Hydrol. Earth Syst. Sci. Discuss., 8, 3599–3629, 2011
www.hydrol-earth-syst-sci-discuss.net/8/3599/2011/
doi:10.5194/hessd-8-3599-2011

Hydrology and Earth System Sciences Discussions

Classification and flow prediction in a data-scarce watershed of the Equatorial Nile region

J.-M. Kileshye Onema[1,2], A. Taigbenu[1], and J. Ndiritu[1]

[1]Water Research Group, School of Civil and Environmental Engineering, University of Witwatersrand, Private Bag 3, Witwatersrand, WITS 2050, South Africa
[2]WaterNet, Box MP 600 Mount Pleasant, Harare, Zimbabwe

Received: 31 March 2011 – Accepted: 1 April 2011 – Published: 14 April 2011

Correspondence to: J.-M. Kileshye Onema (jmkilo3@yahoo.co.uk)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Continuous developments and investigations in flow prediction are of interest in watershed hydrology especially where watercourses are poorly gauged and data are scarce like in most parts of Africa. Thus, this paper reports on two approaches to generate local monthly runoff of the data-scarce Semliki watershed. The Semliki River is part of the upper drainage of the Albert Nile. With an average annual local runoff of 4.622 km^3, the Semliki watershed contributes up to 20% of the flows of the White Nile. The watershed was sub-divided in 21 subcatchments (S3 to S23); eight physiographic attributes from remotely sensed acquired datasets and limited ground information were generated for each subcatchments and used to forecast monthly volumes. One ordination technique, the Principal Component Analysis (PCA) and the tree clustering analysis of the landform attributes was performed to study the data structure and spot physiographic similarities between subcatchments. The PCA revealed the existence of two major groups of subcatchments. Multi-linear and polynomial regressions were the two modeling approaches used to predict the long-term monthly mean of discharges for the two types of subcatchments identified in the Semliki watershed. The ranges of multiple R, the multiple R^2, and the adjusted R^2 for the multi-linear and the polynomial models were, respectively 0.96–0.99; 0.93–0.99 and 0.92–0.99. The linearity assumption provided less accurate predictions.

1 Introduction

Numerous approaches exist for flow prediction in natural river reaches. Flow forecasting has significant interest both from research as well as from an operational point of view. The choice of methods depends on data availability and the type of application. While continuous developments strive at enhancing our predictive capability for streamflow, we are often faced with the challenge of predictions in ungauged basin (Sivapalan et al., 2003). Reliable and accurate estimates of hydrologic components

HESSD

8, 3599–3629, 2011

are not only important for water resources planning and management but are also increasingly relevant to environmental studies (Schröder, 2006). Several studies have reported on the use of catchment descriptors and regionalization of parameters for flow prediction in ungauged basins. Among the most recent studies are those of Sefton and Howard (1998), Mwakalila (2003), Xu (2003), Merz and Blöschl (2004), McIntyre et al. (2005), Sanborn and Bledsoe (2006), Yadav et al. (2007), Sharda et al. (2008), Kwon et al. (2009) and Shao et al. (2009). In their comparison of linear regression with artificial neural networks, Heuvelmans et al. (2006) indicated the need for well-informed choice of physical catchment descriptors as a first condition for a successful parameter regionalization. Cheng et al. (2006) reported on the importance and usefulness of parsimonious models for runoff prediction in data-poor environment as these models are characterized by few numbers of parameters. Reducing uncertainty associated with predictions in ungauged basin is critical as reported by Uhlenbrook and Siebert (2005), Koutsoyiannis (2005a,b), as well as Zhang et al. (2008).

Lately, Koutsoyiannis et al. (2008) indicated that analogue modeling techniques for simulation are also used for prediction with impressive performance due to the advances achieved with non linear dynamical systems (chaotic systems). The major drawback is the fact that these approaches are data intensive and work as black boxes, thus no process insight is provided.

Relevant spatial and temporal scale to flow prediction continues to be a subject of discussions and investigations in watershed hydrology (Kundzewicz, 2007). This paper reports on multi-linear and polynomial regression modeling approaches for flow prediction in a medium size watershed of the Equatorial Nile region. These approaches provide tools that will ultimately contribute towards better water resources planning in a humid and data-scarce environment.

2 Study area

These analyses are conducted within the Semliki watershed of the Equatorial Nile region (Fig. 1). The catchment studied covers an estimated area of 7699 km^2.

The Semliki drains the basins of lakes Edwards and George, and a contributing area downstream that includes the western slopes of the Ruwenzori range. The watershed receives an average rainfall of 1245 mm per annum, with peaks occurring in May (95 mm) and October (205 mm). An average annual local runoff of 4.622 km^3 has been estimated from records at Bweramule (Sutcliffe and Parks, 1999). The elevations comprise flat areas and ice-caped mountains, climbing up to 4862 m above the sea level. The flora and the fauna of the watershed constitute one of the unique and distinct ecosystems of the Albertine Rift region. The vegetation predominantly comprises medium altitude moist evergreen to semi deciduous forest. Five distinct vegetation zones have been documented under the mount Ruwenzori and they occur with changes in altitude. Detailed information on landscape physiographic attributes is reported in Kileshye Onema and Taigbenu (2009).

3 Methods and materials

The landscape of any catchment is made up of several combinations of physiographic attributes. These combinations are usually variable among catchments, giving rise to different hydrological responses. Table 1 presents the eight physiographic attributes extracted from the 21 subcatchments that form the Semliki watershed (S3–S23) (Fig. 2). The remotely sensed acquired NDVI were extracted and processed from decadal NOAA-AVHHR maximum composite imagery with the image display and analysis software WinDisp 5.1. The satellite derived rainfall were provided by the National Oceanic and Atmospheric Administration (NOAA) through the Famine Early Warning system Network (FEWS-Net). Details about the processing of remotely sensed acquired datasets is reported in Kileshye Onema and Taigbenu (2009). The 90 m Digital

HESSD
8, 3599–3629, 2011

Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM) along with traditional maps were used to extract the remaining landform properties. Twenty eight (28) years of historical flow measurement was used to determine the monthly average volume.

The Principal component analysis (PCA) as Indirect Gradient Analysis in association with clustering analysis was used as the exploratory technique to study the structure of the data. Both approaches were further utilised in the identification of similarities among sub catchments. Multi linear and polynomial regressions were performed on the eight physiographic attributes selected to derive the models as illustrated in Fig. 3 except that the normality test was not performed for polynomial optimization.

The coefficient of multiple correlation (*Multiple R*), *R*-square (R^2) that defines the coefficient of multiple determination and which measures the reduction in the total variation of the dependent variable due to the (multiple) independent variables, together with the adjusted R^2 were used to assess the accuracy of the two models produced through multi-linear and polynomial regressions.

4 Results and discussions

4.1 Principal components analysis (PCA)

The descriptive statistics below (Table 2) show that the variables included in the PCA are measured at significantly varying scales, however this does not affect the results of the analysis as the matrix analyzed is the scale invariant correlation matrix of the variables (as opposed to the covariance matrix).

The correlations between the variables are summarized in Table 3. There are some high correlations (greater than 0.5), implying that there is a correlation structure that can potentially be modeled or further explored using PCA. If all the correlations were low there would be no need to try to model the correlation structure using principal components analysis.

HESSD
8, 3599–3629, 2011

The value of phi for this data (0.4) (Table 4) suggests that there is considerable redundancy or complexity in the group of variables which warrants further examination using PCA. Bartlett's sphericity test is used to test the null hypothesis that the correlation matrix of the group of variables is a zero identity matrix i.e. none of the variables are correlated. If we obtain a p-value for the Bartlett's test which is greater than 0.05 we should not carry out PCA. The p-value obtained (Table 4) is very low indicating that we can carry out the PCA.

According to the Kaiser criterion when the principal components have been calculated using correlation coefficients is to retain the principal components with an eigenvalue > 1. Therefore, based on the results shown in Table 5, we would retain the first 3 principal components. These 3 principal components account for 76% of the variation in the data.

The Eigenvectors are the coefficients that relate the scaled original variables to the derived factors. The scaled original variables are defined as follows:

$$x_i = \frac{X_i - \mu_i}{\sigma_i} \tag{1}$$

where; x_i = the scaled variable; X_i = the original variable; μ_i = the mean of the original variable and σ_i = the standard deviation of the original variable.

For instance, the first principal component is:

$$\text{Factor1} = -0.16x_1 - 0.28x_2 + 0.24x_3 + 0.47x_4 + 0.27x_5 + 0.51x_6 - 0.42x_7 - 0.34x_8 \tag{2}$$

where the x_i in Eq. (2) represent the scaled form of the variables:

X_1: Stream length
X_2: Drainage density
X_3: Mean stream slope
X_4: Maximum elevation of the subcatchment
X_5: Minimum elevation of the subcatchment
X_6: Weighted average elevation of the area

HESSD
8, 3599–3629, 2011

X_7: Monthly NDVI

X_8: Monthly precipitation

Inspection of the eigenvectors for the first factor shows that avg_slope, max_elev, min_elev and avg_elev play contrasting roles to strm_len, drainage, monthly_prec and monthly_NDVI. This factor explains 44% of the variation in the data. The eigenvectors of the three factors are shown in Table 6.

The factor loadings are the correlations between the variables and the factors. Factor1 is most highly correlated to the maximum elevation and the average elevation; whereas factor 3 is most highly correlated to the average slope and the minimum elevation (Table 7).

Figure 4 provides an illustration of the projection of the variables on a factor plane using an alternative criteria for the PCA of variables. Each quadrant represents a similar group of variables.

Further use of the PCA is made in order to identify groupings of subcatchments by assessing the projection of cases onto the factor plane (Fig. 5).

The figure distinguishes four groups of subcatchments in term of variability of catchment descriptors. However, these groupings are reduced to two (2) main groups or categories (Table 8) so as to simplify the prediction equations of the runoff, which is the main goal of this paper. The physiographic attributes, namely mean stream slope, minimum elevation, maximum elevation, weighted average elevation that provide this major categorization from the PCA are located in the first quadrant of Fig. 4. The two main categories of subcatchments identified through the PCA are further illustrated in Table 9. All subcatchments in group I for instance are characterized by elevations whose minimum values do not exceed 703 m, while group II subcatchments on the other hand show weighted average of elevations that are above 1122 m. These similarities in physiographic attributes further support the grouping of subcatchments that is produced by the PCA and the tree clustering (Fig. 7).

4.2 Multi-linear and polynomial regression

Multi-linear and polynomial regressions were the modeling approaches used for the determination of the monthly local runoff equations. The multi-linear approach assumed linearity between catchments descriptors and the volume generated within the watershed. Therefore, several normality tests were performed and the results, reported in Table 10, show that the Anderson Darling test was the only one that rejected the null hypothesis at 20%.

The general models generated from the multi-linear and polynomial regressions are represented in Eqs. (3) and (4). The optimum monthly parameters for each category and the two models are reported in Tables 12 and 13. Figure 6a,b illustrates the multi-linear model performance for the two categories of watershed in February while Fig. 6c,d shows results from polynomial regression for the same month.

$$Y = a_o + A_i^T X_i \tag{3}$$

T denotes the transpose.

Where:

$$A_i = \begin{Bmatrix} \alpha_1 \\ \alpha_2 \\ . \\ . \\ . \\ \alpha_8 \end{Bmatrix} \quad \text{and} \quad X_i = \begin{Bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_8 \end{Bmatrix}$$

$$Y = a_o + A_{i\,j}^T Z_{i\,j} \tag{4}$$

HESSD
8, 3599–3629, 2011

$$A_{i\,j} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ . & . \\ . & . \\ . & . \\ \alpha_{71} & \alpha_{72} \\ \alpha_{81} & \alpha_{82} \end{bmatrix} \quad \text{and} \quad Z_{i\,j} = \begin{bmatrix} x_1 & x_1^2 \\ x_2 & x_2^2 \\ . & . \\ . & . \\ . & . \\ x_7 & x_7^2 \\ x_8 & x_8^2 \end{bmatrix}$$

With a_o: the intercept and other variables X_i are as defined earlier.

5 Conclusions

This study undertaken in the data-scarce Semliki watershed of the Equatorial Nile reported on the use of two modeling approaches for the prediction of flows. The principal component analysis performed identified variables that explained most of the variability in the dataset investigated. Furthermore, similar subcatchments in terms of physiographic attributes were identified giving rise to two categories of subcatchments that were used to generate runoff through multilinear and polynomial regressions. The dimensionless statistics (multiple R, multiple R^2, adjusted R^2) for the two modelling approaches indicated that the polynomial regression approach slightly outperformed the multi-linear regression. While few hydrological studies have documented the first approach as one of the germane ways in the determination of flow modelling, this study has illustrated the fact that the linearity assumption between catchment descriptors and the discharges is adequate for Semliki and hydrologically similar regions. This calls for further investigation to find out if the success of the linear model could add to the understanding of the hydrological processes in the basin. In addition, these two approaches provide useful models for long-term planning in the data-poor environment of the Semliki watershed.

HESSD
8, 3599–3629, 2011

References

Cheng, Q., Ko, C., Yuan, Y., Ge, Y., and Zhang, S.: GIS modeling for predicting river runoff volume in ungauged drainages in the greater Toronto area, Canada, Comput. Geosci., 32, 1108–1119, 2006.

Heuvelmans, G., Muys, B., and Feyen, J.: Regionalization of the parameters of a hydrological model: comparison of linear regression models with artificial nets, J. Hydrol., 319, 245–265, 2006.

Kileshye Onema, J.-M. and Taigbenu, A. E.: NDVI–rainfall relationship in the Semliki watershed of the Equatorial Nile, Phys. Chem. Earth, 34, 711–721, 2009.

Koutsoyiannis, D.: Uncertainty, entropy, scaling and hydrological stochastics, 1, Marginal distributional properties of hydrological processes and state scaling, Hydrolog. Sci. J., 50, 381–404, 2005a.

Koutsoyiannis, D.: Uncertainty, entropy, scaling and hydrological stochastics, 2, Time dependence of hydrological processes and time scaling, Hydrolog. Sci. J., 50, 405–426, 2005b.

Koutsoyiannis, D., Yao, H., and Georgakakos, A.: Medium-range flow prediction for the Nile: a comparison of stochastic and deterministic methods, Hydrolog. Sci. J., 53, 405–426, 2008.

Kundzewicz, Z. W.: Prediction in Ungauged Basins - a Systemic Perspective, IAHS publication no. 309, in: Predictions in Ungauged Basins: PUB Kick-off, edited by: Schertzer, D., ISBN-1-901502-83-1, IAHS Press, Institute of Hydrology, Wallingford, Oxfordshire, UK, 2007.

Kwon, H.-H., Brown, C., Xu, K., and Lall, U.: Seasonal and annual maximum streamflow forecasting using climate information: application to the Three Gorges Dam in the Yangtze River Basin, China, Hydrolog. Sci. J., 54, 606–622, 2009.

McIntyre, N., Lee, H., Wheater, H., Young, A., and Wagener, T.: Ensemble predictions of runoff in ungauged catchments, Water Resour. Res., 41, W12434, doi:10.1029/2005WR004289, 2005.

Merz, R. and Blöschl, G.: Regionalization of catchment model parameters, J. Hydrol., 287, 95–123, 2004.

Mwakalila, S.: Estimation of stream flows of ungauged catchments for river basin management, Phys. Chem. Earth, 28, 935–942, 2003.

Sanborn, S. C., and Bledsoe, B. P.: Predicting streamflow regime metrics for ungauged streams in Colorado, Washington and Oregon, J. Hydrol., 325, 241–261, 2006.

Schröder, B.: Pattern, process, and function in landscape ecology and catchment hydrology – how can quantitative landscape ecology support predictions in ungauged basins?, Hydrol. Earth Syst. Sci., 10, 967–979, doi:10.5194/hess-10-967-2006, 2006.

Sefton, C. E. M. and Howard, S. M.: Relationship between dynamic response characteristics and physical descriptors of catchments in England and Wales, J. Hydrol., 211, 1–16, 1998.

Shao, Q., Zhang, L., Chen, Y. D., and Singh, V. P.: A new method for modeling flow duration curves and predicting streamflow regimes under altered land-use conditions, Hydrolog. Sci. J., 54, 582–595, 2009.

Sharda, V. N., Prasher, S. O., Patel, R. M., Ojasvi, P. R., and Prakash, C.: Performance of Multivariate Adaptive Regression Splines (MARS) in predicting runoff in Mid-Himalayan micro-watersheds with limited data, Hydrolog. Sci. J., 53(6), 1165–1175, 2008.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., Mcdonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S. and Zehe, E.: IAHS decade on Predictions in Ungauged Basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences, Hydrolog. Sci. J., 48(6), 857–880, 2003.

Sutcliffe, J. V. and Parks, Y. P.: The Hydrology of the Nile, IAHS special publication no. 5, ISBN 1-901502-75-9; IAHS Press, Institute of Hydrology, Wallingford, Oxfordshire, UK, 1999.

Uhlenbrook, S. and Sieber, A.: On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model, Environ. Modell. Softw., 20, 29–32, 2005.

Xu, C.: Testing the transferability of regression equations derived from small sub-catchments to a large area in central Sweden, Hydrol. Earth Syst. Sci., 7, 317–324, doi:10.5194/hess-7-317-2003, 2003.

Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, Adv. Water Resour., 30, 1756–1774, 2007.

Zhang, Z., Wagener, T., Reed, P., and Bhushan, R.: Reducing uncertainty in prediction in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization, Water Resour. Res., 44, W00B04, doi:10.1029/2008WR006833, 2008.

Table 1. Physiographic attributes generated for subcatchments (S3–S23).

Physiographic attribute	Unit	Abbreviation
Stream length	M	Strm_len
Drainage density	km km^{-2}	Drainage
Mean stream slope	%	avg_slope
Max elevation of the subcatchment	M	Max_elev
Min elevation of the subcatchment	M	Min_elev
Weighted average elevation of the area	M	avg_elev
Monthly precipitation	mm month^{-1}	monthly_prec
Monthly NDVI	–	monthly_NDVI

HESSD
8, 3599–3629, 2011

Data-scarce watershed of the Equatorial Nile region

J.-M. Kileshye Onema et al.

Table 2. Descriptive statistics of variables.

Variables	Count	Mean	Standard deviation
Stream length	21	29.46	20.66
Drainage	21	7.85×10^{-2}	3.52×10^{-2}
Average slope	21	0.2	0.23
Maximum elevation	21	2577.9	1273.84
Minimum elevation	21	733.18	99.78
Average elevation	21	1164.82	392.54

HESSD
8, 3599–3629, 2011

Table 3. Coefficients of correlations between variables.

	Stream length	Drainage	Average slope	Maximum elevation	Minimum elevation	Average elevation	Monthly precipitation	Monthly NDVI
Stream length	1	0.344	−0.2	−0.05	−0.29	−0.29	−0.08	0.07
Drainage		1	−0.19	−0.46	−0.24	−0.45	0.07	0.16
Average slope			1	0.37	0.13	0.54	0.08	−0.15
Maximum elevation				1	0.23	0.87	−0.5	−0.63
Minimum elevation					1	0.46	−0.25	−0.23
Average elevation						1	−0.45	−0.64
Monthly precipitation							1	0.78
Monthly NDVI								1

HESSD
8, 3599–3629, 2011

HESSD

8, 3599–3629, 2011

Data-scarce watershed of the Equatorial Nile region

J.-M. Kileshye Onema et al.

Table 4. Bartlett test and Glaeson–Staelin (Phi).

Bartlett test	DF	p-value	Glaeson–Staelin (Phi)
81.98	28	0.00000	0.395415

Table 5. Eigenvalues of components.

No.	Eigenvalue	Individual percent	Cumulative percent
1	3.49	43.59	43.59
2	1.56	19.46	63.04
3	1.02	12.71	75.75
4	0.78	9.80	85.55
5	0.64	7.97	93.51
6	0.30	3.74	97.25
7	0.17	2.13	99.38
8	0.05	0.62	100.00

HESSD
8, 3599–3629, 2011

J.-M. Kileshye Onema et al.

Table 6. Eigenvectors of principal components.

Variables	Factor1	Factor2	Factor3
Strm_len	−0.16	−0.52	−0.42
Drainage	−0.28	−0.36	−0.10
avg_slope	0.24	0.37	−0.60
Max_elev	0.47	−0.08	−0.27
Min_elev	0.27	0.17	0.56
avg_elev	0.51	0.10	−0.14
monthly__prec	−0.34	0.52	−0.22
monthly_NDVI	−0.42	0.37	−0.05

HESSD
8, 3599–3629, 2011

Table 7. Factor loadings of principal components.

Variables	Factor1	Factor2	Factor3
Strm_len	−0.29	−0.65	−0.42
Drainage	−0.53	−0.45	−0.10
avg_slope	0.45	0.46	−0.60
Max_elev	0.87	−0.10	−0.27
Min_elev	0.50	0.21	0.56
avg_elev	0.94	0.12	−0.14
monthly_prec	−0.63	0.65	−0.22
monthly_NDVI	−0.78	0.46	−0.05

HESSD
8, 3599–3629, 2011

Table 8. Grouping of Semliki subcatchments from the PCA.

Category I	Category II
S3	S14
S4	S15
S5	S16
S6	S17
S7	S18
S8	S20
S9	S21
S10	S22
S11	S23
S12	
S13	
S19	

HESSD

8, 3599–3629, 2011

Table 9. Physiographic attributes of the two groups of Semliki subcatchments as identified by the PCA.

	Str. leng(Km)	Drain. Dsity	Strm slope (%)	Max elev. (m)	Min elev. (m)	W avge elev. (m)	Monthly prec.(mm)	Monthly NDVI
S3	24.9	0.2284	**0.0121**	640	**617**	625	87.1	**0.48**
S4	55.0	0.0886	**0.0909**	1774	**624**	1040	99.1	**0.59**
S5	7.3	0.0482	**0.0953**	836	**615**	675	91.7	**0.54**
S6	29.4	0.1118	1.3927	1476	692	1111	115.3	0.68
S7	10.9	0.1007	0.0365	1307	690	888	112.5	0.69
S8	74.4	0.1148	**0.0632**	2640	**615**	877	96.9	**0.58**
S9	10.8	0.1328	0.0739	861	695	767	106.0	0.67
S10	82.4	0.1039	**0.0594**	2983	**624**	812	102.5	**0.61**
S11	39.0	0.0671	0.0564	2069	667	879	106.6	0.67
S12	15.7	0.0351	0.4454	1186	703	891	114.0	0.70
S13	16.7	0.0577	0.0240	2221	701	858	107.0	0.67
S14	25.7	0.0543	0.4667	3920	669	1600	98.9	0.57
S15	14.0	0.0457	0.2705	4727	710	1915	98.4	0.53
S16	10.1	0.0541	0.2470	4329	719	1502	98.1	0.41
S17	17.7	0.0772	0.5196	4862	717	1566	99.4	0.63
S18	12.9	0.0441	0.4280	4792	806	1913	94.6	0.43
S19	47.6	0.1143	0.1953	1887	804	1036	107.6	0.63
S20	22.0	0.0630	0.3048	4197	899	1718	92.0	0.57
S21	17.0	0.0567	0.0118	2264	894	1122	99.8	0.58
S22	29.3	0.0965	0.4238	2382	901	1499	102.1	0.59
S23	36.5	0.0508	0.0329	2542	895	1201	89.6	0.57

HESSD
8, 3599–3629, 2011

Data-scarce watershed of the Equatorial Nile region

J.-M. Kileshye Onema et al.

Table 10. Normality test.

Test name	Test value	Prob level	Reject H_0 At alpha = 20%
Shapiro Wilk	0.96	0.44	No
Anderson Darling	0.53	0.18	Yes
D'Agostino Skewness	0.64	0.52	No
D'Agostino Kurtosis	0.35	0.73	No
D'Agostino Omnibus	0.54	0.76	No

HESSD
8, 3599–3629, 2011

Table 11. Performance statistics (minimum and maximum).

Method	Multiple R	Multiple R^2	Adjusted R^2
Polynomial regression	0.99	0.98–0.99	0.98–0.99
Multi-regression	0.96–0.99	0.93–0.99	0.92–0.99

HESSD
8, 3599–3629, 2011

Table 12. Monthly equations parameters from multi-linear regression optimization for Semliki subcatchments.

		a (10^5)	$\alpha1(10^4)$	$\alpha2(10^6)$	$\alpha3(10^4)$	$\alpha4$	$\alpha5(10^2)$	$\alpha6$	$\alpha7(10^4)$	$\alpha8$
Group 1	Jan	56.01	16.48	−35.53	−38.21	−719.88	−14.00	−201.95	150.01	1222.84
	Feb	47.30	14.83	−31.33	−36.16	−642.21	−9.88	−112.01	127.40	520.72
	Mar	53.90	16.08	−32.70	−42.33	−892.36	−22.72	−607.55	319.28	957.68
	Apr	48.70	17.08	−36.28	−39.13	−753.17	−10.80	−334.74	196.1	1535.05
	May	64.00	18.82	−41.39	−40.67	−708.66	−8.29	−10.07	6.02	1800.68
	Jun	37.02	16.78	−35.29	−34.38	−725.02	−8.79	−508.86	374.50	2821.7
	Jul	46.78	17.70	−37.45	−35.58	−791.07	−8.06	−553.03	259.50	5097.92
	Aug	52.58	18.22	−38.60	−29.86	−757.17	−15.04	−317.23	284.80	1160.25
	Sep	47.31	19.19	−38.54	−46.60	−1098.71	−28.99	−1020.17	599.10	1544.80
	Oct	52.66	19.60	−40.86	−45.63	−893.83	−22.71	−771.69	427.00	2983.79
	Nov	31.05	20.99	−41.82	−74.25	−1097.14	−25.40	−796.84	787.60	3515.37
	Dec	49.29	19.71	−40.81	−59.94	−927.30	−20.72	−252.88	436.20	1411.49
Group 2	Jan	51.14	33.54	−90.44	−40.88	322.03	−2.79	−872.59	−13.15	10.57
	Feb	40.09	27.02	−73.36	−30.08	258.47	−1.53	−700.24	2.10	250.21
	Mar	44.52	29.951	−81.58	−31.94	283.73	−1.66	−773.52	6.96	30.78
	Apr	49.45	33.26	−90.33	−37.20	317.06	−2.03	−855.94	4.15	48.04
	May	55.15	37.47	−102.95	−36.29	350.33	−2.00	−961.26	32.87	224.64
	Jun	48.73	32.69	−88.79	−36.16	311.26	−1.97	−841.62	2.76	33.69
	Jul	51.09	34.34	−93.25	−38.65	326.76	−2.06	−881.12	2.89	131.12
	Aug	53.36	35.70	−97.01	−41.57	336.48	−2.57	−902.37	8.37	13.27
	Sep	52.06	35.21	−95.96	−39.75	332.06	−2.22	−888.41	13.12	83.98
	Oct	55.17	37.63	−103.09	−38.74	358.54	−1.99	−980.17	26.12	175.35
	Nov	57.42	39.37	−107.70	−38.50	373.67	−1.76	−1015.20	19.92	220.07
	Dec	56.28	38.10	−103.68	−41.36	365.54	−2.11	−988.24	7.84	349.41

HESSD
8, 3599–3629, 2011

Data-scarce watershed of the Equatorial Nile region

J.-M. Kileshye Onema et al.

Table 13. Monthly equations parameters from polynomial regression optimization for Semliki subcatchments.

		a(10^5)	α11 (10^4)	α12 (10^2)	α21 (10^6)	α22 (10^6)	α31 (10^5)	α32 (10^5)	α61 (10^3)	α62	α71 (10^5)	α72 (10^5)	α81 (10^2)	α82
	Jan	−179.33	27.42	−13.19	−79.98	203.10	25.40	−11.35	54.10	−32.40	9.09	− 3.73	19.61	14.40
	Feb	−118.54	25.06	−12.12	−68.99	168.03	24.14	−12.38	42.10	−25.58	−40.51	50.03	−241.12	185.93
	Mar	−150.17	25.52	−12.17	−74.31	188.10	22.86	−10.63	45.30	−27.38	7.84	1.42	48.26	−18.41
	Apr	−143.93	28.52	−13.61	−81.86	205.83	26.53	−13.22	47.70	−28.95	−50.78	53.93	52.32	−10.81
	May	−169.92	32.11	−15.57	−90.91	227.42	30.58	−14.14	62.60	−37.59	−111.9	87.66	−7.12	4.14
Group 1	Jun	−163.19	27.80	−13.29	−80.49	203.69	24.94	−11.58	49.40	−29.83	−4.82	21.71	30.39	−5.31
	Jul	−163.99	29.33	−14.19	−81.02	198.27	29.47	−14.09	54.90	−33.12	−88.71	78.21	216.21	−95.50
	Aug	−204.07	30.54	−14.79	−86.94	218.71	28.42	−12.83	57.80	−34.80	−12.30	17.66	299.44	−116.12
	Sep	−117.78	29.73	−13.55	−89.93	228.62	24.60	−12.54	42.60	−26.01	−68.60	82.39	118.95	−46.91
	Oct	−209.42	31.78	−14.87	−96.46	250.31	27.53	−12.79	55.80	−33.69	93.60	−52.37	−10.49	−1.30
	Nov	−179.07	33.17	−15.59	−95.83	240.55	29.50	−14.41	58.50	−35.32	−48.37	55.26	34.17	−8.54
	Dec	−161.32	31.65	−14.66	−94.94	241.92	28.57	−13.92	57.60	−34.48	−104.8	96.09	205.63	−145.43
	Jan	56.68	18.73	30.28	20.14	−750.34	–	–	−4.68	1.65	−11.55	15.65	20.96	−14.76
	Feb	46.88	14.52	25.53	16.31	−607.02	–	–	−3.73	1.31	8.53	−5.78	−178.52	167.26
	Mar	45.09	17.08	26.54	9.61	−616.86	–	–	−3.92	1.37	−11.43	16.03	109.56	−41.38
	Apr	47.44	17.20	32.51	8.26	−673.04	–	–	−3.49	1.26	16.45	−8.55	−32.64	13.57
	May	62.60	21.45	32.78	9.48	−764.34	–	–	−4.23	1.52	−17.01	21.49	−14.79	5.19
Group 1	Jun	56.65	18.60	28.81	8.66	−666.08	–	–	−4.12	1.46	−14.27	18.72	−17.87	20.93
	Jul	60.49	19.63	30.52	9.05	−697.21	–	–	−4.83	1.70	−11.85	18.07	26.76	−2.25
	Aug	49.50	19.76	32.64	15.67	−766.07	–	–	−4.46	1.58	12.78	−10.72	87.65	−37.73
	Sep	53.42	19.54	31.86	9.54	−718.87	–	–	−4.09	1.47	4.43	1.31	10.27	−4.31
	Oct	52.32	20.60	34.65	15.74	−803.16	–	–	−4.44	1.58	19.72	−13.31	18.25	−8.99
	Nov	60.45	22.26	35.03	21.00	−870.98	–	–	−5.18	1.85	15.39	−8.48	−47.82	18.46
	Dec	60.70	20.72	35.53	21.76	−850.59	–	–	−4.99	1.78	−21.07	25.22	125.89	−82.99

HESSD

8, 3599–3629, 2011

Data-scarce watershed of the Equatorial Nile region

J.-M. Kileshye Onema et al.

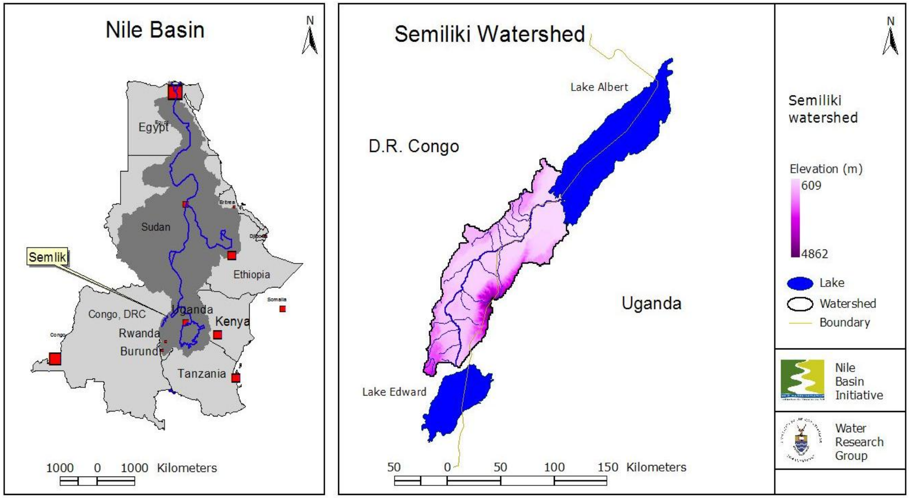

Fig. 1. Semliki watershed.

HESSD
8, 3599–3629, 2011

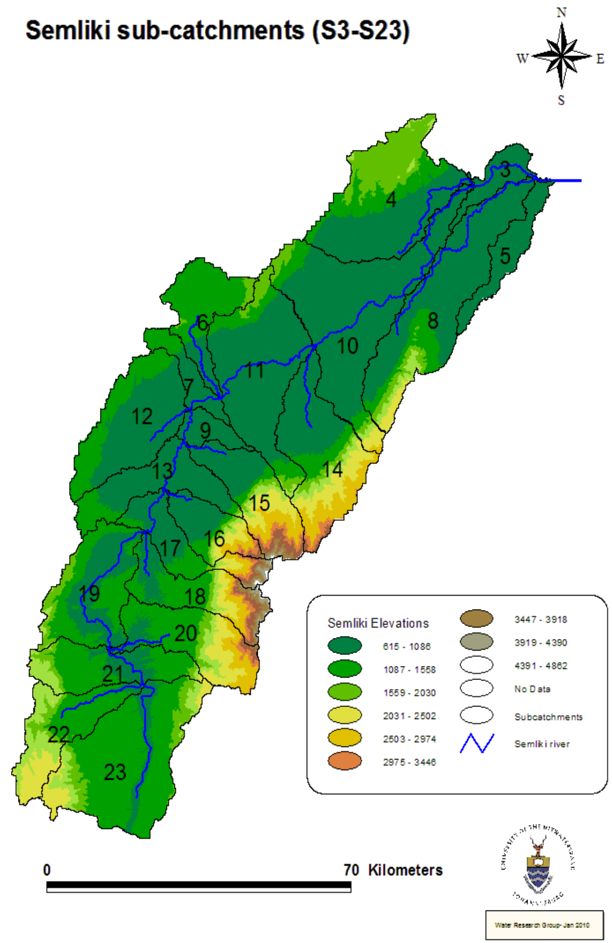

Fig. 2. Semliki subcatchments (S3–S23).

HESSD
8, 3599–3629, 2011

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

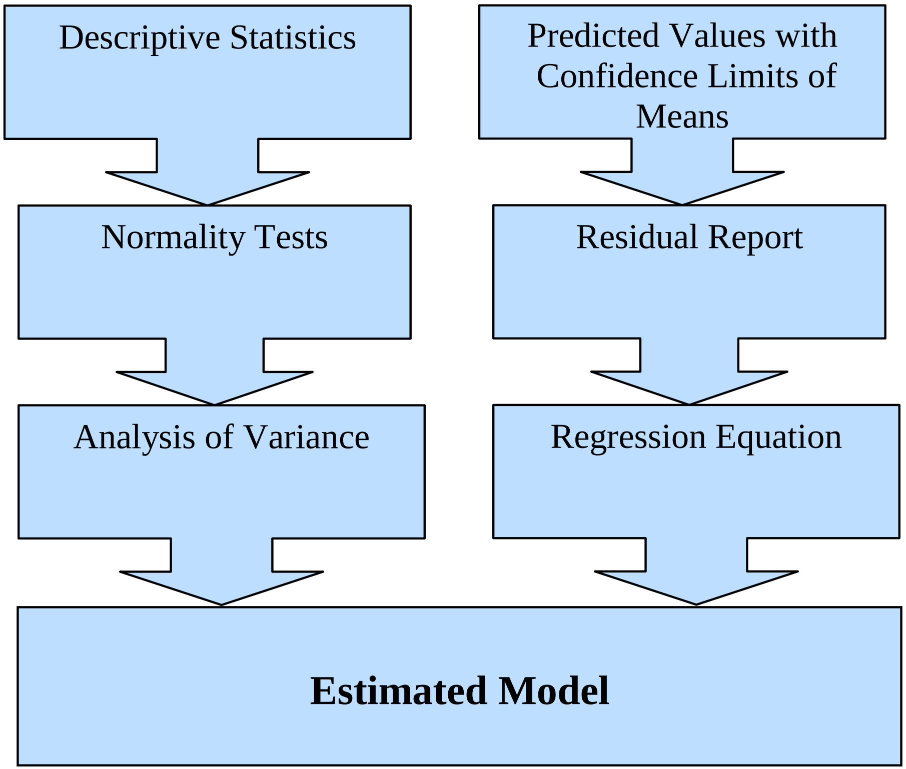

Fig. 3. Multiple regression analysis.

HESSD
8, 3599–3629, 2011

J.-M. Kileshye Onema et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

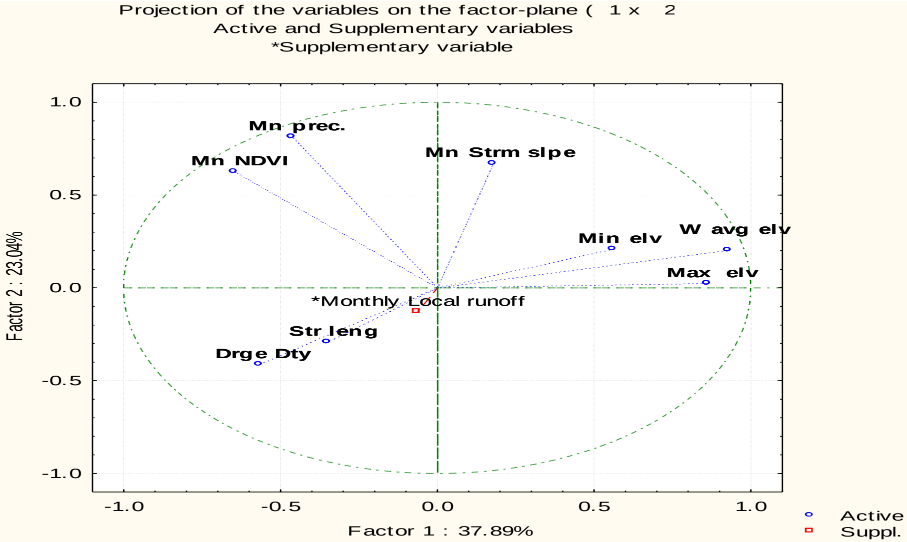

Fig. 4. PCA of variables.

HESSD
8, 3599–3629, 2011

Data-scarce watershed of the Equatorial Nile region

J.-M. Kileshye Onema et al.

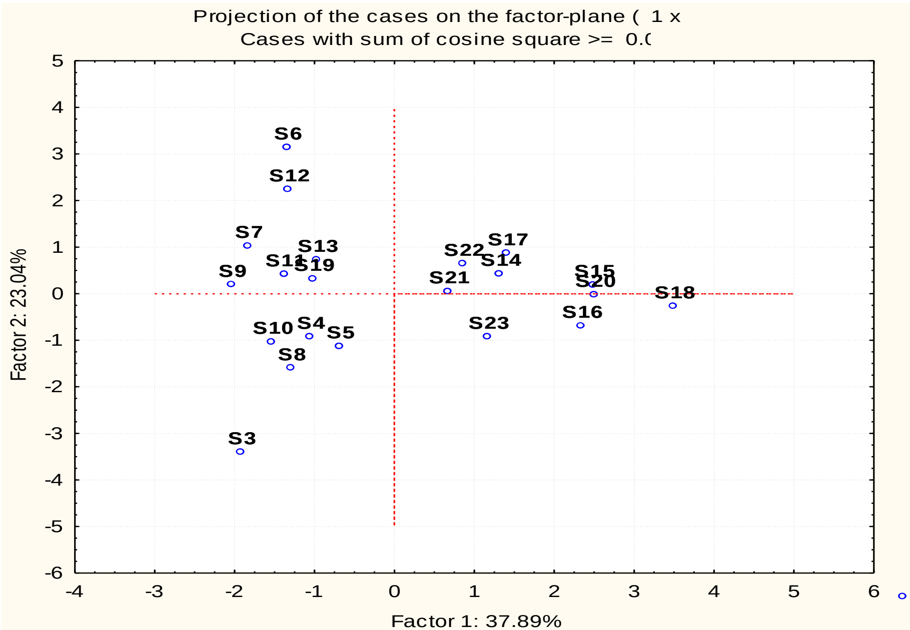

Fig. 5. Projection of cases from the PCA of variables.

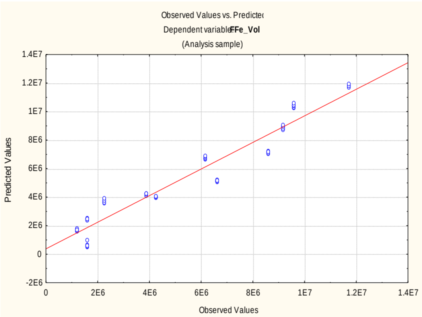

Figure 6 (a)

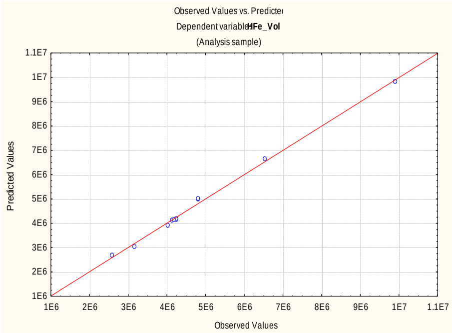

Figure 6 (b)

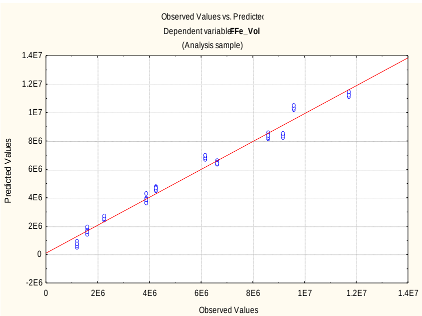

Figure 6 (c)

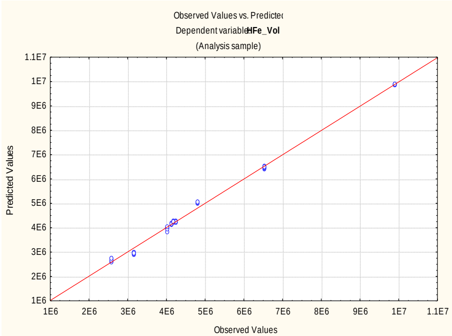

Figure 6 (d)

Fig. 6. Model predictions in February.

HESSD
8, 3599–3629, 2011

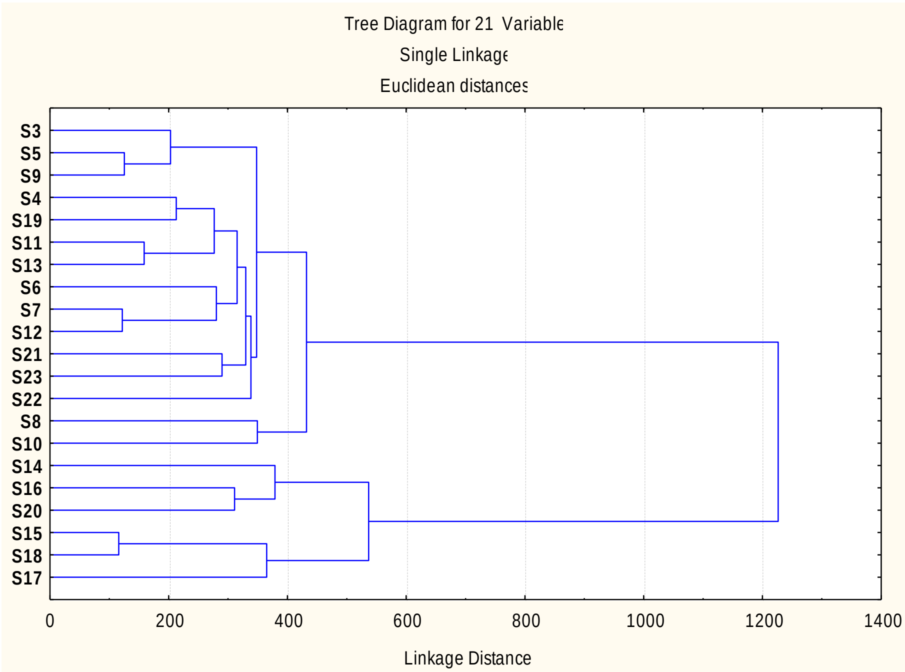

Fig. 7. Tree diagram clustering of Semliki subcatchments.