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

J.-M. Kileshye Onema1,2, A. E. Taigbenu1, and J. Ndiritu1
1Water Research Group, School of Civil and Environmental Engineering, University of Witwatersrand, Private Bag 3, WITS 2050, South Africa
2WaterNet, Box MP 600 Mount Pleasant, Harare, Zimbabwe

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

Received: 31 March 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 14 April 2011
Revised: 14 April 2012 – Accepted: 21 April 2012 – Published: 15 May 2012

Abstract. Continuous developments and investigations in flow predictions 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 km3/annum, the Semliki watershed contributes up to 20 % of the flows of the White Nile. The watershed was sub-divided in 21 sub-catchments (S3 to S23). Using eight physiographic and meteorological variables, generated from remotely sensed acquired datasets and limited catchment data, monthly runoffs were estimated. One ordination technique, the Principal Component Analysis (PCA), and the tree cluster analysis of the landform attributes were performed to study the data structure and spot physiographic similarities between sub-catchments. The PCA revealed the existence of two major groups of sub-catchments – flat (Group I) and hilly (Group II). Linear and nonlinear regression models were used to predict the long-term monthly mean discharges for the two groups of sub-catchments, and their performance evaluated by the Nash-Sutcliffe Efficiency (NSE), Percent bias (PBIAS) and root mean square error to the standard deviation ratio (RSR). The dimensionless indices used for model evaluation indicate that the non-linear model provides better prediction of the flows than the linear one.

1 Introduction

Numerous approaches exist for streamflow prediction in natural river reaches. Streamflow forecasting has significant interest both from a research and an operational point of view. The choice of methods depends on data availability and the type of application. While continuous research efforts strive at enhancing our predictive capability for streamflow, we are often faced with the challenge of making such predictions in basins that are poorly gauged or not gauged at all (Sivapalan et al., 2003). Reliable and accurate estimates of hydrologic components 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 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. Many authors have also identified the reduction of uncertainty associated with predictions in ungauged basins as being very important (Uhlenbrook and Siebert, 2005; Koutsoyiannis, 2005a, b; Zhang et al., 2008).
Table 1. Statistics of physiographic variables.

<table>
<thead>
<tr>
<th>Variables</th>
<th>No.</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stream Length ($x_1$)</td>
<td>21</td>
<td>29.46</td>
<td>20.66</td>
</tr>
<tr>
<td>Drainage density ($x_2$)</td>
<td>21</td>
<td>$7.85 \times 10^{-2}$</td>
<td>$3.52 \times 10^{-2}$</td>
</tr>
<tr>
<td>Average stream Slope ($x_3$)</td>
<td>21</td>
<td>0.2</td>
<td>0.23</td>
</tr>
<tr>
<td>Maximum elevation ($x_4$)</td>
<td>21</td>
<td>2577.9</td>
<td>1273.84</td>
</tr>
<tr>
<td>Minimum elevation ($x_5$)</td>
<td>21</td>
<td>733.18</td>
<td>99.78</td>
</tr>
<tr>
<td>Weighted average elevation ($x_6$)</td>
<td>21</td>
<td>1164.82</td>
<td>392.54</td>
</tr>
</tbody>
</table>

Lately, Koutsoyiannis et al. (2008) indicated the use of analogue modeling techniques for flow prediction which give impressive performance due to advances in nonlinear dynamical systems (chaotic systems). The major drawbacks of these approaches are that they are data intensive and work as black boxes, thus provide no insight into the hydrological processes.

This paper reports on linear and nonlinear regression modeling approaches for flow prediction in a medium size watershed of the equatorial Nile region (Semliki catchment), where very little hydro-meteorological data are available. These approaches attempt to provide monthly flow estimates that relate similar catchment physiographic attributes to generated flows.

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/month) and October (205 mm/month). 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-capped 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 are reported in Kileshye Onema and Taigbenu (2009).

3 Data and methods

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. The Semliki catchment was delineated into 21 sub-catchments (S3–S23) that are presented in Fig. 2. This delineation was achieved using the 90m Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM) along with traditional maps. From six physiographic attributes, the 21 sub-catchments were grouped...
into two categories using the principal component analysis (PCA) and cluster analysis. Making use of the assumption that the physiographic variables are stationary, a quasitemporal prediction of the flow was achieved with two additional variables (NDVI and rainfall) that have both spatial and temporal variation. These two latter variables were generated for the catchment from remote-sensed data. The physiographical characteristics of these sub-catchments were identified by six variables, namely the stream length \((x_1)\), the drainage density \((x_2)\), the mean stream slope \((x_3)\), the maximum elevation \((x_4)\), the minimum elevation \((x_5)\), and the weighted average elevation \((x_6)\). The statistical properties of these physiographic variables are presented in Table 1 and used subsequently in the grouping of these sub-catchments. In predicting monthly flows in these sub-catchments, it is assumed that these variables are stationary, constant over time but not over space. Two additional variables, generated for these sub-catchments, are the monthly rainfall \((x_7)\) and NDVI \((x_8)\) which have both spatial and temporal variation. The remotely-sensed NDVI data covered the period from 1982 to 2008 and were extracted from decadal maximum composite imagery provided by the National Oceanic and Atmospheric Administration-Advance Very High Resolution Radiometer (NOAA-AVHRR) and processed with the image display and analysis software WinDisp 5.1. The satellite derived rainfall data were provided by the National Oceanic and Atmospheric Administration (NOAA) through the Famine Early Warning System Network (FEWS-Net) for the period 2001–2007, and were refined by the only point daily rainfall data in the catchment obtained from a station located at Beni which covered the period from 1973 to 2008. Additional details on the processing of remotely sensed acquired datasets...
Table 2. Generated landscape attributes for subcatchments (S3–S23).

<table>
<thead>
<tr>
<th>Description</th>
<th>Source</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landform</td>
<td>SOTERCAF</td>
<td>GIS processing (WINDISP, ARC View3.3, Excel 07-10, Statistica 8.0)</td>
</tr>
<tr>
<td>Lithology</td>
<td>SOTERCAF</td>
<td>GIS processing (WINDISP, ARC View3.3, Excel 07-10, Statistica 8.0)</td>
</tr>
<tr>
<td>Soils</td>
<td>SOTERCAF</td>
<td>GIS processing (WINDISP, ARC View3.3, Excel 07-10, Statistica 8.0)</td>
</tr>
<tr>
<td>Drainage Density (Dd)</td>
<td>SRTM 90m-DEM, SWAT pre-processor</td>
<td>Subcatchments areas generated from SWAT preprocessor (WINDISP, ARC View3.3, Excel 07-10, Statistica 8.0, SWAT)</td>
</tr>
<tr>
<td>Stream Length</td>
<td>SRTM 90m-DEM, topographical map (1/50,000), SWAT pre-processor</td>
<td>Generated from the SWAT preprocessor and cross validation with traditional map (WINDISP, ARC View3.3, Excel 07-10, Statistica 8.0)</td>
</tr>
<tr>
<td>Stream slope</td>
<td>SRTM 90m DEM</td>
<td>Remotely sensed acquired (ARC View3.3, Excel 07-10, Statistica 8.0)</td>
</tr>
<tr>
<td>Elevations: Minimum</td>
<td>SRTM 90m-DEM, topographical map (1/50,000)</td>
<td>Remotely sensed acquired and cross validation with traditional map. (Arc View 3.3, Excel 07-10, Statistica 8.0)</td>
</tr>
<tr>
<td>Maximum</td>
<td>Area-weighted average</td>
<td></td>
</tr>
<tr>
<td>NDVI</td>
<td>NOAA-AVHRR (1982–2008)</td>
<td>Remotely sensed acquired and correlated with rainfall. (WINDISP, ARC View3.3, Excel 07-10, Statistica 8.0)</td>
</tr>
</tbody>
</table>

Fig. 6. Tree cluster of the physiographic variables.

can be found in Kileshye Onema and Taigbenu (2009). At the outlet of the Semliki watershed is a gauging station located at Bweramule (Fig. 2) which provided historical monthly flows from 1950 to 1978 that were used for the calibration of both the linear and the non-linear flow prediction models that are subsequently described in this paper. The periods of available data on climate-related variables, namely rainfall and NDVI, that can be considered causal to the observed stream flow are presented in Fig. 3. The figure reveals the challenge posed in the flow modelling of this catchment which, apart from the lack of data for flows in the tributaries of the Semliki river and climatic variables for the 21 sub-catchments, the only available point rainfall data from Beni, which is not representative of the entire catchment, overlaps the record of monthly flow measurements at Bweramule for only six years from 1973 to 1978. Our problem statement, therefore, precludes the use of regionalization techniques that are based on flow duration curves (FDCs) that can be constructed by statistical or parametric or graphical approaches (Castellarin et al., 2004; Quipo et al., 1983). For each subcatchment, a historical 28-yr monthly mean volume was computed proportionally to the subcathment area and was labeled as “control”. This historical approach is similar to the one undertaken by Asadullah et al. (2008) in data-scarce regions of central and Eastern Africa. The runoff generation approach is supported by the fact that in humid basins like the Semliki as opposed
Table 3. Coefficients of correlations between variables.

<table>
<thead>
<tr>
<th></th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
<th>x5</th>
<th>x6</th>
<th>x7</th>
<th>x8</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>1</td>
<td>0.344</td>
<td>-0.2</td>
<td>-0.05</td>
<td>-0.29</td>
<td>-0.29</td>
<td>-0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>x2</td>
<td>1</td>
<td>0.19</td>
<td>0.46</td>
<td>-0.24</td>
<td>-0.45</td>
<td>0.07</td>
<td>-0.16</td>
<td></td>
</tr>
<tr>
<td>x3</td>
<td>1</td>
<td>0.37</td>
<td>0.13</td>
<td>0.54</td>
<td>0.08</td>
<td>-0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>x4</td>
<td>1</td>
<td>0.23</td>
<td>0.87</td>
<td>-0.5</td>
<td>-0.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x5</td>
<td>1</td>
<td>0.46</td>
<td>-0.25</td>
<td>-0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x6</td>
<td>1</td>
<td>-0.45</td>
<td>-0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x7</td>
<td>1</td>
<td>0.78</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x8</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Eigenvalues of components.

<table>
<thead>
<tr>
<th>Eigenvalue</th>
<th>Cumulative Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.49</td>
<td>43.59</td>
</tr>
<tr>
<td>2</td>
<td>1.56</td>
<td>63.04</td>
</tr>
<tr>
<td>3</td>
<td>1.02</td>
<td>75.75</td>
</tr>
<tr>
<td>4</td>
<td>0.78</td>
<td>85.55</td>
</tr>
<tr>
<td>5</td>
<td>0.64</td>
<td>93.51</td>
</tr>
<tr>
<td>6</td>
<td>0.30</td>
<td>97.25</td>
</tr>
<tr>
<td>7</td>
<td>0.17</td>
<td>99.38</td>
</tr>
<tr>
<td>8</td>
<td>0.05</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Table 5. Eigenvectors of principal components.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Factor1</th>
<th>Factor2</th>
<th>Factor3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>-0.16</td>
<td>-0.52</td>
<td>-0.42</td>
</tr>
<tr>
<td>x2</td>
<td>-0.28</td>
<td>-0.36</td>
<td>-0.10</td>
</tr>
<tr>
<td>x3</td>
<td>0.24</td>
<td>0.37</td>
<td>-0.60</td>
</tr>
<tr>
<td>x4</td>
<td>0.47</td>
<td>-0.08</td>
<td>-0.27</td>
</tr>
<tr>
<td>x5</td>
<td>0.27</td>
<td>0.17</td>
<td>0.56</td>
</tr>
<tr>
<td>x6</td>
<td>0.51</td>
<td>0.10</td>
<td>-0.14</td>
</tr>
<tr>
<td>x7</td>
<td>-0.34</td>
<td>0.52</td>
<td>-0.22</td>
</tr>
<tr>
<td>x8</td>
<td>-0.42</td>
<td>0.37</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

to arid and semi-arid regions, “stream flows increase in the downstream direction, and the spatial distribution of average monthly or seasonal rainfall is more or less the same from one part of the river basin to another, hence the runoff per unit land area is assumed constant over space. In these situations, estimated flows are usually based on the watershed areas, as contributing flow to those sites, and the corresponding streamflows and watershed areas above the nearest or most representative gauge sites” (Loucks et al., 1981; Loucks and Van Beek, 2005). Furthermore, rainfall for the equatorial region, one of the main driving force behind runoff generation has not changed in a statistical significant way since 1950 despite reported seasonal and interannual variations (Nicholson and Entekhabi, 1987; Bigot et al., 1998; Lienou et al., 2008). Table 2 provides landscape attributes, their sources and software used and guides any further work that intends to follow the approach undertaken in this paper.

The data generated from the physiographic variables of the 21 sub-catchments were examined to explore their groupings and similarities in their data structure. To achieve these objectives, the principal component analysis (PCA) and cluster analysis were used as exploratory techniques to study the structure and similarities of the data. The PCA analysis performed in this study used the scale invariant correlation matrix of the variables as opposed to the covariance matrix. The scaled variables are obtained by the expression:

\[ X_i = \frac{x_i - \mu_i}{\sigma_i} \]  

where \( X_i \) is the scaled variable of the original variable \( x_i \), and \( \mu_i \) and \( \sigma_i \) are the mean and standard deviation of the original variable.

The Glaeson-Staelin redundancy statistic and the Bartlett’s sphericity test were performed as pre-requisite to the PCA. The Glaeson-Staelin redundancy statistic (\( \phi \)) measures the level of interrelation between a group of variables (Magingaxa et al., 2009) with a zero value indicating uncorrelated variables and unity indicating perfect correlation between the variables (Jolliffe, 2002). The Bartlett’s sphericity test is used to test the null hypothesis that the variables of the correlation matrix are not correlated (Sousa et al., 2007). When p-values of the Bartlett’s test are greater than 0.05 it is not prudent to proceed with the PCA (Peres-Neto et al., 2005). The Kaiser criterion was used to identify the number of principal components for consideration as providing significant variance in the data (Jolliffe, 2002). The general suggested rule associated with this criterion is to retain principal components with eigenvalue greater than 1 (Chen and Chen, 2003), and to consider those less than 1 as trivial.

The monthly local runoff of the catchment was predicted by linear and nonlinear regression models that make use of the physiographic and meteorological variables. They are described by the relationships below. For the linear model of the Bartlett’s test, the Nash-Sutcliffe efficiency (NSE), the percent bias (PBIAS) and the

\[ Q = a_0 + A_T^T x_i \]  

\[ Q = b_0 + B_{i}^T z_{ij} \]  

where \( Q \) represents the monthly flows, the superscript \( T \) denotes the transpose, \( a_0 \) and \( A_T = \{ a_1, a_2, ..., a_8 \} \) are the regression parameters, and \( x_i^T = \{ x_1, x_2, ..., x_8 \} \). The expression for the flow by the nonlinear model is given by

\[ B_{ij} = \{ \beta_{11}, \beta_{21}, ..., \beta_{81}, \beta_{12}, \beta_{22}, ..., \beta_{82} \} \]  

\[ z_{ij} = \{ x_1, x_2, ..., x_8 \} \]  

To evaluate the performance of these two models, the Nash-Sutcliffe efficiency (NSE), the percent bias (PBIAS) and the
Table 6. Grouping of Semliki sub-catchments from the PCA.

<table>
<thead>
<tr>
<th>Group I</th>
<th>Group II</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3</td>
<td>S14</td>
</tr>
<tr>
<td>S4</td>
<td>S15</td>
</tr>
<tr>
<td>S5</td>
<td>S16</td>
</tr>
<tr>
<td>S6</td>
<td>S17</td>
</tr>
<tr>
<td>S7</td>
<td>S18</td>
</tr>
<tr>
<td>S8</td>
<td>S20</td>
</tr>
<tr>
<td>S9</td>
<td>S21</td>
</tr>
<tr>
<td>S10</td>
<td>S22</td>
</tr>
<tr>
<td>S11</td>
<td>S23</td>
</tr>
<tr>
<td>S12</td>
<td></td>
</tr>
<tr>
<td>S13</td>
<td></td>
</tr>
<tr>
<td>S19</td>
<td></td>
</tr>
</tbody>
</table>

RMSE to observation standard deviation ratio (RSR) are used. NSE is a normalized index that defines the relative magnitude of the residual variance compared to that of the measured, and expressed as (Nash and Sutcliffe, 1970)

\[
\text{NSE} = 1 - \frac{\sum_{i=1}^{N} (Q_i^{\text{obs}} - Q_i^{\text{sim}})^2}{\sum_{i=1}^{N} (Q_i^{\text{obs}} - Q_i^{\text{mean}})^2}
\]  

where \(Q_i^{\text{obs}}\) and \(Q_i^{\text{mean}}\) are respectively the \(i\)th observed flow and its mean, \(Q_i^{\text{sim}}\) is the \(i\)th simulated flow, and \(N\) is the number of observations. NSE has a range between \(-\infty\) and 1, with its optimal value of 1. Negative values of NSE indicate poor performance of the model, suggesting that the observed mean is a better predictor of the flows than those of the model (Moriasi et al., 2007). The PBIAS is expressed as

\[
\text{PBIAS} = \frac{\sum_{i=1}^{N} (Q_i^{\text{obs}} - Q_i^{\text{sim}})}{\sum_{i=1}^{N} Q_i^{\text{obs}}} \times 100
\]  

The optimal value of PBIAS is 0, with low-magnitude values indicating good predictive ability of the model. Positive values indicate underestimation of the observed flows by the model, and conversely for negative values (Gupta et al., 1999). The RMSE-observation standard deviation ratio (RSR) is an error statistics that normalizes the RMSE with the standard deviation of the observed flows. It is expressed as

\[
\text{RSR} = \frac{\text{RMSE}}{\text{STDEV}_{\text{obs}}} = \sqrt{\frac{\sum_{i=1}^{N} (Q_i^{\text{obs}} - Q_i^{\text{sim}})^2}{\sum_{i=1}^{N} (Q_i^{\text{obs}} - Q_i^{\text{mean}})^2}}
\]  

The RSR is a normalized error index that utilizes the benefits of the RMSE which is one of the frequently used error index statistics. Excellent model performance will produce a zero value of RSR or RMSE (Moriasi et al., 2007).

4 Results and discussions

4.1 Principal Components Analysis (PCA)

The scaled variables of the physiographic and meteorological variables, expressed by Eq. (1), are incorporated into the PCA. In this way the correlation matrix of the variables is scale invariant. The correlations between the variables are summarized in Table 3. Italicized correlation coefficients are significant at 95% confidence level. 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 the PCA. The Glaeson-Staelin redundancy test yields a value of \(\phi = 0.395\), suggesting that there is considerable complexity in the data of these variables which warrants further examination using the PCA. The p-value from Bartlett’s sphericity test is 0.00000 which indicates that the null hypothesis can be rejected and there is significant strength in the relationship among the variables to warrant carrying a PCA.

The eigenvalues of the principal components are presented in Table 4, and making use of the Kaiser criterion allows for the retention of the first three principal components which account for 76% of the variation in the data. The factor loadings of these three principal components reflect the contributions and roles of these variables in the correlation of the data. The eigenvectors of the three loading factors are presented in Table 5. The factor loadings are the correlations between the variables and the factors. Factor 1 showed best correlation with the maximum elevation and the average elevation, whereas factor 3 showed best correlation with the
average slope and the minimum elevation. Figure 4 provides an illustration of the projection of the variables on a factor plane using an alternative criteria for the PCA of the variables. Each quadrant represents a similar group of variables. Further use of the PCA is made in order to identify groupings of sub-catchments by assessing the projection of cases onto the factor plane (Fig. 5). The figure establishes four groups of sub-catchments that are further reduced to two main groups or categories (Table 6) 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. All sub-catchments in group I, except S19, are characterized by low elevations, while those in group II are generally characterized with high elevations with weighted values that are above 1122 m. The

<table>
<thead>
<tr>
<th></th>
<th>Linear model</th>
<th>Nonlinear model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NSE</td>
<td>PBIAS</td>
</tr>
<tr>
<td>Group I (Flat)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan</td>
<td>0.933</td>
<td>0.097</td>
</tr>
<tr>
<td>Feb</td>
<td>0.932</td>
<td>0.032</td>
</tr>
<tr>
<td>Mar</td>
<td>0.930</td>
<td>−4.449</td>
</tr>
<tr>
<td>Apr</td>
<td>0.932</td>
<td>−0.030</td>
</tr>
<tr>
<td>May</td>
<td>0.932</td>
<td>0.023</td>
</tr>
<tr>
<td>Jun</td>
<td>0.934</td>
<td>0.023</td>
</tr>
<tr>
<td>Jul</td>
<td>0.934</td>
<td>0.020</td>
</tr>
<tr>
<td>Aug</td>
<td>0.933</td>
<td>−0.004</td>
</tr>
<tr>
<td>Sep</td>
<td>0.937</td>
<td>0.004</td>
</tr>
<tr>
<td>Oct</td>
<td>0.934</td>
<td>0.025</td>
</tr>
<tr>
<td>Nov</td>
<td>0.938</td>
<td>0.027</td>
</tr>
<tr>
<td>Dec</td>
<td>0.934</td>
<td>0.016</td>
</tr>
<tr>
<td>Group II (Hilly)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan</td>
<td>0.997</td>
<td>−1.060</td>
</tr>
<tr>
<td>Feb</td>
<td>0.998</td>
<td>0.006</td>
</tr>
<tr>
<td>Mar</td>
<td>0.998</td>
<td>0.009</td>
</tr>
<tr>
<td>Apr</td>
<td>0.998</td>
<td>−0.015</td>
</tr>
<tr>
<td>May</td>
<td>0.997</td>
<td>−0.905</td>
</tr>
<tr>
<td>Jun</td>
<td>0.991</td>
<td>−3.094</td>
</tr>
<tr>
<td>Jul</td>
<td>0.998</td>
<td>0.008</td>
</tr>
<tr>
<td>Aug</td>
<td>0.998</td>
<td>−0.042</td>
</tr>
<tr>
<td>Sep</td>
<td>0.998</td>
<td>−0.003</td>
</tr>
<tr>
<td>Oct</td>
<td>0.998</td>
<td>0.001</td>
</tr>
<tr>
<td>Nov</td>
<td>0.996</td>
<td>1.638</td>
</tr>
<tr>
<td>Dec</td>
<td>0.998</td>
<td>−0.006</td>
</tr>
</tbody>
</table>

Fig. 8. Predicted flows for April in Group II sub-catchments.

Fig. 9. RSR Performance statistics.
The support from the Applied Training Project of the Nile Basin Initiative (ATP/NBI) and WaterNet is highly appreciated. The authors extend also their gratitude to the reviewers of the initial version of this manuscript. The opinions and results presented in this paper are those of the authors and do not necessarily represent the donors or participating institutions.

Edited by: A. Castellarin

References


Fig. 10. NSE Performance statistics.


Sutcliffe, J. V. and Parks, Y. P.: The hydrology of the Nile, IAHS special publication No. 5, 1999.


