomenon such as flood prediction and catchment water resources management requires quantitative rainfall. This kind of quantitative data is not readily available at all point for estimation of river discharges (Kibiiy J.K., 2005).

Monitoring of environmental phenomenon with a network of stations is not adequate without correlation of the stations to capture spatial variability of the event and demand for high spatial sampling rate (Maxim A.B.,et al., 2004). To capture spatial distribution using other techniques which are not GIS based will demand require impractically large number of recording stations. Quantitative rainfall measurement requires that the location of the rain gauge stations are set in such a way that they capture a representative rainfall over the entire region of concern (Mani and Panigraphy, 2001).

The rainfall data at ungauged points are modeled and linked to recording stations to suite the quantitative data requirements (Kibiiy J. K., et al 2005). The records of the Kenya Metrological Department indicate unequal distribution of rain gauges in the country. The concentration of the gauges in agricultural areas higher as compared to areas with less agricultural activities. In Nzoia river basin, Kitale has a higher concentration of rain gauges as compared to lower parts of the Nzoia river basin in Budala’ngi. It can be noted from the KMD (2004) records that isolation of areas lying outside the areas of higher concentration of rain gauges. Prediction of hydrological phenomenon like river flooding can be done with accurate rainfall data, however the data currently available has missing gaps both in spatial nature and time. The spatial correlation is not well understood link of the data from one station to another and at ungauged stations (Armstrong D., et al.1992). To be able to estimate the rainfall in ungauged points a form of interpolation model is applied.

This study intended to optimize the existing rain gauge sites in order improve the interpolation of measured rainfall at ungauged points it also aims to provide information useful to enhance water resources development and management in the study area. The study adopted the spatial analysis Kriging technique based on Geographical Information System (GIS) analysis techniques using the USGIS model for the Nzoia river basin. The study enhances the understanding of geographic approaches to study water engineering by developing new tools for the analysis of geographically referenced information (Mark D. M., et al., 1997). It is upon this background that the study of estimation of spatial rainfall variability in a river basin from point rainfall data using gis

2. METHODOLOGY

TheNzoia river basin (Figure1) covers an area of 12000 km^2 (Kibiiy J. K., et al. 2005). It has different topographical zones with distinct climatic conditions. Four zones were identified on basis of elevation and uniformity of climatic conditions. The zones were the slopes of Mount Elgon and Cherengani hills, Eldoret and Kitale plateau, the low land of kakamega and the border of Lake Victoria in Budala'ngi.

The daily rainfall records were obtained from the Kenya Meteorological Department (KMD). The mean rainfall data was used for spatial analysis, whereas the daily rainfall data for stream flow simulation.

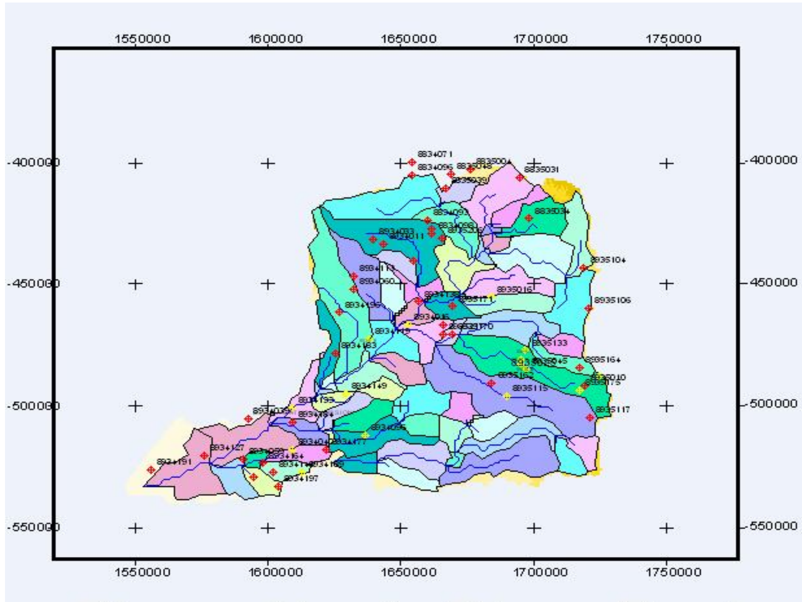


Figure1 The Nzoia river basin. (prepared from the Kenya DEM using Arc View GIS).

The study was conducted by a quantitative method. Fifty three raingauge stations were identified using purposeful sampling among the approximately six hundred existing stations. The selected stations had continues rainfall data in the years for which the study considered. The basin topographical map of 1:250,000 were obtained from Survey of Kenya.The maps indicated the georeferencing delineation of the basin and the point where the stations were located could be identified from the map.

Initially information of the rainfall and stations that were either installed and or operated in the basin between the years 1928 and 2003 was obtained from KMD in Nairobi. The information showed the names of the stations, their reference numbers, the georeferencing location in the latlong projection, year opened and whether the stations were operational or when they were closed. The following years were sampled for purpose of the study; (1988, 1989, 1990, 1997 and 1999).

The mean rainfall data was analysed for the five years, two years, annual and monthly.

The spatial distribution of rainfall was analyzed by variogram models that were adopted for interpolation. The analysis for spatial variability was performed using IDRISI 32 . The UTM format was suitable for surface analysis using Kriging technique that relies on the separation distances between sample points. The geostatistical analysis of the rainfall depth was carried out from the GIS module of the IDRISI 32. The geostatistical module has a spatial dependence modeler. It formed the first interface in Kriging analysis. Sample variograms were calculated for all directions (Omni directional variogram). The variograms were fitted to existing variogram models.

The best fitting model was identified by adjusting the nugget, range, sill and anisotropy factor of the chosen model. Model variogram was used to develop Kriged surface to predict spatial continuity.

The problem of ascertaining the optimum number of rain gauges in various basins is of statistical nature and depends on the spatial variation of rainfall. Thus, the coefficient of spatial variation of rainfall from existing stations was utilized for determining the optimum number of rain gauges. To identify the key station in a group of existing stations, first, the correlation coefficient between the average of storm rainfall and individual station rainfall was determined. The stations were then arranged in the order of their decreasing correlation coefficients, a station exhibiting highest correlation coefficient was referred to as first key station and its data was removed for determination of next key station. The procedure was repeated by considering the average rainfall of the remaining stations. The station indicating the highest correlation coefficient after removing the data of the first key station becomes the second key station. Similarly, the third and successive key stations are determined after removing the data of already selected key stations. Finally, the sum of squares of deviation of the estimated values of average rainfall from the actual rainfall in respect of 1st, 2nd, 3rd and successive key stations was determined and a graph was plotted of variance and the corresponding rain gauges. The plot can show a stationary point where increase in station population does not reduce a corresponding variance. The number of stations at that point were representative of the sub-basin (Mani and Panigraphy, 2001).

The stream flow model was performed by using Arc-view GIS. The module was used alongside other extensions and data on soil and land cover (Geostatistical stream flow model manual, 2005).

3. Results

The spatial location of the rain gauges is presented in Figure 2. The figure is a plot of the spatial rian gauge distribution of the Nzoia basin.

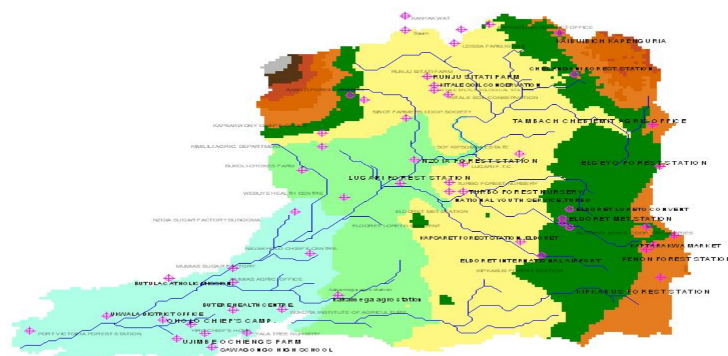


Figure 2 Spatial distribution of the rain gauges in the Nzoia river basin.

Source: Prepared from Nzoia river basin DEM and Rain Gauges location shape file using Arcview GIS

The Variation of color shades is an indication of the change in elevation from point to point.

(Y-axis V(x)- semivariance (mm)^2, X-axis Distance in Meters).

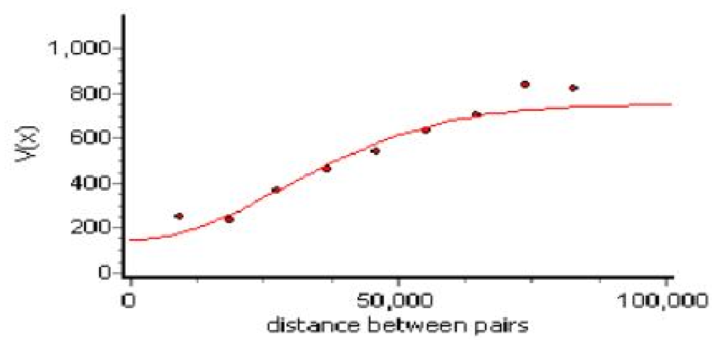


Figure 3 Best Fitting Gaussian Variogram for the three years mean rainfall data for all the fifty-three station in the whole Nzoia river Basin.

The variogram in Figure 3 was fitted to three years mean rainfall data for the entire basin. Table 2 presents a summary of the parameters that were obtained from experimental fitting of the models to surface rainfall data recorded at stations in the study area. It was observed that the range falls between 83km and 92km. A higher nugget effect was observed. The highest value of 200mm^2 was recorded for the two years mean rainfall. The Gaussian model was fitting in all the years considered.

Table 2 The best fitting models generated from the mean rainfall data for entire Nzoia basin.

Time	Model	Range ,a, (km)	Sill ,c, mm^2	Nugget ,co, mm^2
Three	**Gaussian**	**92**	**735**	**55**
Two	**Gaussian**	**90**	**750**	**200**
Annual	**Gaussian**	**83**	**1164**	**63**

The data for mean monthly rainfall was analyzed for the zones. The entire Nzoia river basin does not record rainfall instantaneouly hence the analysis based on the entire basin became impracticable when short period mean rainfall was considered. Table 1.presents the range, sill and nugget effect. It was therefore, logical to consider data for zones that were smaller than the entire basin. This was achieved by Zoning.

3.1 Zone I (The Upper Region of Nzoia River Basin)

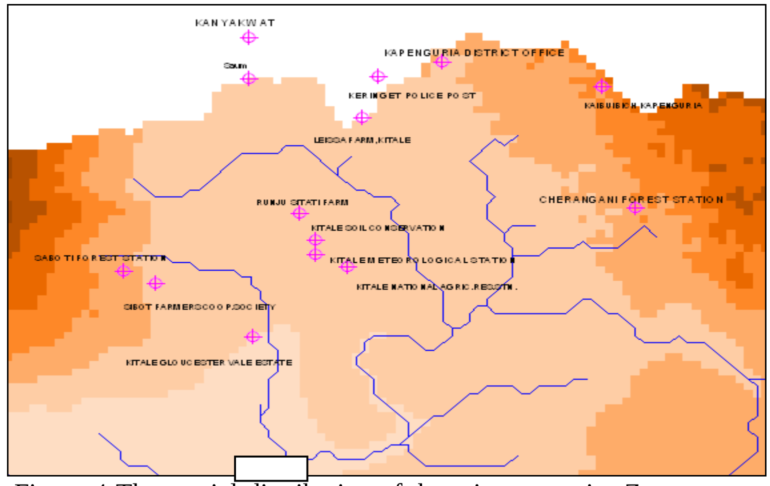


Figure 4 The spatial distribution of the rain gauges in Zone I.

Majority of the stations are located either at agricultural farming estates, meteorological stations or at forest stations. Five stations are located at the agricultural estates or farms. In this zone the highest density of rain gauge stations was in Kitale. The spatial distribution of the stations was presented in Figure 4. A total number of 14 stations were identified in this zone. Analysis indicates low spatial correlation of the rainfall mean data for the period exceeding a year.

It was noted that the average range, sill and nugget for all the months considered were 7.6km, 1.37mm^2 and 0.11mm^2 respectively. The results indicate much lower values of the range, sill or nugget as compared to the ones obtained for the entire basin. Gaussian model fited majority of the mean monthly rainfall.

Table 3 Summary for the variogram models and the parameters upper zone of the Nzoia river basin.

Time	Model	Range (a)km	Sill(c) mm^2	Nugget (co) mm^2
3 Years	**None**	**-**	**-**	**-**
2 Years	**None**	**-**	**-**	**-**
1 Year	**None**	**-**	**-**	**-**
March	Spherical.	**90**	**2.76**	**1.4**
March	Gaussian.	**92**	**1.43**	**0.0**

April	Gaussian.	**90**	**1.43**	**1.2**
May	Gaussian.	**58**	**1.01**	**1.7**
October	Gaussian.	**58**	**0.73**	**1.5**
November	Gaussian.	**91**	**1.32**	**1.13**
December	Gaussian.	**55**	**0.93**	**1.30**

It was noted that the rainfall spatial distribution obeys the Gaussian model. An inferential analysis was carried out for the historical data from the stations to identify key stations. Analysis conducted to investigate the rate of decrease of variability with increasing of the number of station used in the observation of spatial rainfall. Figure 5 shows a combination of the rate of decrease in variability of rainfall recorded with the increase of the number of raingauges. Therefore it was noted that in zone I a minimum number of nine stations were adequate for estimation of rainfall. In order to investigate the validity of these results a modeling process was undertaken by using the rainfall stations recommended above in simulating stream discharges. The third objective was to simulate the stream flows from the rainfall data of the selected key station. To accomplish this objective for Zone I the simulation was undertaken for the stream discharges a hydrograph was plotted as shown in figure 6. The result presented in Figure 6 shows that the discharge simulated for the streams in zone I were comparing well for selected stations and all the installed stations.

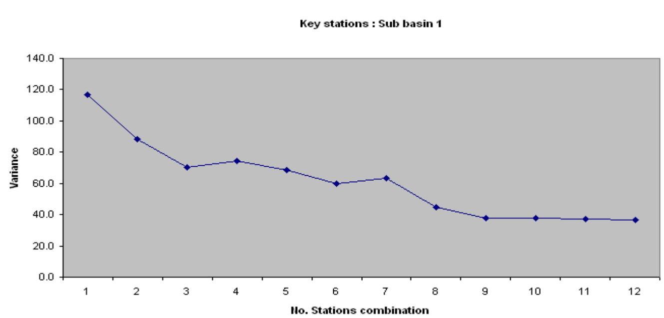


Figure 5 Spatial rainfall variance in Zone I

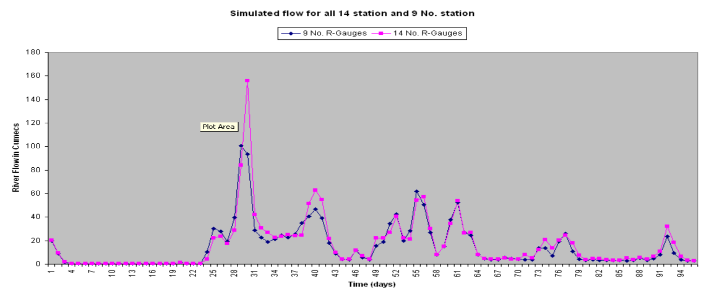


Figure 6 Simulated stream flows for 14 stations and 9 proposed stations generated by FEWS-SFM.

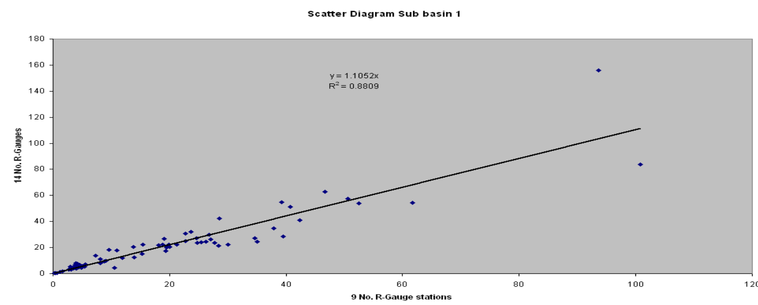


Figure 7 Scatter diagram zone I

The correlation value of 0.88 was obtained when the discharges in the streams for all the stations and the proposed optimum stations were compared. Figure 7 presents the relationship. It was recommended that Zone I was adequately represented by nine rain gauge stations. The same procedure was used for the other remaining Zones.

3.3 Zone II (Middle Plateau around Eldoret)

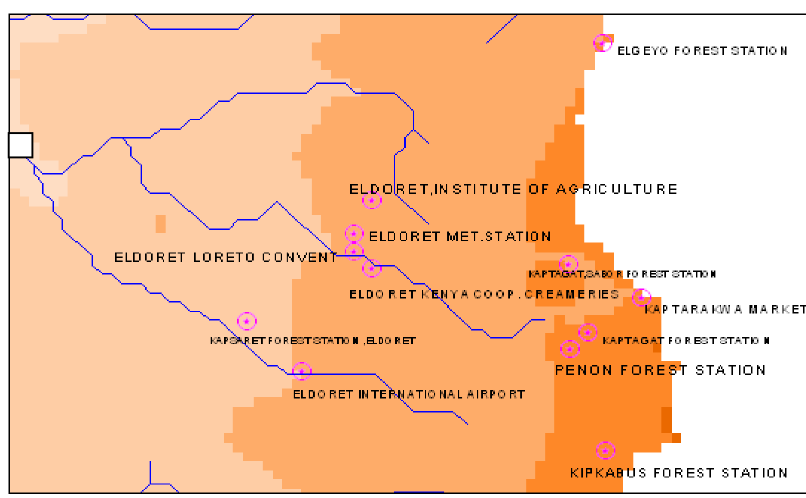


Figure 8 The spatial distribution of the rain gauges in Zone II.

Table 4 Summary for the variogram models and the parameters fitted for the Zone II.

Time	Model	Range (a)km	Sill(c) mm^2	Nugget (co) mm^2
3 Years	None	-	-	-
2 Years	None	-	-	-
1Year	None	-	-	-
March	Gaussian.	5	1.0	0.0
April	Gaussian.	6	4.2	1.0
May	Gaussian.	5	2.6	1.3
October	Gaussian.	6	3.8	1.2
November	Gaussian.	6	2.1	1.8
December	Gaussian.	5	3.6	0.9

It can be noted that Gaussian model was generally fitting to the monthly rainfall data for all cases considered. The range for the model was close for all the months considered as it falls between 5000 to 6000m. The sill for the mean monthly rainfall was not exceeding 5.0(mm)^2.

The rate of rainfall spatial variance decreasing until the introduction of the tenth station. Implying that ten rain gauge stations are adequate in Zone II.

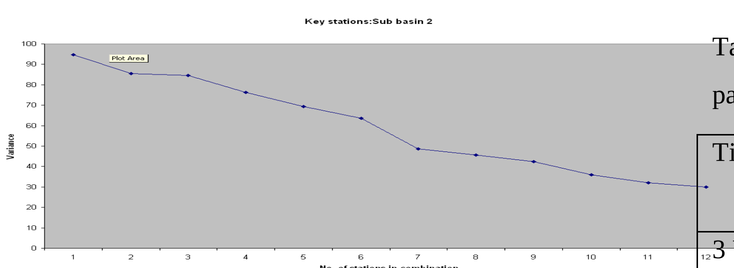
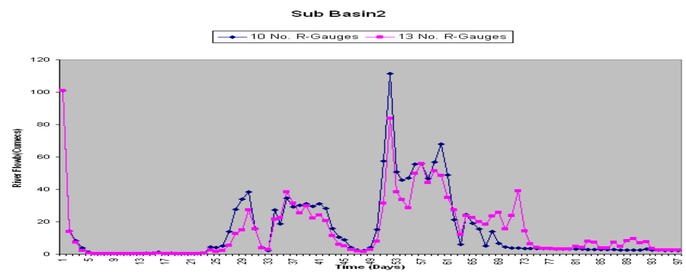


Figure 9 Spatial rainfall variance Zone II

Figure 11 The analysis produced correlation value of R^2=0.858.



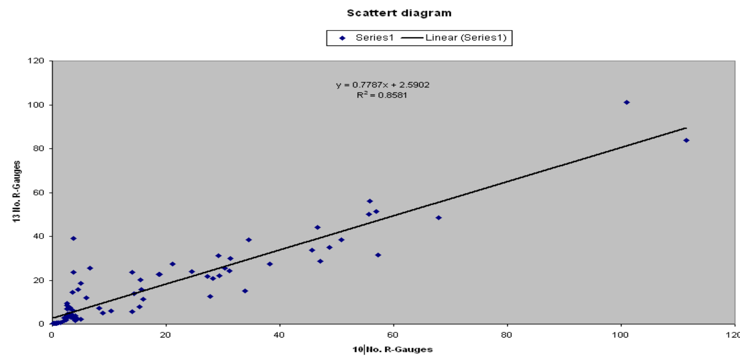


Figure 13 Scatter diagram zone II.

3.4 Zone III (Middle plateau around Kakamega)

It covered the middle plateau around Kakamega Spatial density of rain gauges plotted in figure 5-16.

Figure14 The rain gauges station identified for Zone III.

The Area lies on the western side of the Middle plateau of the Nzoia river basin.

Results of variogram analysis are presented figure 5-17 and 5-18.

Table 5 Summary for the variogram models and the parameters fitted for the Zone III

Time	Model	Range (a) km	Sill(c) mm^2	Nugget (co) mm^2
3 Years	Gaussian .	13	100	50
2 Years	Gaussian .	12	120	70
1Year	Gaussian .	14	110	97
March	Gaussian .	14	2.3	0.0
April	Gaussian .	15	2.025	0.0
May	Gaussian	16	2.6	1.2

	.			
October	Gaussian .	15	2.4	1.01
November	Gaussian .	15	2.07	0.98
December	Gaussian .	15	1.15	0.90

Zone III has a range falling between 14 and 17k m. The range is higher than in Zone II. The sill and nugget for a period exceeding a year are higher compared to the same attributes for the monthly periods. Selection of the optimum number of the stations in Zone III is presented in table 5-9. Simulation of river flows based on these stations and all the stations in the sub basin is given in figure 5-29. The key station analysis was carried out and the graph showing the rate of decrease of variance of spatail rainfall averages with increasing number of stations is given in figure 15

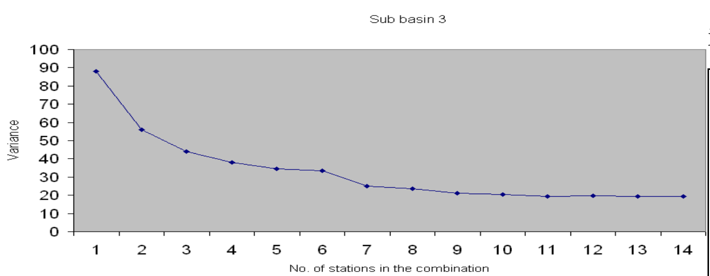


Figure 15 Spatial rainfall variance in Zone III of the Nzoia river basin.

Figure 15 Shows that ten key stations are adequate for spatial

The simulated stream flows for optimum number of stations in the zone is presented in Figure 16. The figure presents the daily variation of stream flows at the outlet of a selected sub basin in Zone III of the Nzoia river basin.

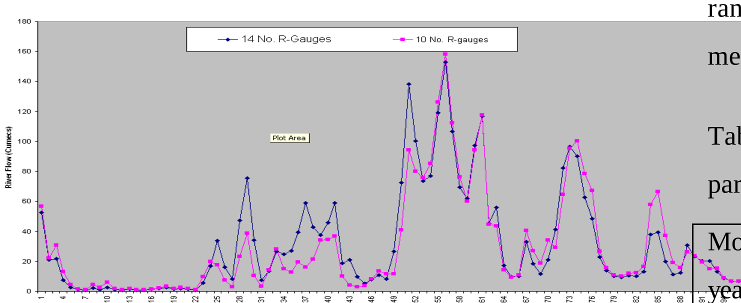


Figure 16 simulated stream flows generated by FEWS-SFM model for 14 stations and 10 proposed stations.

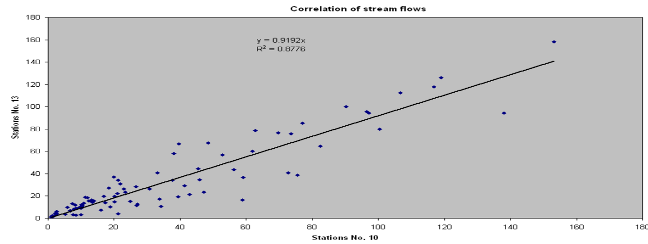


Figure 17 Scatter diagram Zone III

Correlation in Figure17 has a value of $R^2 = 0.8776$

3.5 Sub Basin IV (Lower Part of The Nzoia Basin around Lake Victoria)

This falls on the lower parts of the Nzoia river basin. It covers the area around Busia and Siaya Districts. Budalangi area which is prone to floods is also in this zone.

The spatial locations of rain gauges in zone IV is presented in Figure 5-19.

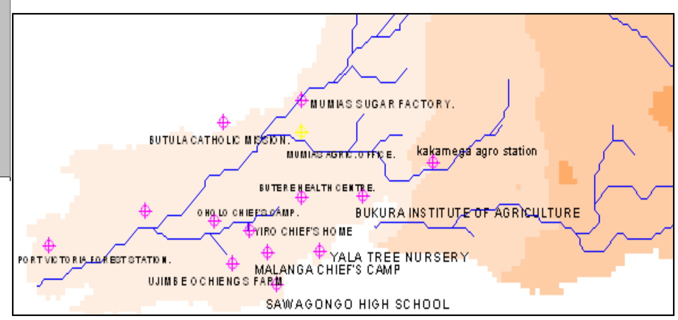


Figure 18 The spatial distribution of rain gauges station identified for Zone IV.

A summary of the fitted variograms is presented in the table 5-6. The Gaussian model fits in zone IV. The range is moderate with the lowest range of 9.5km and the highest range of 12km. None of the variograms model could fit to mean rainfall data averaged for a period exceeding a year.

Table 6 Summary for the variogram models and the parameters fitted for the Zone IV

Month/ year	Model	Range (a) km	Sill(c) mm^2	Nugget (co) mm^2

3 Years	None	-	-	-
2 Years	None	-	-	-
1Year	None	-	-	-
March	Gaussian.	12	9.0	5.0
April	Gaussian.	12	15.1	6.0
May	Gaussian.	108	12.0	3.0
October	Gaussian.	9500.3	11.0	5.4
November	Gaus.	10200.5	14.6	6.1
December	Gaus.	10000.4	8.6	7.3

The range falls between 9km and 12km in Zone IV.

.

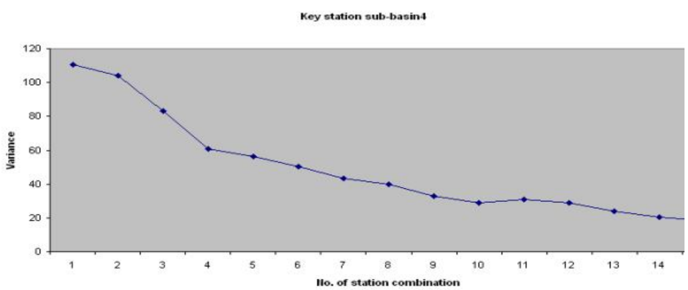


Figure 19 Spatial rainfall Variance in Zone IV

Figure 19 shows that variance decreases steadily throughout. The introduction of the 13th station shows a steady but lower rate of decrease. It suggests that the 12 stations may be considered adequate in estimation. It is reasonable to note that most of the installed stations contribute effectively in estimation of rainfall in this sub basin. This can noted that out of fourteen stations, twelve of them should be retained when estimating which is 86% of the total number of stations.

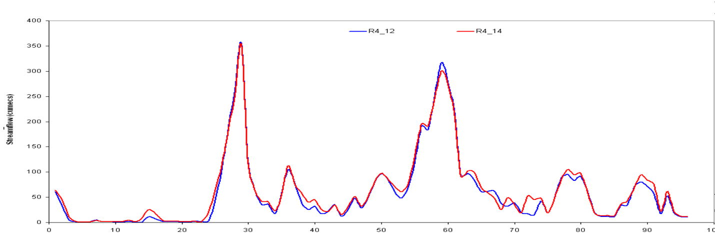


Figure 20 Simulated stream flows FEWS-SFM Model for 14 stations and 12 proposed stations.

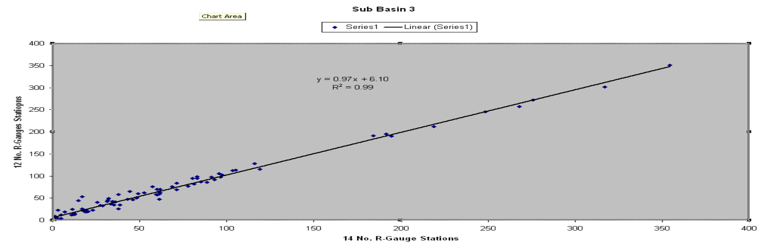


Figure 21 Scatter diagram for Zone IV of the lower parts of the Nzoia river basin.

Figure 21 shows the correlation of the simulated discharge values for the fourteen and twelve stations. The correlation analysis gives a value of R^2 =0.99 .The analysis results strongly points out that the zone requires almost all the stations in estimation of the rainfall. It should be noted that the rate of change of variance continues to decrease even after combining all the fourteen stations. This is a strong indication that an increase in the stations in the zone could lead to a better estimation of rainfall data. The zone IV requires additional stations to improve rainfall estimation. Kakamega requires additional rain gauges installation.

4. Discussions and conclusion

The study investigated the identification of representative rainfall observation stations in a network of existing stations. The aim was to select an optimal number of stations that may serve as a representative of all the stations in the Nzoia river basin. It was in relation to the need of collecting a more reliable rainfall data in all parts of the basin that may be used development and management of water resources. The study specifically sought to establish the spatial distribution pattern of rainfall. The distribution of the observed rainfall data over years was used to identify optimal rain gauge sites in the existing network, hence the most representative rain gauge stations in a network of existing stations from statistical analysis. The selected stations were finally used in modelling the stream flows and comparing the discharges based on the proposed optimal stations and sampled existing stations.

When data was analysed, it was noted that in a total of fifty three sampled stations in the Nzoia river basin proposed forty two were optimum for recording rainfall and subsequently for stream flow estimation. A further analysis indicated that the upper zone of the Nzoia river basin had fourteen stations out of which nine were recommended to adequately be used for recording rainfall and estimation of the stream flows.

In Zone II,. The rain gauges in this zone were crowded around Eldoret town. The crowding could lead to duplication of rainfall data recorded. The analysis of rainfall data in zone II indicates that ten stations could be used to adequately record the rainfall data.

Zone III of the Nzoia river basin lies in the area around Kakamega town. The zone had fairly better spatial distribution of rain gauges. The results from the analysis show that ten rain gauge stations may be adequate for rainfall recording. The correlation of the simulated stream discharges based on the optimal number of ten stations and all the stations used in the study for this zone is high and gives a value of (R^2=0.881). This is an indication of higher reliability of the records based on the proposed optimal stations. The lower parts of the Nzoia river basin borders Lake Victoria. This part of Nzoia river basin was covered under Zone IV. It covers parts of Busia and Siaya districts. The zone indicates a continuous decrease in variance in key station analysis. This implied that more stations were required in this zone to achieve a more reliable rainfall data recording and subsequent estimation of the stream flows. Although, it shows a higher correlation value in stream flows, it is reasonable to note that the key station analysis indicated that almost all the stations should be retained in this zone. In view of these findings, the study concludes that a total number of stations used to estimate spatial rainfall In the Nzoia basin can be reduced to a proposed optimum number of forty-two from the fifty-three sampled stations. The proposed optimum number of stations can be used both for recording rainfall and estimation of stream rate of discharge.

The spatial distribution of rainfall can be described by the Gaussian model with a range falling between 5-12 kilometres between the selected rain gauge sites. It is against this back ground that the following recommendations are made. Despite its limitation the study may be significant for water resources planning in engineering and agriculture. The study area is an agricultural zone in Kenya. The Nzoia river basin thus requires reliable rainfall information. Water resources planning entails both rainfall and river discharge data.

Basing generalisation on the study findings, it may be recommend that;

- The method used in optimizing the location of rain gauges in this study area may also be extended to other basins/sub basins in Kenya.
- The study should be extended to river gauging stations to optimize their location and applied in flood prediction.
- Further research should be conducted to establish the relationship of spatial rainfall distribution with change in topography and land use/cover.

5 REFERENCES

Armstrong D.,Renard J., Rivoirard P. (1992). Geostatistics for fish survey data.

Davis W. (1996). *The estimation of hydrological processes.* Water Resource. Res., 31(15),

Goodall, J. (2001). *Determining the spatial distribution of Benthos in the Western Artic Ocean, GIS in water resources,* Term project Presentation 2001.

Kibiiy J.K., Saenyi W., Kivuva J., Dulo S.M.,(2005). Flood and drought Forecasting and early warning program for Nile basin –UNESCO IHE. Kenya: Nairobi University.

Mani P.and Panigraphy (2001). *Precipitation Network Design for Pagladiya BasinNational Institute of Hydrology Center for Flood Management Studies(Brahmaputra Basin) Guwahatic Assam India.* Hydrology journal 24(3) 2001 33-38.

Mark D.M., Egenhofer M.J.,BianL.,Rogerson P., Vena J. (1997). Spatio-temporal GIS analysis for environmental health.

McCuen R.H and Synyder M.W (1986). *Hydrologic Modelling:Statistical Methods and Applications* . Georgia USA.

Microsoft Encarta Encyclopedia. (2005). *World Climatic Conditions*.

Monestiez P., Sampson P.D., Guttorp P. (1993). Modelling of heterogeneous spatial correlation structure by spatial deformation. USA:Washington University.

MOWD and JICA (1998) *The interim report national water master plan. Publication by MOWD,* Nairobi Kenya

Nile, proceedings of the 4th Kenya meterological society workshop on meterological resources application and services. mombasa Kenya.

Pebesma. E.J and C.G Wesseling (1998). *Gstat, aprogram for geostatistical modeling, prediction and simulation. Idrisi32 Giude*.