
2007

Modeling and Understanding the Relationship between Vegetation and Rainfall of a Tropical Watershed using Remote Sensing Data and GIS

Follow this and additional works at: <https://scholarsarchive.byu.edu/josh>

BYU ScholarsArchive Citation

(2007) "Modeling and Understanding the Relationship between Vegetation and Rainfall of a Tropical Watershed using Remote Sensing Data and GIS," *Journal of Spatial Hydrology*. Vol. 7 : No. 2 , Article 4. Available at: <https://scholarsarchive.byu.edu/josh/vol7/iss2/4>

This Article is brought to you for free and open access by the Journals at BYU ScholarsArchive. It has been accepted for inclusion in Journal of Spatial Hydrology by an authorized editor of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.



Modeling and Understanding the Relationship between Vegetation and Rainfall of a Tropical Watershed using Remote Sensing Data and GIS

Ron T. Chaoka¹, Berhanu F. Alemaw^{*1} and Demeke M. Tsige²

ABSTRACT

A model is developed to understand the relationship between satellite-derived NDVI and rainfall data in a large tropical catchment. Two Fourier-based modeling techniques with a seasonal component, viz. a seasonal model (SM) and a linear perturbation model (LPM) are tested, and their performance in reproducing the observed NDVI was evaluated. The methodology makes use of 15 years of 10-day composite time series data of rainfall and NDVI, which is estimated from NOAA-AVHRR data, both of which constitute concurrent data from 1982-96. The models are applied to a large catchment system of the Rufiji basin in Tanzania, with a network of 26 stations rainfall record and Thiessen polygon-interpolated spatially averaged NDVI data. The application of the SM model in forecasting NDVI and the LPM in relating NDVI and Rainfall at the 26 stations in the basin has been tested using the Nash and Sutcliffe (1970) model efficiency criterion. The linear perturbation model performed better than the simple seasonal model. The average model efficiency at the 26 stations considered during calibration and verification, are 0.64 and 0.54 for the LPM, and 0.62 and 0.49 for the SM, respectively. The approach can be used to improve our understanding of vegetation-rainfall relationships as well soil-vegetation-atmospheric processes, thus contributing to enhance hydrologic modeling of tropical watersheds.

(KEYWORDS: NDVI, rainfall, modelling, calibration, verification)

INTRODUCTION

Extensive research on land use and land cover changes have been conducted so far (Lambin *et al.*, 2001; Shrestha and Zinck, 2001), for it is one important environmental resource. The large-scale study results show that, due to increasing deforestation and agricultural intensification forest covers are rapidly decreasing. Deforestation has caused the loss of 1.16 billion ha of forest while farmland thereafter increased by 1.24 billion ha in only 300 years (Lambin *et al.*, 2001). However, if the research is measured by a different scale in different regions, results may be significantly different (Jianchu, 2005).

The importance and extend of vegetation information and monitoring for increasing our understanding of the potential of forest recovery for biodiversity conservation, the sequestering of carbon, as well as the environmental consequences has been studied (Mather, 1990; Zhang *et al.*, 2000; McNeely, 2003). Long-term population growth and economic development usually do not take place without intensification and agricultural growth where agricultural monitoring becomes important as stated in several works (Reynolds *et al.*, 2000; Lambin *et al.*, 2001; Tomich *et al.*, 2004).

Development of a model to reproduce observed vegetation dynamics and to analyse the temporal vegetation dynamics and its relationship with rainfall in the Rufiji basin are of particular interest to

¹Hydrogeology Programme, Geology Department, Private Bag 0022, University of Botswana,

²Hydrologic Consultant, Stewart Scott Ltd, PO Box 1066, Pietermaritzburg 3200, South Africa

*corresponding author's e-mail: alemaw@mopipi.ub.bw,

understand atmosphere–ecosystem interaction in a typical tropical climatic region of Africa. The time-series observations of the normalized difference vegetation index (NDVI) from the Advanced Very High Resolution Radiometer (AVHRR) on the NOAA-series of polar orbiting environmental satellites, which is computed from visible and near infrared spectral reflectance, are often used to monitor the vegetation dynamics (Kobayashi and Dye, 2005).

The monitoring of vegetation in southern Africa using AVHRR NDVI data has become increasingly important for various applications e.g. in monitoring large scale basin vegetation dynamics (Kobayashi and Dye, 2005), in the retrieval of quantitative properties of agricultural landscapes (e.g. Hill and Donald, 2003), in agricultural production (Diallo *et al.*, 1991; Reynolds *et al.*, 2000), and in climate change studies with implications for wildlife management and tourism (Sannier *et al.*, 1998).

For many years, the NDVI archive has been one of the most important remote sensing databases for monitoring the response of vegetation to weather conditions, by subtracting the current NDVI image from the average historical NDVI image of the same period. The resultant difference image shows regional vegetation anomalies and the relative severity of drought, but these difference images do not provide governments with quantitative crop assessments.

Since the launching of NOAA-7 satellite, a lot of studies have been carried out on utilising the remotely sensed information obtained from this pioneer satellite and all the subsequently launched satellites. Vegetation indices derived from (NOAA-AVHRR) sensor have been employed for qualitative and quantitative studies. Examples are (1) on the expansion of the Sahara desert (Tucker *et al.*, 1991), and (2) on the calculation of bio-physical parameter for climate models (Sellers *et al.*, 1995). The necessity to consider the effects of land cover changes in watershed modeling using remote sensing has been recognized (Tao and Kouwen, 1989; Duchon *et al.*, 1992). Also, the space platforms provide us with spatially distributed remotely sensed data to augment our limited, ground-based, point data to monitor land use changes. Semi- and fully distributed hydrologic models, designed to take advantage of these spatially distributed remote sensing data, should be able to quantify basin hydrologic processes more accurately than before.

Thus the main objective of this study is investigation of the relationship between vegetation cover, available from monitored remote sensing data as NDVI and rainfall by applying system type models and developing a model that can be useful for monitoring vegetation cover in the Rufiji basin of Tanzania (Figure 1). The Rufiji valley presents an interesting paradox that it is an area of good agricultural potential, but also suffers from recurrent droughts and famines. Understanding of the spatial distribution of vegetation and its link with rainfall is also an important aspect to further enhance our understanding of the soil-vegetation-atmospheric processes and to develop a process-based hydrologic model of this tropical large drainage basin.

This paper addresses two main objectives: (i) investigation of the relationship between vegetation cover and rainfall by applying Fourier-based models to model the seasonal components of the variables studied, and (ii) evaluation of the performance of the models and make recommendation on an appropriate model type to improve our understanding of vegetation-rainfall relationships in the Rufiji basin. A detailed account of the methodology used to achieve the stated objectives along with description of data used in the study is provided in the following sections.

DATA AND METHODS

The study area

The Rufiji river basin lies between longitudes 32.5° and 40 ° East and latitudes 5.5° and 10.5° South. It is the largest river basin in Tanzania (Figure 1) covering an area of about 177,000 km² and extending 700kms from Mbeya region in the West to the Indian Ocean in the East. The Rufiji river basin constitutes three main river sub-basins viz., the Great Ruaha, the Leivegu, and the Kilombero with sizes of 47%, 18%, and 20% of the total area of the Rufiji basin.

The basin is situated dominantly in the semi-arid belt, which runs from North to South through the central portion of Tanzania. The land cover of the study area is dominated by 92 scattered forest reserves in the Rufiji Basin (MOA, 1987), in which 47.5% is forest and woodland, 43.0% is bush land and grassland, 6.5% is cultivated, and the remaining 3% is open land and water bodies.

There are three general physiographical zones the coastal belt which has elevation ranging from sea level to 1000m and apparently located along the coast and nearby, the central plateau with elevation ranging from 1000 to 1500 and the high lands with elevation of 1500 and above. The general climate of the basin ranges from tropical humid heat of the coastal belts to the temperate condition in the southern highlands. The wet season is between November and April with most of the rain falling during December and January and the dry season, which lasts generally from May to October. The rainfall is always in the form of showers with high intensities. The range of monthly temperature values is very low. The prevailing winds are the Easterly from the Indian Ocean. In the transition month between seasons a westerly wind can also predominate.

NDVI and rainfall data

At each rain gauge station, we created the Thiessen polygons spatially averaged by extracting the 10-day composite Pathfinder AVHRR Land (PAL) NDVI data (James and Kalluri, 1994). The PAL NDVI data are calculated from reflectance values after corrections for ozone absorption and Rayleigh scattering. The Normalized Difference Vegetation Index (NDVI) data are available at 8 km grid resolution. The distribution of the 26 rain gauge stations and the relative areas delineated by Thiessen polygon to determine the average 10-day composite NDVI data corresponding to each rain-gauge station and the 10-day composite rainfall record, which are extracted from daily rainfall records, are shown in Figure 1. The Thiessen polygons were produced according to the algorithm and software code presented by Eastman (2001).

The potential use of the AVHRR for vegetation monitoring was more realised after the satellite NOAA -7 became operational in 1981, and this study was conducted using NDVI data available for 1982-96 and corresponding rainfall record.

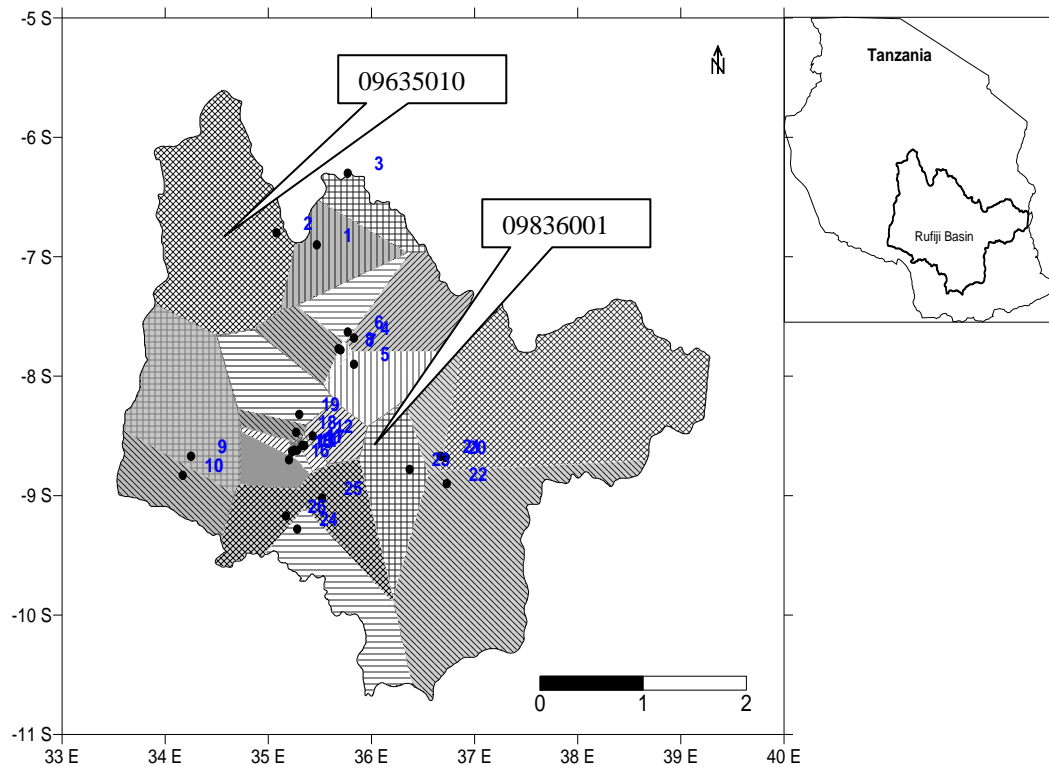


Figure 1 Location of the Rufiji basin showing rain-gauge sites with the Thiessen polygon subdivisions for NDVI spatial interpolations (The numbers correspond to the respective station codes shown in Table 1)

Data preparation

Due to uneven spatial distribution of the rainfall stations in the basin as seen in Figure 1, generalising of the outcomes of the analysis at a basin scale is too difficult. It was therefore prudent to consider Thiessen polygons for areal averaging the NDVI over a point rain gauge station. This gave us consistent results compared to using individual pixels that correspond to a rain gauge and associated NDVI. We also attempted to divide the basin into 100 blocks each having 30 pixels and the average rainfall on each block were inputs for case of the application of linear perturbation model (LPM). The average rainfall is interpolated from Isohyetal maps of the basin, developed through kriging algorithm (Cressie, 1990), which was re-sampled for each of the 30 pixels block surrounding a rain gauge station, where the corresponding 10 day composite time series data of NDVI was retrieved.

No significant difference in result or improvement was obtained using the later approach compared to the Thiessen polygon average results. This perhaps was due to the pertinent forest and bush land cover types that dominant the study area. A more reasonable approach could be to use less number of pixels in a block such as 3x3 pixels surrounding each rainfall station, which could improve results. This however is a further aggregation of the original AVHRR data of 1.1 km resolution at the satellite sub-point, which was further re-sampled to Global Area Coverage Resolution (GAC), which the current study relies on.

In the retrieval of NDVI time series data that associated with a give rain gauge, the area covered by water bodies and wetlands was masked out and averaging was only within the remaining area of a given polygon.

One of the limitations of the study was that NDVI specific to the various land uses was not considered. This was perhaps with a fair assumption that the predominant land cover is forest and bush land in the case of the Rufiji basin. Otherwise, it is imperative to consider and underscore the need to accurately incorporate them in hyperspectral analysis and vegetation analysis applications of remote sensing (e.g. Melesse and Jordan, 2003; Thenkabail, 2004; Thenkabail et al., 2004).

METHODOLOGY

The methodology involves developing Fourier-based models to monitor vegetation from historical NDVI data and rainfall. Two models are optimised, calibrated and validated using 15 years of 10-day composite time series NDVI and rainfall data for the period 1982-1996. The first set of 9 years data (1982-1990) was used for model calibration, with 36 decadal time series per annual cycle. Another set of the data of the last 6 years (1991-96) was used for model verification. The two models tested are the seasonal model (SM) and the linear perturbation model (LPM) highlights of which are discussed in the following sections.

The seasonal model (SM)

The first modelling technique is the systems approach, called here as the seasonal model (SM) in which NDVI is modelled with the premise of seasonal fluctuations, which takes the form of Fourier model. In this approach, the assumption is that during a year in which the rainfall is identical to its seasonal expectation, the corresponding NDVI is also identical to its seasonal expectation. It is ideally with the assumption that NDVI between years does not change significantly, which might be true for forest land covers and highly seasonal prevailing rainfall patterns. This assumption is fairly reasonable as the predominant land cover is forest and bush land in the case of the Rufiji basin.

The NDVI, $Y(t)$ is considered as manifestation of aggregate effects of a long-term seasonal mean (S), and an error term (ξ), all of which vary with time, as:

$$Y(t) = S(t) + \xi(t) \quad (1)$$

The seasonal cycle, using decadal or 10-day time scale, is modelled by Shutler's periodogram (Matalas, 1967), which is modified for single harmonic case as:

$$S(t) = \Theta + A\sin(2\pi t/L) + B\cos(2\pi t/L) \quad (2)$$

In which $L/2$ is the half period or half a year in our case, θ is the mean, and A and B are parameters of a single harmonic given by:

$$A = \frac{2}{L} \sum_{t=1}^L S(t) \cos(2\pi t / L) \quad (3)$$

$$B = \frac{2}{L} \sum_{t=1}^L S(t) \sin(2\pi t / L)$$

The seasonal expression of y on a dekad d , Y_d to be modelled as $S(t)$ in Equation 2, is obtained as:

$$Y_d = \frac{1}{L} \sum_{r=1}^L y_{d,r} \quad (4)$$

Where $y_{d,r}$ is the observed NDVI on dekad d (i.e. 1 to 36), in the year r and L is the number of years in the calibration period.

The values are changing from dekad to dekad, and a model represented by 36 dekads of 10 days composite time series remains non-parsimonious. The seasonal model considers smoothed function of the above Y_d series using a Fourier function. The above model does not account for small fluctuations or departures from the seasonal mean values, which are further smoothed by Fourier function.

The linear perturbation model (LPM)

The departures from the long term mean, between the observed NDVI, $Y(t)$ and the seasonal mean Y_d values modelled as $S(t)$ as determined in the Seasonal Model (SM) may not be insignificant. However, in all other years, unlike the SM, when the rainfall and the NDVI values depart from their respective seasonal expectations, these departures series are assumed to be related by a linear time invariant system. This is explained by a multiple linear regression model that can be used to model the possible persistence effect of the deviations. We borrowed the concept from a similar technique known as the Linear Perturbation Model as originally known, which was developed by Nash and Barsi (1983) for rainfall-runoff transformation. It was applied in this study in order to reduce the dependence on the linearity and increase the dependence on the observed seasonal behaviour of vegetation and rainfall.

The model is based on the following assumptions:

- (i) If the expected value of a 10-day composite rainfall and NDVI on dekad d are denoted by I_d and Y_d respectively, then the input I_d of rainfall gives the output Y_d of NDVI, each of which are modelled as decadal expectations based on the smoothed mean given by the first harmonic Periodogram given (Equation 2).
- (ii) Perturbations, or departures from the expected input value (I_d), are linearly related to the corresponding perturbation from the expected output (Y_d). The expected values of rainfall (I_d) and NDVI (Y_d) are smoothed mean values of the first harmonic Periodogram of Equation 2.
- (iii) For departures from the seasonal mean of the NDVI and rainfall values of $Z_i = Y_i - Y_d$ and $X_i = I_i - I_d$ respectively, a multiple linear regression model, can be written as:

$$Z_i = \sum_{j=1}^m X_{i-j+1} h_j + e_i \tag{5}$$

where $i = 1, 2, 3, \dots, n$ is the number of observations, and h is the transfer function ordinates, same as the multiple linear regression coefficients with index $j=1, m$, and m is the memory length. The memory length could be considered an indicator of the lag effect in phenology, i.e. by which the green flush follows rainfall as it is a factor that defines the relationship between rainfall and NDVI departures from their seasonal mean values.

If the operation is represented by Equation 5, then the memory length, m must be determined by the least-squares estimation method. Having determined the seasonal component Y_d , based on single harmonic Fourier function, the model (Equation 5) is calibrated by the use of ordinary least squares (OLS) where by the memory length is determined accordingly. In this method, an assumed memory length, m , is used to solve coefficients h_j in Equation 5, and model accuracy or during optimization the objective function in terms of R^2 and MSE is computed. Sequentially increased values of m and corresponding objective function values in the iterative process provides an m value which is selected based on a minimum value of the objective function in terms of both R^2 and MSE .

In optimising the memory length through solving Equation 5 based an OLS, the Nash and Sutcliffe (1970) model efficiency (R^2) criteria and the mean-square error (MSE) were used. The former is

analogous to the coefficient of determination used in linear regression analysis. According to Equation 6, in essence the model efficiency (R^2) (Nash and Sutcliffe, 1970) indicates the residual variance unaccounted for by a model. It is one minus the ratio of the residual variance (F) or the mean square error (MSE) to the initial variance (F_o) given by:

$$R^2 = 1 - \frac{F}{F_o} = 1 - \frac{MSE}{F_o} \quad (6)$$

With

$$F_o = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \quad (7)$$

Another measure of model performance used is the root mean-square error (RMSE) which is the square root of the mean-square error (MSE), which measures the sum of squares of deviations between the observed and modelled results given by:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (8)$$

In expressions (7) and (8), y_i is the observed and \hat{y}_i estimated time series of NDVI departures, n is the total length of NDVI values over the calibration period which represents the number of dekads i.e. 1982-90, and \bar{y} is the mean of the y_i series over the calibration period.

RESULTS AND DISCUSSION

As a rationale of using the seasonality of NDVI and rainfall and their seasonal relationship between them, Thiessen polygon-based spatially averaged NDVI time series data was retrieved from a network of 26 rain-gauge stations in the Rufiji basin (Figure 1). spatially weighted decadal or 10 days composite time series NDVI data and the corresponding rainfall series at a station 09635010, located at 35.47°E and 6.90°S, are shown in Figure 2, in which marked seasonal variation is evident.

The apparent nature of seasonality of these two variables has definitely contributed for the good results that could be obtained using the proposed models: the seasonal model (SM) and linear perturbation model (LPM).

By virtue of this the LPM and Seasonal Model (SM) are expected to have better performance and comparison between them is also made and the results are summarised in Table 1 for the model calibration and verification periods. The number of harmonics is fixed to one as shown in the Shutler's periodogram (Equation 2).

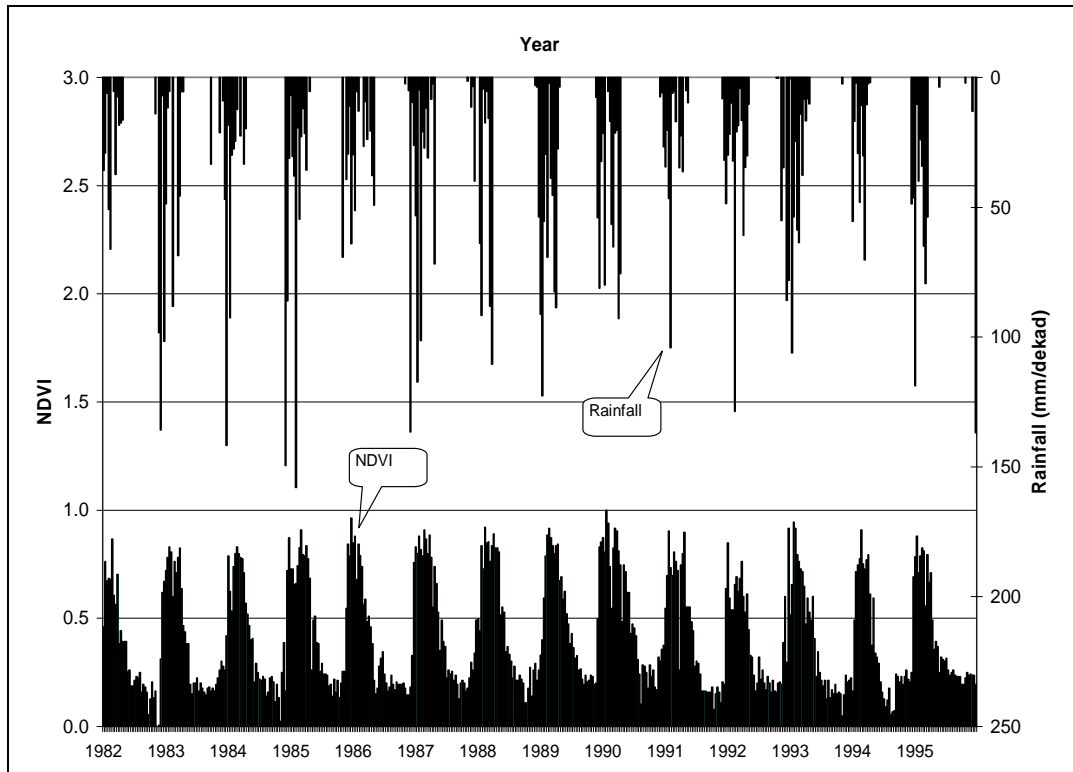


Figure 2 Relationship between 10-day composites of spatial averaged NDVI and rainfall data for a station 09635010

According to the analysis on the 26 rainfall stations considered to be appropriate the Linear Perturbation Model (LPM) has performed better in expressing the rainfall-NDVI relationship. Applying the LPM up to 0.84 in the calibration and 0.82 in the verification period efficiencies are observed at station 09835011 and 09635010 respectively. The analysis on the application of Seasonal Model and Linear Perturbation Model for the 26 stations is summarized in Table 1. A summary of the average R^2 for the 26 stations is shown in the last row. The optimized memory length m , in Equation 5, is also shown for each station of Table 1. On the other hand, Table 2 shows the root mean-square error (RMSE), between the observed and modelled results of the two models for the 26 stations considered.

When summarised from the above tables, in the calibration period the Linear Perturbation Model has performed best on 18 stations and the Seasonal Model on the remaining 8 stations. Considering the verification period on 12 stations the Seasonal Model is better than the Linear Perturbation Model. In most of the stations considered, where there were superior results of Seasonal Model in the calibration period there were also accompanying good results for the verification period of model application.

On the average, model efficiency in terms of R^2 at the 26 stations considered for the calibration and verification periods are 0.64 and 0.54, for the LPM. The corresponding calibration and verification period results for SM are 0.62 and 0.49, respectively.

Table 1 Comparison of the R² results of LPM with the corresponding results of SM for calibration and verification periods at the 26 stations

No.	Station Code	Location		Memory length, m (dekad)	Calibration R ² (1982-90)		Verification R ² (1991-96)	
		Long. (°E)	Lat. (°S)		SM	LPM	SM	LPM
1	09635010	35.47	6.90	4	0.71	0.79	0.52	0.82
2	09635011	35.08	6.80	12	0.59	0.84	0.46	0.54
3	09635014	35.77	6.30	10	0.86	0.81	0.60	0.78
4	09735002	35.83	7.68	12	0.74	0.83	0.56	0.73
5	09735007	35.83	7.90	4	0.71	0.77	0.77	0.63
6	09735013	35.77	7.63	10	0.74	0.75	0.56	0.74
7	09735014	35.70	7.78	3	0.65	0.76	0.77	0.66
8	09735015	35.68	7.77	4	0.65	0.78	0.77	0.53
9	09834008	34.25	8.67	1	0.56	0.74	0.64	0.60
10	09834010	34.17	8.83	5	0.63	0.71	0.68	0.71
11	09835009	35.33	8.58	7	0.68	0.49	0.61	0.42
12	09835019	35.43	8.50	1	0.73	0.40	0.49	0.37
13	09835021	35.25	8.62	10	0.67	0.72	0.68	0.55
14	09835022	35.28	8.62	4	0.69	0.30	0.59	0.38
15	09835024	35.23	8.63	6	0.67	0.80	0.68	0.50
16	09835025	35.20	8.70	10	0.52	0.58	0.18	0.62
17	09835034	35.35	8.58	1	0.68	0.46	0.61	0.61
18	09835036	35.27	8.47	2	0.66	0.33	0.62	0.24
19	09835039	35.30	8.32	2	0.64	0.48	0.54	0.15
20	09836001	36.72	8.68	1	0.56	0.62	0.15	0.14
21	09836002	36.67	8.67	15	0.56	0.61	0.15	0.40
22	09836003	36.73	8.90	10	0.48	0.65	0.44	0.65
23	09836011	36.37	8.78	13	0.81	0.69	0.39	0.60
24	09935004	35.28	9.28	3	0.21	0.40	0.04	0.50
25	09935007	35.52	9.02	2	0.31	0.64	0.14	0.76
26	09935009	35.17	9.17	1	0.27	0.64	0.02	0.37
Average of 26 stations considered					0.62	0.64	0.49	0.54

From Table 2, it can be noted also that model performance in terms of RMSE provides accuracy evaluation in fitting the two models. The average root mean-square error of the two models, the seasonal model and the linear perturbation model, for the calibration period is found to be 0.029 and 0.028, respectively. The corresponding average root mean-square error of the two models for the verification period is found to be 0.029 and 0.028, respectively.

Table 2 Comparison of the root mean-square error (RMSE) results of LPM with the corresponding results of SM for calibration and verification periods at the 26 stations

No.	Station Code	RMSE Calibration (1982-90)		RMSE Verification (1991-96)	
		SM	LPM	SM	LPM
1	09635010	0.027	0.025	0.029	0.023
2	09635011	0.030	0.024	0.030	0.029
3	09635014	0.023	0.025	0.028	0.024
4	09735002	0.027	0.024	0.029	0.025
5	09735007	0.027	0.026	0.024	0.027
6	09735013	0.027	0.026	0.029	0.025
7	09735014	0.029	0.026	0.024	0.027
8	09735015	0.029	0.025	0.024	0.029
9	09834008	0.030	0.027	0.027	0.028
10	09834010	0.029	0.027	0.026	0.026
11	09835009	0.028	0.032	0.028	0.031
12	09835019	0.027	0.033	0.030	0.031
13	09835021	0.028	0.027	0.026	0.029
14	09835022	0.028	0.034	0.028	0.031
15	09835024	0.028	0.025	0.026	0.030
16	09835025	0.031	0.030	0.033	0.028
17	09835034	0.028	0.032	0.028	0.028
18	09835036	0.028	0.034	0.028	0.033
19	09835039	0.029	0.032	0.029	0.034
20	09836001	0.030	0.029	0.034	0.034
21	09836002	0.030	0.029	0.034	0.031
22	09836003	0.032	0.029	0.030	0.027
23	09836011	0.025	0.028	0.031	0.028
24	09935004	0.035	0.033	0.035	0.029
25	09935007	0.034	0.029	0.034	0.025
26	09935009	0.034	0.029	0.035	0.031
Average of 26 stations considered		0.0290	0.0284	0.0292	0.0285

Figure 5 shows the distribution of errors in the LPM model in relation to the observed 10-day composite of NDVI for the model calibration and validation set of data at a typical station no. 09836001.

Even though, numerical evaluation of the performance of the two models indicate that Linear Perturbation Model (LPM) performed better than Seasonal Model (SM), as seen from the above table (Table 1) and figures (Figure 3 and Figure 5), there are still some errors which are not yet accounted for, and the improvement obtained using LPM over the SM seems insignificant. This issue takes one to an argument that a model will not be taken for granted by the primary model accuracy criteria alone

without further error diagnostic checks of the model. Such diagnostic checks are used to examine if any systematic errors are not unaccounted for due to any deficiencies / inadequacies of the proposed model structure.

Three diagnostic checks of model adequacy testing namely seasonality, non-linearity and persistence are normally conducted in hydrological literature (e.g. Kachroo, 1992), which might be conducted through the visualization of these features from the error diagrams, such as the one presented in Figure 5. As can be seen in Figure 5, the series of residuals arising from the application of Linear Perturbation Model seem not to portray marked seasonality, linearity or persistence which has not been accounted for by LPM. Therefore, there are no deficiencies in terms of mathematical formulation of the proposed model structure, which LPM could have incorporated to cater for additional seasonality or linearity of the relationship between rainfall and vegetation in the Rufiji basin.

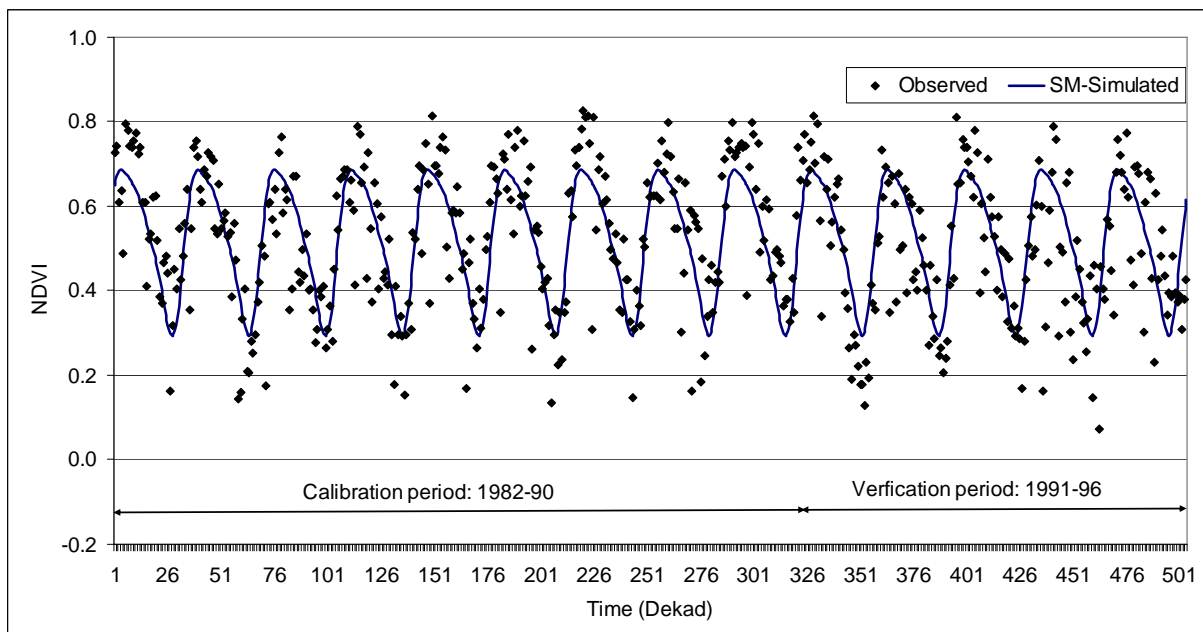


Figure 3 Comparison between observed and seasonal model-estimated NDVI at station 09836001

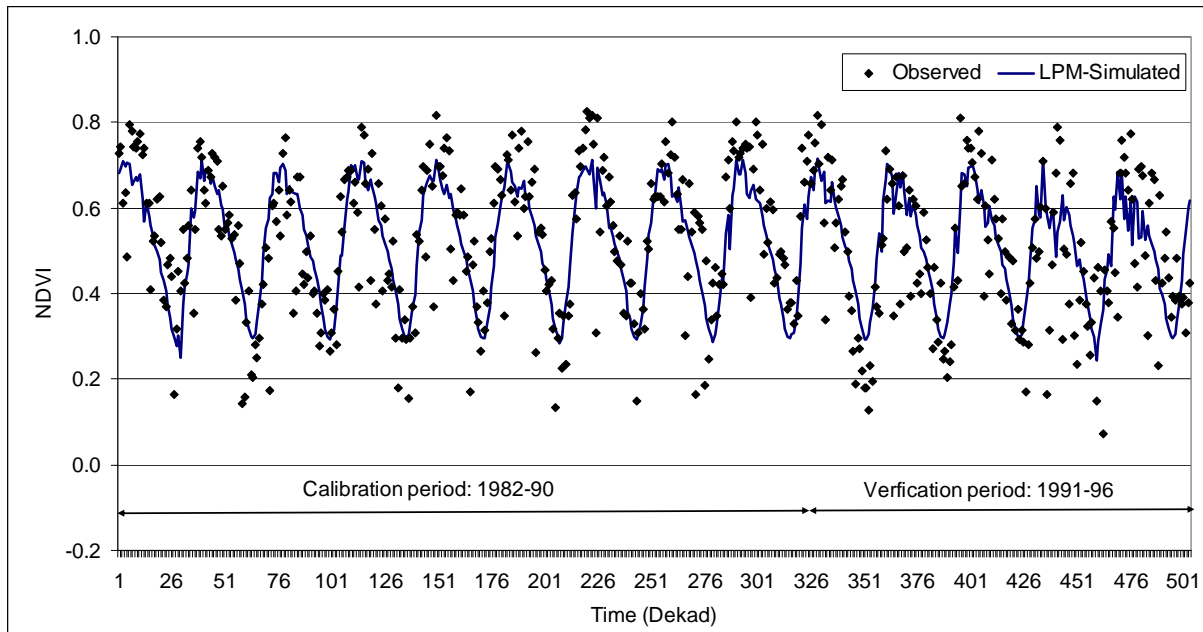


Figure 4 Comparison between observed and LPM-estimated NDVI at station 09836001

For real-time or timely rainfall data gathered and availed, it might be possible to monitor vegetation in the Rufiji basin using the Linear Perturbation Model. This effort should be considered to complement real-time monitoring of vegetation directly from remote sensing data. Moreover, the approach can be useful to infill missed information for the periodic gap until remote sensing data is archived, processed and used for availing vegetation and NDVI information. This approach can be useful for drought monitoring and famine early warning applications. It is important to reiterate that if the models are recalibrated for crop-specific vegetation cover, the approach can also be potentially useful for crop monitoring and specifically LPM can be considered appropriate model type for use in the monitoring of vegetation.

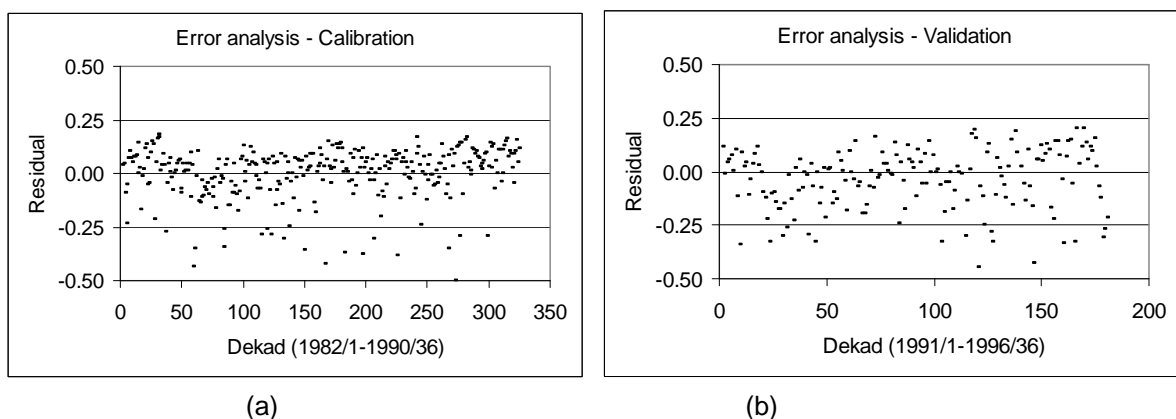


Figure 5 Distribution of errors in LPM model-simulated NDVI in relation to the observed 10-day composite of NDVI for the model (a) calibration and (b) validation set of data at station 09836001

ACKNOWLEDGEMENTS

Thanks are extended to the team of NDVI datasets distributors and the Goddard Space Flight Center for NOAA/NASA Pathfinder Land data. This research was partly funded through a German Government Scholarship, the Deutsche Academic Exchange Programme (DAAD), which the authors are indebted. Some support was provided from ORD-R528 during the course of developing the paper. The authors appreciate the support of the University of Dares Salaam, for hosting the MSc research of the third author and guidance of the second author. The first author improved the literature and diagnostic error analyses. Special thanks are due to the anonymous reviewers for their comments and suggestions on earlier versions of the manuscript which improved the quality of the manuscript.

REFERENCES

- Cressie, N.A.C. (1990) The Origins of Kriging. *Mathematical Geology*, 22, 239-252.
- Diallo, O., Diouf, A., Hanan, N.P., and Prevost, Y. (1991) AVHRR monitoring of savana primary production in Senegal, West Africa; 1987-1988. *International Journal of Remote Sensing*, 12(16), 1259-1279.
- Duchon, C. E., Salisbury, J.M., Lee Williams, T. H., and Nicks, A. N. (1992) An example of using Landsat and GOES data in a water budget model. *Water Resource Research*, 28(2), 527-538.
- Eastman, J.R. (2001) *Guide to GIS and Image Processing - IDRISI32, Volume 1. Idrisi Source Code*, Clark Labs, Clark University, Worcester, MA, pp. 171
- Hill, M.J., and Donald, G.E. (2003) Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series, *Remote Sensing of Environment*, 84, 367–384.
- James, M.E., and Kalluri, S.N.V. (1994) The pathfinder AVHRR land data set: An improved coarse resolution data set for terrestrial monitoring, *International Journal of Remote Sensing*, 15(17), 3347–3363.
- Jianchu, X., Xihui, A., and Xiqing, D. (2005) Exploring the spatial and temporal dynamics of land use in Xizhuang watershed of Yunnan, southwest China, *International Journal of Applied Earth Observation and Geoinformation*, 7: (2005) 299–309.
- Kobayashi, H., and Dye, D.G. (2005) Atmospheric conditions for monitoring the long-term vegetation dynamics in the Amazon using normalized difference vegetation index, *Remote Sensing of Environment*, 97, 519– 525.
- Lambin, E.F., Turner II, B.L., Geist, H., Agbola, S., Angelsen, A., Bruce, J.W., Coomes, O., Dirzo, R., Fischer, G., Folke, C., George, P.S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E.F., Mortimore, M., Ramakrishnan, P.S., Richards, J.F., Skaanes, H., Steffen, W., Stone, G.D., Svedin, U., Veldkamp, T., Vogel, C., and Xu, J. (2001) Our emerging understanding of the causes of land use and -cover change. *Global Environmental Change*, 11, 261–269.
- Matalas, N.C. (1967). Time series analysis. *Water Resources Research* 3: 817-829.
- Mather, A. (1990) *Global Forest Resources*, Bellhaven, London, pp. 30–38.

McNeely, J., 2003. Conservation in areas of armed conflict. In: Price, S. (Ed.), War and Tropical Forests. Haworth, New York, pp. 9–11.

Melesse, A.M. and Jordan, J.D. (2003) Spatially distributed watershed mapping and modeling: Part 1. Thermal maps and vegetation indices to enhance land cover and surface microclimate mapping. *Journal of Spatial Hydrology*. 3(2), 1-29

MOA (1987) Basic Data on Agriculture and Livestock Sector, Government of Republic of Tanzania, Ministry of Agriculture, Publication No. 1887.

Nash, J.E., and Barsi, B.I., (1983) A hybrid model for flow forecasting on large catchments. *Journal of Hydrology*, 66, 125-137.

Nash, J.E., and Sutcliffe, J. (1970) River flow forecasting through conceptual models. Part I. A discussion of principles. *Journal of Hydrology*, 10, 282-290.

Kachroo, R.K. (1992) River flow forecasting. Part 1: A discussion of the principles, *Journal of Hydrology*, 133, 1-15.

Reynolds, C. A., Yitayew, M., Slack, D.C., Hutchinson, C.F., Huete, A., Petersen, M.S. (2000) Estimating crop yields and production by integrating the FAO Crop Specific Water Balance model with real-time satellite data and ground-based ancillary data, *International Journal of Remote Sensing*, 21(18), 3487-3508.

Sannier, C.A.D., Taylor, J.C., Du Plessis, W., and Campbell, K. (1998) Real-time vegetation monitoring with NOAA- AVHRR in southern Africa for wild life management and for food security assessment, *International Journal of Remote Sensing*, 19(4), 621-639.

Sellers, P., Los, S., Justice, C., Dazlich, D., Collatz, G., and Randall, D. (1996). A revised land surface parameterization (SiB-2) for atmospheric GCMs. Part 2: The generation of global fields of terrestrial biophysical parameters from satellite data, *Journal of Climate*, 9, 706-737.

Shrestha, D.P., and Zinck, J.A. (2001) Land use classification in mountainous areas: integration of image processing, digital elevation data and field knowledge, *International Journal of Applied Earth Observation and Geoinformation*, 3(1), 78–85.

Tao, T. and Kouwen, N. (1989) Remote sensing and fully distributed modeling for flood forecasting. *Journal of Water Resources Planning Management Division of ASCE*, 115(6), 809-823.

Thenkabail, P. S. (2004). Inter-sensor relationships between IKONOS and Landsat-7 ETM+ NDVI data in three ecoregions of Africa. *International Journal of Remote Sensing*, 25(2), 389– 408.

Thenkabail, P.S, Enclona, E.A. Ashton, M.S., and Van Der Meer, B. (2004). Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sensing of Environment* 91, 354–376.

Tucker, C.J., Dregne, H.E., and Newcomb, W.W. (1991) Expansion and Contraction of the Sahara Desert from 1980 to 1990, *Science*, 253, 299 - 302.

Tomich, T.P., Cattaneo, A., Chater, S., Geist, H.J., Gockowski, J., Lambin, E.F., Lewis, J., Palm, C., Stolle, F., Valentim, J., van Noordwijk, M., and Vosti, S.A. (2004) Balancing agricultural development and environmental objectives: assessing trade-offs in the humid tropics. In: Palm, C.A., Sanchez, S.A., Vosti, P.J., Erickson, P.A., Joo, A.S.R. (Eds.), *Slash and Burn: The Search for Alternatives*. Columbia University Press, New York.

Zhang, P.C., Shao, G., Zhao, G., Le Master, D.C., Parker, G.R., Dunning Jr., J.B., and Li, Q. (2000) China's forest policy for the 21st century, *Science*, 288 (5474), 2135–2136.