Hydrol. Earth Syst. Sci. Discuss., 7, 4089–4111, 2010
www.hydrol-earth-syst-sci-discuss.net/7/4089/2010/
doi:10.5194/hessd-7-4089-2010

Hydrology and
Earth System
Sciences
Discussions

# Hydrological characterization of watersheds in the Blue Nile Basin

**S. G. Gebrehiwot[1,4], U. Ilstedt[2], A. I. Gärdenas[3], and K. Bishop[1]**

[1]Departmen of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, P.O. Box 7050, 750 07 Uppsala, Sweden
[2]Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, 901 83 Umeå, Sweden
[3]Department of Soil and Environment, Swedish University of Agricultural Sciences, P.O. Box 7001, 750 07 Uppsala, Sweden
[4]Department of Geography and Environmental Studies, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia

Received: 31 May 2010 – Accepted: 13 June 2010 – Published: 2 July 2010

Correspondence to: S. G. Gebrehiwot (solomon.gebreyohannis@vatten.slu.se)

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

## Abstract

We made a hydrological characterization of 32 watersheds (31–4350 km$^2$) in the Blue Nile Basin, using data from a study of water and land resources in the Blue Nile Basin, Ethiopia published in 1964 by the US Bureau of Reclamation (USBR). The USBR document contains data on flow, climate, topography, geology, soil type, and land use for the period from 1959 to 1963. The aim of the study was to identify which combination of watershed variables best explain the variation in the hydrological regime, with special focus to low flow and, what kind of land use low flow might benefit from. Principal Component Analysis (PCA) and Partial Least Square (PLS) were used to analyze the relationship between hydrologic variables (total flow, maximum flow, minimum flow, runoff coefficient, and low flow index) and 30 potential watershed variables. We found that three groups of watershed variables – climate and topography, geology and soil, and land use had almost equal influence on the variation in the hydrologic variables ($R^2$ values ranging from 0.3 to 0.5). The individual variables which were selected based on statistical significance from all groups of explanatory variables were better in explaining the variation. Low flow was positively correlated most strongly to wetland, wood land, rainfall, luvisols, and alluvial soils. Low flow was negatively correlated to grazing land, bush land, tuffs/basalts, eutric-vertisols and riverine forest. We concluded that low flow benefits from the land use types that preserve soil quality and water storage, such as wetland, savannah and woodland, while it was lower in land use resulting in soil degradation. Therefore it provides support to the theory that some land use such as grassland, can promote higher low flow

## 1 Introduction

Despite having learned so much about hydrology, the complexity of watershed response to rainfall still defies ready prediction (McDonnell et al., 2007). General classes of watershed characteristics attributed to the differences in hydrological response to rainfall are watershed size, local climate, vegetation cover, soil properties, geology, anthropogenic activities, relief, and drainage characteristics (Black, 1997; Uhlenbrook, 2003; Sivapalan, 2005). Some watershed characteristics are more important than others in ways that are specific to different watersheds and scales. One way to advance in watershed hydrology is to classify watersheds based on most influential factors (McDonnell et al., 2007). Characterization of watersheds provides a basis for planning of land management issues in order to develop and secure water resources. Moreover, characterization of watersheds with reference to their spatial differences in hydrological response is a worthwhile complement to modeling (Yadav et al., 2007); since parameterization of the differences in characteristics of watersheds is one of the inhibitors to the progress of hydrological modeling (Hauhs and Lange, 2008).

It is vitally important to characterize hydrological response of watersheds in areas such as the Blue Nile where the well being of most of the population depends on the ability to manage water scarcity. Although 60% of Nile flow at Aswan comes from the Blue Nile, the local population can still suffer from water shortages during the dry season. There is a great interest in managing the landscape to reduce seasonal water shortages. But there are other factors that influence water availability besides those which can be *managed*, such as topography, geology and local climate. Management plans, and expectations for success in those plans need to be based on an understanding of the full range of factors controlling watershed response to rainfall, in addition to those that management can influence.

A characterization of watersheds with respect to the peak flows has been done for the Blue Nile Basin that delineates hydrologically homogenous units based on Hoskin L-moments for the purpose of flood frequency analysis in the basin (Abebe, 2007). Abebe (2007) found five generalized regional flood frequency curves for the basin. Adane and Foerch (2006) discussed characterization of watersheds for prediction of base flow index in *Wabi Shebele*, another $1.2\times10^5$ km$^2$ basin in Ethiopia which adjoins the Blue Nile Basin to the south-east and flows to the east of the country. This particular study comprises catchment size, stream density, climate index, hypsometric integral, Normalized Difference Vegetation Index (NDVI) and geological parameters to generate the base flow index. They concluded that negative correlation between base flow index and NDVI is highly pronounced. Given the importance of the Blue Nile, there remains scope for a more complete characterization of watersheds of the Blue Nile Basin, especially with regards to the critical low flow period.

Desirable as spatial characterization of watersheds in the Blue Nile is, it requires data. Systematic and extensive flow data were collected on the Blue Nile from 1959 to 1963 at the time of the building of the Aswan Dam on the Nile. The land and water resource investigation initiated by the Ethiopian government of that period and US Department of Interior's, Bureau of Reclamation (USBR). Hydrological data were collected from 59 gauge stations; and summaries of monthly flow data along with descriptions of soils, geology, land use and topography were published for 35 stations (US Department of Interior, 1964). This paper is the first effort to apply multivariate methods to these published data in order to characterize the hydrological response of the region (both high flows, annual flows and dry season flows) in terms of catchment geology, topography, land use, soil and climate.

The aim of this study is to determine which combinations of watershed variables best describe the variability of the hydrological regime, in particular low flow, for different watersheds in the Blue Nile Basin. Moreover, this study aims to identify variables which may be susceptible to management policies for developing and securing water resources in the dry periods.

## 2 Materials and methods

### 2.1 Site description of the Blue Nile Basin

The Blue Nile Basin, $2\times10^5$ km$^2$ in area size, contributes ca 60% of the Nile's flow at Aswan, Egypt; though the Blue Nile comprises only ca 8% of the Nile's catchment area. Rainfall in the Blue Nile ranged from 880 to 2070 mm per annum (1960–1964). Most of the basin area is covered with mountains. The catchments of the basin have a range of sizes, slopes, climatic patterns, topography, drainage patterns, geological formations, soils, vegetation cover and anthropogenic activities. There are three broad topographical divisions; the highland plateau, steep slopes, and the western low lands which are more flat lands. The basin is also characterized by three different groups of geological formations; the crystalline basement exposed over 32% of the area, sedimentary formations covering about 11% visible in the deep valleys of major southern tributaries, and finally the volcanic formations covering about 52% of the area in the north, central and east part of the basin (Ministry of Water Resources, 1998). The dominant soil type of the basin is clay. The special type of shrinking and expanding clay – *Vertisol*, covers about 15% of the basin.

## 2.2 The USBR study and selection of watersheds

From 1958 to 1963 the US Department of Interior, Bureau of Reclamation (USBR), and the Ethiopian Government at that time, studied the land and water resources of the Blue Nile Basin (Fig. 1). In 1964, the USBR published “Land and Water Resource of the Blue Nile” in 7 volumes, which comprises one main report and 6 appendixes: Plans and estimates, Hydrology, Geology, Land classification, Power, as well as Agriculture and Economics (US Department of Interior, 1964). We used data of the main report, and appendixes for Hydrology, Geology, and Land classification.

The USBR document contains data for 35 watersheds. All these 35 watersheds were considered in our study from the beginning. Three watersheds were excluded: two because they were comprised largely by other watersheds in the data set, and one because of insufficient flow data. The selected watersheds are documented in Table 1. The codes indicated in Table 1 are used in the subsequent graphs and tables.

### 2.3 Data

In total, 6 different hydrological variables and 30 different potential characterizing watershed variables were extracted from the USBR document. The potential characterizing variables of the watersheds are called simply watershed variables from here on.

#### 2.3.1 Hydrology, climate and geography

Hydrologic variables used in this study were average annual flow (Qt), minimum month flow (Ql), maximum month flow (Qh) and low flow index (LFI). The LFI is the ratio of minimum flow (during the driest month) to the total flow for the year. Flow data were transformed from the S.I. units (monthly cubic meters) to mm based on respective watershed sizes. Maximum and minimum flow, low flow index, and runoff coefficient were computed from the monthly flow data from 1959 to 1963 published in the USBR report. The runoff coefficient (C) is the proportion of total flow to rainfall.

The climatic data for this period, rainfall ($P$, mm/yr), potential evaporation (ET, mm/day) and temperature ($T$, °C/day) data were collected by the National Meteorological Service Agency of Ethiopia.

Geographical information was taken from the main report of the document. Geographical data includes latitude (Lat) and longitude (Long) (both in degree-decimal) of the centre of each watershed, area of the watersheds (Area, km$^2$), average elevation (El, m) and average slope class (Slope, %). All variables are summarized from text and map information in the USBR document.

#### 2.3.2 Geology, soil and land use types

While the hydrologic, climatic, relief and geographical data are all continuous, so-called numerical descriptors; the geology, soil and land use types are not continuous, so-called nominal descriptors. Moreover, the different classes within the groups of these variables are dependent on each other as the sum of the coverage of the different classes within a group is 100%. We used fuzzy coding – a special type of assigning dummy variables (Upton and Cook, 2006), to transform the nominal descriptors. For example, four different geological classes were identified; these were alluvials (Allu), sandstones/limestones (S/L stones), tuffs/basalts (Tu/Ba), and metamorphic stones (Meta). The four fractions of the geology classes should sum to one (100%) in each watershed. The same procedure was done for soil and land use types.

Nine different soil types were taken into account; (i) shallow leptosols (Lepto), soils with a very high content of coarse materials, (ii) cambisols (Cambi), a cambic B horizon or an umbric A horizon, (iii) arenosols (Areno), soils with albic, argillic, cambic or oxic, (iv) luvisols (Luvi), soils with an argillic horizon, and highly basic, (v) alisols (Ali), acidic and infertile soils, (vi) eutric-vertisols (EutV), clay rich soils, vii) regosols (Rego), soils formed on unconsolidated material except for recent alluvial deposits, (viii) acrisols (Acri), soils with an argillic B horizon and less basic and (ix) wetlands (Wet), soils having voids filled with water. Assigning numbers was done as for the geological groups using fuzzy coding.

Nine different land use classes were selected; (i) cultivated land (CU), arable land for seasonal crop production, (ii) dense wet forest (DWF), a humid highland forest where annual rainfall exceeds 1500 mm, (iii) dense dry forest (DDF), a montane forest where annual rainfall ranges from 500–1500 mm, (iv) open woodland (OWL), a sparsely distributed trees and shrubs with dominant grassland, (v) woodland (WL), a drought resistant vegetation type dominated by trees and shrubs, (vi) savannah grassland (SGL), lowland grasslands where the height of grasses are up to 1.3 m, (vii) grazing land (GL), a land designated for an open grazing system, (viii) riverine forest (RF), a continuous forest cover along river banks where the height of trees could be greater than 10 m, and ix) bush land (BL), a vegetation which includes shrubs, succulents and grasses with shallow degraded soil. Soil and land use components were directly taken as they were reported in the document, since temporal changes because of natural and human activities within the watersheds were assumed to be insignificant in the study period (1959–1963). Values of land use classes were also fuzzy-dummy variables.

## 2.4 Statistical analysis

### 2.4.1 Correlation

Correlation was used to analyze the relationships among the watershed variables, among the hydrological variables, as well as between the watershed and hydrological variables. Correlation was run using JMP (SAS Institute Inc., 2007).

### 2.4.2 Multivariate analysis

Principal Component Analysis (PCA) and Partial Least Square (PLS) were used to explore multivariate relationships in the watershed variables and hydrologic variables. PCA and PLS find new and independent latent variables (principal components) that maximize the explained variance among all variables (PCA) or the $y$-variables (PLS), where the $y$-variables are hydrologic variables in this case. In PCA, the two types of variables, hydrologic and watershed variables are treated together. While in the PLS analysis, the hydrologic variables were assigned as response variables, and the watershed variables are explanatory. All watershed variables were centered and scaled by their respective variance as these variables have different units and different data types. Transformation was tried on some of the variables. This was not found to be helpful for improving the explained variance; so transformations were not used.

The PCA and PLS runs were made in two steps. In the first step all explanatory variables were used to identify the most significant. Then, significant variables were used for the second stage in the PCA and PLS run. In both analyses, PCA and PLS, a variable was removed if it failed to be significant at the 90%-confidence-interval by Jack-Knifing in at least one of the model components. Significance, based on 95%-confidence-interval, was also determined. First, a PLS was conducted for all the response variables separately. Then, PLS models were run for two grouped response variables, (i) the total/high flow (total flow, maximum flow and runoff coefficient) and (ii) the low flow (minimum flow and low flow index) regimes. The explanatory variables

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

were also grouped into three conceptual groups: climate and topography/geography, geology and soil, as well as land use; then PLS was run for each conceptual explanatory group. This was done to assess if the different groups of explanatory variables differ in explaining the hydrologic variables. A comparison of these groups has been done based on the percent of variation of hydrologic variables explained by the different model ($R^2$ – a measure of fit of the data to the model in SIMCA). The assumed perfect limit for $R^2$ is 1. Both PCA and PLS were run using SIMCA 12.0.1 (UMETRICS, 2009).

## 3 Results

### 3.1 Correlation of variables

Many of the variables had less degree of correlation with others, Pearson correlation coefficient ($r$)$<|0.7|$. In Table 2 highly correlated variables, with $r \geq |0.7|$ were reported. Some of the highly correlated variables were obvious; like evapotranspiration and mean temperature, average elevation and mean temperature, or rainfall and dense wet forest. The remaining pairs of variables, which were not reported in Table 2 were correlated with $r<|0.7|$.

### 3.2 Multivariate analysis

In PCA and PLS runs significant and non-significant variables were identified. Then, significant variables from each of four components were recruited. The non-significant variables were listed in Table 3. Then, final plots of PCA and PLS were done excluding these non-significant variables.

#### 3.2.1 Principal component analysis (PCA)

Most of the variables (ca 60% of all the variables) were significant in explaining the variation between the watersheds with respect to the hydrological regime (Table 3 and Fig. 2). The total variation explained by the first two components was $R^2$=0.36 (Fig. 2). Total flow, maximum flow and runoff coefficient increased with increment of tuffs/basalts, alisols and grazing land, and decreased with the area of wetland, woodland and latitude. Minimum flow increased with increased rainfall, dense wet forest and alisols, though the positive relationship was weak. Low flow was negatively correlated to dense dry forest, wetland, woodland, latitude, alluvials, regosols, bush land, eutric-vertisols, cultivated land and luvisols. The LFI, which normalizes the low flow to annual flow, was positively related to savannah grassland, arenosols, wetland, woodland and latitude; and negatively related to grazing land and tuffs/basalts.

There are also important relationships existing between land use variables. There were negative correlations between grazing land and savannah as well as between wetland and woodland; the same was true for cultivated land in relation to both riverine forest and open woodland.

### 3.2.2 Partial Least Square (PLS)

The variables explaining the variation and potential predictors of the total/high flow regime were grazing land, tuffs/basalts, rainfall, alisols, leptosols, regosols, area, and eutric-vertisols (Fig. 3). The cumulative variation explained by the first two components was $R^2$=0.45.

Land use variables woodland, dense wet forest, wetland and bush land were significant for explaining the variation in the low flow regime (Ql, and LFI). The significant soil, climate, location and topography variables were alluvials, rainfall, luvisols, eutric-vertisols, longitude, tuffs/basalts and slope (Fig. 3). The variability explained by the first two components was $R^2$=0.44.

In the PLS analysis of expanatory variables (Table 3) ranifall was positively related with both groups of hydrologic variables. Geology, soils and landuse variables, however, showed some contrasting relationships to the high and low flow regimes. Some variables that were positively related to the total/high flow group were negatively correlated to the low flow group and vice versa, such as alisols, alluvials, tuffs/basalts, grazing land and woodland. Comparison of the groups of the explanatory variables showed that the geology and soil group better explained the variation and predictions in the total/high flow group. All groups of variables (climate and topography, land use, geology and soils) had almost an equal influence when it came to the variation and prediction of the low flow regime (Table 3).

## 4 Discussion

Different watersheds have different properties which influence the flow regime. Knowledge of this can help in planning action on variables which need management measures. To facilitate this, it is helpful to classify and regionalize a basin based on similarity and differences of their relation to the hydrologic regimes. Recent developments in hydrological science seek to find a way for understanding the interconnection of hydrological processes at the catchment scale by characterizing watershed variables (Sivapalan, 2005). This study addressed characterization of different watershed variables, relationship of variables to each other and to hydrologic variables, as well as possible predictors for the hydrological regime in the Blue Nile Basin. A land and water resource study document on the Blue Nile Basin produced by the US Department of Interior, Bureau of Reclamation (USBR) in 1964, which comprises data and information from 1959–1963 has been used for this study.

Among the 30 watershed variables, 13 were highly correlated to each other (Table 2). The PCA analysis explained 60% of the variation in the high flow regime. Of that explained variation, the watershed variables explained 80% (or roughly half of the total variation). Each of the variables could have possible explanation for being positively or negatively related to total/high flow regime. Vegetation with high leaf area, deep roots and long growing season normally has high total water loss by transpiration and interception (Bruijnzeel, 2004). For the variables related positively to the total/high flow regime, flow could be faster with less infiltration, and therefore decreasing groundwater storage. While, for the variables related negatively to total/high flow, water is flowing slowly with higher infiltration rate. Cultivation and woodland are the main correlates to reduced total/high flow regime. So management to promote water infiltration, such as woodland, could result in lower high flows, which is in line with other studies as summarized by Bruijnzeel (2004).

The major variables influencing the low flow regime (Ql and LFI) were wetland, wood land, rainfall, luvisols, and alluvial positively; while, grazing land, bush land, tuffs/basalts, eutric-vertisols, riverine forest and slope negatively correlated to the low flow regime. Alluvial soils and wetlands could allow for better water storage in the watershed. Grazing land and bush land are features of land degradation in the region (Demel and Tesfaye, 2002), which promote rapid runoff. The reduction of low flow with increment of slope and eutric-vertisols is likely related with promotion of rapid runoff in steep terrain. As for the riverine forests, Adane and Foerch (2006) also found that base flow index was negatively related to normalized difference vegetation index (NDVI); where they analyzed near-stream forest for the vegetation index. This could be related to higher water uptake by riverine forest during dry season.

The summarized PLS results from different groups of explanatory variables showed how the variables were doing in regard to explaining the hydrologic variables respective to each group of explanatory variable. From the climate and topography group rainfall was the single factor most positively correlated with both total/high flow and low flow regimes. Moreover, latitude and longitude had a negative influence on both flow regimes. Longitude though, could be replaced by elevation or evapotranspiration, because of their high correlation (Table 2).

In contrast to rainfall; geology, soil and land use were positively correlated to one of the flow regimes, and negatively to the other. The three groups of explanatory variables (climate and topography, geology and soil, and land use) have almost equal strength in influencing the variability of flow regimes, and also in predicting the flow regimes. It should be also noted that separate groups of variables (climate and geography, or geology and soil, or land use) explain less variation than selected variables from all groups based on level of significance: almost two times the variation explained by the significant variables – $R^2$ increased from 0.3 to 0.6 (Table 3).

Forest, woodlands and savannah were negatively correlated to total/high flow and positively correlated to the low flow regime. There are also forest related land uses like, bush land and riverine forest, which were negatively related to the low flow regime. Dense wet forest, woodland and savannah grasslands were the three dominant land use features positively related to low flow regime. The betterment of low flow regime because of woodland and grasslands was hypothesized for tropical areas by Bruijnzeel (2004) and Malmer et al. (2009). Land use such as grazing land and bush land are features which others have found to indicate degradation of land resources (forests and soils) (Gete, 2000; Demel and Tesfaye, 2002). It should be noted that grazing land was expanding at the expense of wetland and woodland; while cultivated land was expanding at the expense of riverine forest and open woodland (Fig. 2). So, land use management tradeoff between woodland, savannah and wetland on the one hand and bush land and grazing on the other, need to be considered for betterment of low flow management. Woodland, grassland and wetland ecosystems should be conserved well for development and sustenance of low flow in the landscape. Low flow has been found as a key risk for existing and future water availability (Smahktin, 2001).

Low flow was also positively related with rainfall (Figs. 2 and 3). Though rainfall is the source of water for every flow regime, some watershed variables have a better ability to retain rainfall by soil infiltration rather to lose the rain to rapid runoff. This study found that woodland, grassland and wetland accompanied by deep soil and/or large catchment size were positively related to a catchment's ability to intercept and retain rainfall for later dry season flow. However, this needs to be recalibrated to recent hydrological systems half a century after the USBR study, as most of the soil resource in the region has been degraded through erosion during recent decades (Gete and Hurni, 2001).

## 5 Conclusions

Our analysis provide some rare data which supports the highly debated theory that some land use, such as woodland and savannah, can promote higher low flow even if the total/high flow is decreased. We concluded that for sustaining water availability in the dry periods, conservation of woodland, savannah and wetland is useful, while grazing land, bush land and riverine forest should not be promoted in areas where dry season flow is a problem.

All groups of explanatory variables (climate and topography, geology and soil, and land use) were about equally important in explaining and predicting the hydrologic regime. The individual variables which were selected based on statistical significance from all groups of explanatory variables, in general, were better in explaining the variation, and showed a wider range of explanation of the variation of total/high flow and low flow.

A hydrological characterization like this one could be used as a foundation for regionalization of watershed management in the basin. It could be especially useful for the management of ungauged watersheds. There is a great interest in managing the landscape to reduce seasonal water shortages. But there are other factors that influence water availability besides those which can be "managed", such as topography, geology and local climate. Management plans, and expectations for success in those plans needs to be based on an understanding of the full range of factors controlling watershed response to rainfall, in addition to those that management can influence. As one tries to see human impacts on land use/land cover, there is a need to be able to separate factors such as geology and local climate since there is less that management plans can do to influence these.

Land management cannot change geological formations that are not conducive to infiltration. But this study did show that factors, which can be influenced by land management, such as extent of dense forest or woodland or grassland can enhance the low flow regime. This study also showed that map information, such as relief and

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

soil type can identify, where low flow could be more of a problem for local communities. While this is to be expected, it is still something that has not been documented, and this documentation might serve as the basis for aiding the calibration of hydrological models for management purposes. The results could also be used as a basis for or a complement to process-oriented hydrological models, which cover the temporal variability of specific catchments.

## References

Abebe, S.: Identification and delineation of hydrological homogeneous regions – the case of Blue Nile River Basin, M.Sc. Thesis, ArbaMinch University, Ethiopia, 2007.

Adane, A. and G. Foerch.: Catchment characterization as predictors of baseflow index (BFI) in Wabi Shebele River Basin, East Africa, Conference on International Agricultural Research for Development, University of Bonn, Germany, 11–13 Oct 2006, 2006.

Black, P. E.: Watershed functions, J. Am. Water Resour. As., 33(1), 1–11, 1997.

Bruijnzeel, L. A.: Hydrological functions of tropical forests: not seeing the soil for the trees?, Agr. Ecosyst. Environ., 104, 185–228, 2004.

Demel, T. and Tesfaye, B.: State of Forests and Forestry Research in Ethiopia, Indicators and Tools for Restoration and Sustainable Management of Forests in East Africa, I-TOO working paper No. 1, Addis Ababa, Ethiopia, 2002.

Gete, Z. and Hurni, H.: Implications of land use and land cover dynamics for mountain resource degradation in the Northwestern Ethiopian Highlands, Mt. Res. Dev., 21(2), 184–191, 2001.

Hause, M. and Lange, H.: Classification of runoff in headwater catchments: a physical problem?, Geography Compass, 2(1), 235–254, 2008.

Malmer, A., Murdiyarso, D., Bruijnzeel, L. A. (Sampurno), and Ilstedt, U.: Carbon sequestration in tropical forests and water: a critical look at the basis for commonly used generalizations, Glob. Change Biol., 16(2), 599–604, 2010.

McDonnell, J. J., Sivapalan, M., Vache, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M. L., Selker, J., and Weiler, M.: Moving beyond heterogeneity and process complexity: a new vision for watershed hydrology, Water Resour. Res., 43, 1–6, 2007.

Ministry of Water Resources: Abbay River Basin Integrated Development Master Plan Project: Data Collection – Site Investigation Survey and Analysis, Phase 2. BECOM in collaboration with BRGM and ISL Consulting Engineers, Ministry of Water Resources, Ethiopia, Addis Ababa, Ethiopia, 1998.

SAS Institute Inc.: JMP® 7.0.1, SAS Campus Drive, Cary, North Carolina, 2007.

Sivapalan, M.: Pattern, process and function: elements of a unified theory of hydrology at the catchment scale, in: Encyclopedia of Hydrological Sciences, edited by: Anderson, M. G., John Wiley & Sons, Ltd, 2005.

Smakhtin, V. U.: Low flow hydrology: a review, J. Hydrol., 240, 147–186, 2001.

Uhlenbrook, S.: An empirical approach for delineating spatial units with the same dominating runoff generation processes, Phys. Chem. Earth, 28, 297–303, 2003.

Umetrics AB: SIMCA P+®, Version 12.0.1.0, April, 2009.

Upton, G. and Cook, I.: A Dictionary of Statistics, Oxford University Press, London, UK, 2006.

US Department of the Interior: Land and water resources of the Blue Nile Basin, Ethiopia, Appendixes III, IV, and V, Bureau of Reclamation, US, 1964.

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.

Zeleke, G.: Landscape dynamics and soil erosion process modelling in the Northwestern Ethiopian highlands, Ph.D. thesis, in: African Studies Series, vol. A16, Geographica Bernensia, Berne, Switzerland, 2000.

**Table 1.** Watersheds and their respective variables as extracted from the USBR document (US Department of Interior, 1964)[a].

| Watershed<br>Name | Code | Qt | Ql | Qh | LFI | *C* | *P* | *T* | ET | El | Slope | Allu | S/L stone | Tu/Ba | Meta | GL | CU | OWL | DWF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gilgel Abbay | GA | 981.5 | 3.2 | 355.5 | 0.0032 | 0.63 | 1562 | 17.0 | 3.4 | 2250 | 12.5 | 0 | 0 | 1 | 0 | 0.2 | 0.5 | 0.1 | 0 |
| Koga | Ko | 578.7 | 5.0 | 224.3 | 0.0086 | 0.37 | 1562 | 17.0 | 3.4 | 2250 | 7.5 | 0.1 | 0 | 0.9 | 0 | 0.2 | 0.6 | 0 | 0 |
| Gummera | Gu | 747.6 | 4.1 | 380.3 | 0.0055 | 0.51 | 1460 | 17.0 | 3.4 | 2900 | 7.5 | 0.1 | 0 | 0.9 | 0 | 0.1 | 0.5 | 0.2 | 0 |
| Megetch | Me | 240.9 | 0.5 | 99.1 | 0.0020 | 0.21 | 1170 | 20.4 | 3.8 | 2250 | 22.5 | 0.05 | 0 | 0.95 | 0 | 0.1 | 0.8 | 0.1 | 0 |
| Angereb | An | 374.8 | 2.9 | 168.0 | 0.0078 | 0.30 | 1264 | 17.0 | 3.4 | 2000 | 12.5 | 0.1 | 0 | 0.9 | 0 | 0.1 | 0.7 | 0.05 | 0 |
| Ribb | Ri | 235.9 | 0.2 | 228.4 | 0.0010 | 0.16 | 1460 | 17.0 | 3.4 | 2500 | 10 | 0.15 | 0.05 | 0.8 | 0 | 0.1 | 0.5 | 0.2 | 0 |
| Wizer | Wi | 711.7 | 0.2 | 355.8 | 0.0002 | 0.80 | 887 | 14.2 | 3.3 | 2800 | 5 | 0 | 0 | 1 | 0 | 0.2 | 0.5 | 0 | 0 |
| Beressa | Be | 509.4 | 1.2 | 240.0 | 0.0024 | 0.57 | 887 | 14.2 | 3.3 | 2800 | 15 | 0 | 0 | 1 | 0 | 0.2 | 0.5 | 0 | 0 |
| Muger_chancho | Mc | 289.2 | 0.9 | 137.9 | 0.0030 | 0.24 | 1192 | 13.3 | 3.1 | 2800 | 5 | 0.2 | 0 | 0.8 | 0 | 0.3 | 0.3 | 0 | 0 |
| Guder_Jibat | Gj | 648.4 | 2.5 | 170.3 | 0.0039 | 0.48 | 1352 | 13.3 | 3.1 | 2800 | 7.5 | 0.2 | 0 | 0.8 | 0 | 0.1 | 0.4 | 0 | 0 |
| Guder_Guder | Gg | 792.5 | 1.8 | 318.4 | 0.0023 | 0.59 | 1352 | 13.3 | 3.1 | 2800 | 7.5 | 0.2 | 0 | 0.8 | 0 | 0.15 | 0.45 | 0 | 0 |
| Guder_Melke | Gm | 510.0 | 2.4 | 167.3 | 0.0047 | 0.38 | 1352 | 13.3 | 3.1 | 2800 | 5 | 0.2 | 0 | 0.8 | 0 | 0.15 | 0.45 | 0 | 0 |
| Guder_Fato | Gf | 613.2 | 1.8 | 218.1 | 0.0030 | 0.45 | 1352 | 13.3 | 3.1 | 2800 | 5 | 0.2 | 0 | 0.8 | 0 | 0.05 | 0.45 | 0 | 0 |
| Guder_Idris | Gi | 866.7 | 6.6 | 370.3 | 0.0076 | 0.64 | 1352 | 13.3 | 3.1 | 2800 | 5 | 0.2 | 0 | 0.8 | 0 | 0.2 | 0.5 | 0 | 0 |
| Guder_Bello | Gb | 819.1 | 1.8 | 294.3 | 0.0022 | 0.61 | 1352 | 13.3 | 3.1 | 2800 | 12.5 | 0.2 | 0 | 0.8 | 0 | 0.2 | 0.5 | 0 | 0 |
| Fincha | Fi | 330.3 | 3.1 | 89.1 | 0.0095 | 0.32 | 1025 | 22.5 | 3.9 | 1750 | 22.5 | 0.45 | 0.05 | 0.5 | 0 | 0.1 | 0.3 | 0 | 0 |
| Beles | Bs | 313.4 | 0.7 | 139.8 | 0.0023 | 0.32 | 971 | 23.9 | 4.0 | 750 | 5 | 0 | 0 | 0.75 | 0.25 | 0 | 0.2 | 0.2 | 0 |
| Birr_Jiga | Bj | 668.3 | 0.2 | 364.0 | 0.0004 | 0.49 | 1355 | 19.2 | 3.5 | 2750 | 22.5 | 0 | 0 | 1 | 0 | 0.2 | 0.4 | 0.1 | 0 |
| Birr_Temcha | Bt | 1291.0 | 2.5 | 537.2 | 0.0019 | 0.79 | 1644 | 19.2 | 3.5 | 2750 | 18.5 | 0 | 0 | 1 | 0 | 0.3 | 0.45 | 0 | 0 |
| Temim | Te | 827.4 | 0.3 | 396.7 | 0.0003 | 0.61 | 1355 | 19.2 | 3.5 | 2750 | 18.5 | 0 | 0 | 1 | 0 | 0.15 | 0.45 | 0 | 0 |
| Arera | Ar | 477.9 | 2.9 | 223.9 | 0.0061 | 0.35 | 1355 | 19.2 | 3.5 | 2750 | 7.5 | 0 | 0 | 1 | 0 | 0.15 | 0.45 | 0 | 0 |
| Leza | Le | 415.2 | 2.1 | 252.8 | 0.0052 | 0.31 | 1355 | 19.2 | 3.5 | 2750 | 18.5 | 0 | 0 | 1 | 0 | 0.1 | 0.4 | 0.2 | 0 |
| Ketchem | Kt | 1107.3 | 0.9 | 455.6 | 0.0008 | 0.82 | 1355 | 19.2 | 3.5 | 2750 | 18.5 | 0 | 0 | 1 | 0 | 0.2 | 0.4 | 0.1 | 0 |
| Dabana | Da | 590.6 | 2.6 | 213.4 | 0.0044 | 0.29 | 2071 | 22.6 | 3.8 | 1000 | 5 | 0 | 0 | 1 | 0 | 0 | 0.3 | 0.2 | 0.2 |
| Angar | Ag | 611.0 | 3.0 | 191.6 | 0.0049 | 0.30 | 2071 | 22.6 | 3.8 | 1000 | 7.5 | 0 | 0.25 | 0.25 | 0.5 | 0.1 | 0.3 | 0.2 | 0.1 |
| Chemoga | Ch | 716.7 | 1.7 | 294.8 | 0.0023 | 0.53 | 1355 | 15.7 | 3.4 | 2750 | 18.5 | 0.15 | 0 | 0.85 | 0 | 0.2 | 0.4 | 0 | 0 |
| Djilil | Dj | 480.5 | 1.9 | 217.4 | 0.0039 | 0.35 | 1355 | 15.7 | 3.4 | 2750 | 18.5 | 0.15 | 0 | 0.85 | 0 | 0.1 | 0.6 | 0 | 0 |
| Kulch | Ku | 1157.8 | 4.4 | 574.8 | 0.0038 | 0.70 | 1655 | 15.7 | 3.4 | 2750 | 18.5 | 0.15 | 0 | 0.85 | 0 | 0.2 | 0.5 | 0 | 0 |
| Jedeb | Je | 1100.2 | 3.8 | 517.8 | 0.0035 | 0.35 | 1355 | 15.7 | 3.4 | 2750 | 18.5 | 0.15 | 0 | 0.85 | 0 | 0.2 | 0.4 | 0 | 0 |
| Fettaru | Fe | 1725.6 | 7.0 | 676.1 | 0.0041 | 0.70 | 2071 | 17.9 | 3.3 | 1750 | 10 | 0 | 0 | 1 | 0 | 0.1 | 0.4 | 0.1 | 0.2 |
| Sifa | Si | 754.1 | 3.6 | 177.9 | 0.0047 | 0.81 | 2071 | 17.9 | 3.3 | 1750 | 12.5 | 0 | 0 | 1 | 0 | 0.2 | 0.3 | 0.1 | 0.3 |
| Wama | Wa | 604.1 | 2.9 | 194.9 | 0.0048 | 0.83 | 2071 | 17.9 | 3.3 | 1750 | 12.5 | 0 | 0 | 1 | 0 | 0.2 | 0.3 | 0.1 | 0.3 |

**Table 1.** Continued.

| Watershed | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Name | Code | DDF | SGL | WL | RF | BL | Lepto | Cambi | Areno | Luvi | Ali | EutV | Rego | Acri | Wet | Area | Long | Lat |
| Gilgel Abbay | GA | 0.1 | 0 | 0.05 | 0.05 | 0 | 0 | 0 | 0 | 0.8 | 0 | 0 | 0.2 | 0 | 0 | 1660 | 37.1 | 11.5 |
| Koga | Ko | 0 | 0 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0.9 | 0 | 0 | 0 | 0 | 0.1 | 266 | 37.1 | 11.5 |
| Gummera | Gu | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1239 | 37.6 | 11.8 |
| Megetch | Me | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 519 | 37.5 | 12.7 |
| Angereb | An | 0 | 0.1 | 0 | 0.05 | 0 | 0.9 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0 | 0 | 660 | 37.6 | 11.5 |
| Ribb | Ri | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.2 | 0 | 0.8 | 0 | 0 | 0 | 1497 | 37.7 | 12.0 |
| Wizer | Wi | 0 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0 | 0.8 | 0.2 | 0 | 0 | 0 | 60 | 38.8 | 10.3 |
| Beressa | Be | 0 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0 | 0 | 0.8 | 0.2 | 0 | 0 | 220 | 38.7 | 9.6 |
| Muger_chancho | Mc | 0.3 | 0.1 | 0 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0 | 0.8 | 0.1 | 0 | 0 | 606 | 38.7 | 9.3 |
| Guder_Jibat | Gj | 0.3 | 0.1 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0.3 | 0.5 | 0.2 | 0 | 0 | 0 | 143 | 37.9 | 9.0 |
| Guder_Guder | Gg | 0.3 | 0 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0.3 | 0.45 | 0.25 | 0 | 0 | 0 | 499 | 37.9 | 9.0 |
| Guder_Melke | Gm | 0.3 | 0 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0.35 | 0.5 | 0.15 | 0 | 0 | 0 | 80 | 37.9 | 8.9 |
| Guder_Fato | Gf | 0.3 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0.3 | 0.5 | 0.2 | 0 | 0 | 0 | 98 | 37.8 | 8.9 |
| Guder_Idris | Gi | 0.25 | 0 | 0 | 0 | 0.05 | 0 | 0 | 0 | 0.3 | 0.45 | 0.25 | 0 | 0 | 0 | 76 | 37.9 | 9.0 |
| Guder_Bello | Gb | 0.25 | 0 | 0 | 0 | 0.05 | 0 | 0 | 0 | 0.3 | 0.45 | 0.25 | 0 | 0 | 0 | 244 | 37.8 | 8.9 |
| Fincha | Fi | 0.2 | 0.2 | 0.1 | 0.1 | 0 | 0 | 0.2 | 0.2 | 0 | 0.2 | 0.2 | 0 | 0 | 0.2 | 1390 | 37.5 | 9.5 |
| Beles | Bs | 0.1 | 0.2 | 0 | 0 | 0.3 | 0.2 | 0.5 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 3520 | 36.5 | 11.2 |
| Birr_Jiga | Bj | 0.2 | 0 | 0 | 0.1 | 0 | 0.1 | 0 | 0 | 0.2 | 0.45 | 0.2 | 0 | 0 | 0.05 | 813 | 37.5 | 10.6 |
| Birr_Temcha | Bt | 0.1 | 0 | 0 | 0.15 | 0 | 0.05 | 0 | 0 | 0.1 | 0.6 | 0.15 | 0 | 0 | 0.1 | 350 | 37.6 | 10.2 |
| Temim | Te | 0.2 | 0.1 | 0 | 0.1 | 0 | 0.1 | 0 | 0 | 0.2 | 0.4 | 0.2 | 0.1 | 0 | 0 | 108 | 37.4 | 10.7 |
| Arera | Ar | 0.2 | 0.1 | 0 | 0.1 | 0 | 0.1 | 0.2 | 0 | 0.2 | 0 | 0.2 | 0.3 | 0 | 0 | 31 | 37.3 | 10.7 |
| Leza | Le | 0.2 | 0 | 0 | 0.1 | 0 | 0.1 | 0 | 0 | 0.2 | 0.4 | 0.2 | 0.1 | 0 | 0 | 159 | 37.4 | 10.7 |
| Ketchem | Kt | 0.1 | 0.1 | 0 | 0.1 | 0 | 0.1 | 0 | 0 | 0.2 | 0.7 | 0 | 0 | 0 | 0 | 183 | 37.6 | 10.6 |
| Dabana | Da | 0.2 | 0 | 0 | 0.1 | 0 | 0.1 | 0 | 0 | 0 | 0.8 | 0 | 0 | 0.1 | 0 | 3080 | 36.3 | 8.4 |
| Angar | Ag | 0.2 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0 | 0 | 0.7 | 0 | 0 | 0.3 | 0 | 4350 | 36.4 | 9.4 |
| Chemoga | Ch | 0.1 | 0 | 0.1 | 0.1 | 0.1 | 0 | 0.1 | 0 | 0 | 0.8 | 0.1 | 0 | 0 | 0 | 320 | 37.8 | 10.3 |
| Djilil | Dj | 0.05 | 0.1 | 0 | 0.1 | 0.05 | 0 | 0 | 0 | 0 | 0.7 | 0.2 | 0.05 | 0 | 0.05 | 70 | 37.5 | 10.8 |
| Kulch | Ku | 0.1 | 0.1 | 0 | 0.1 | 0 | 0.05 | 0 | 0 | 0 | 0.55 | 0.3 | 0.05 | 0 | 0.05 | 50 | 37.3 | 11.0 |
| Jedeb | Je | 0.1 | 0 | 0 | 0.1 | 0.2 | 0 | 0.1 | 0 | 0 | 0.8 | 0.1 | 0 | 0 | 0 | 250 | 37.7 | 10.5 |
| Fettaru | Fe | 0 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.7 | 0.3 | 0 | 0 | 0 | 200 | 36.7 | 9.1 |
| Sifa | Si | 0 | 0.1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.8 | 0.2 | 0 | 0 | 0 | 978 | 36.8 | 8.9 |
| Wama | Wa | 0 | 0.1 | 0 | 0 | 0 | 0 | 0.1 | 0 | 0 | 0.8 | 0.1 | 0 | 0 | 0 | 764 | 36.6 | 8.9 |

[a] Codes of the variables represent; Qt total flow, Qh maximum flow, Ql minimum flow, LFI low flow index, *C* runoff coefficient, *P* rainfall, *T* temperature, ET evapotranspiration, El elevation, slope average watershed slope, Tu/Ba tuffs and basalts, S/L stone sand/lime stone, Meta metamorphic rock, Allu Alluvial, GL grazing land, CU cultivated land, OWL open woodland, DWF dense wet forest, DDF dense dry forest, SGL savannah grassland, WL woodland, BL (degraded) bush land, RF riverine forest, Lepto leptosols, Cambi cambisols, Rego regosols, Areno arenosols, Acri acrisols, Ali alisols, EutV Eutric Vertisols, Luvi luvisols, Wet Wetland, Area watershed area, Long Longitude, Lat Latitude.

**Table 2.** Pearson correlation coefficients ($r$) for pairs of hydrologic and watershed variables with an $r \geq |0.7|$.

| Variable | by Variable | Correlation | |
|---|---|---|---|
| Hydrologic variables | Qh | Qt | 0.9 |
| | Ql | LFI | 0.8 |
| Watershed variables | ET | $T$ | 0.9 |
| | El | $T$ | −0.7 |
| | Tu/Ba | S/L stone | −0.8 |
| | Meta | S/L stone | 0.9 |
| | Tu/Ba | Meta | −0.7 |
| | $P$ | DWF | 0.8 |
| | Acri | S/L stone | 0.9 |
| | Acri | Meta | 0.8 |
| | Wet | Areno | 0.8 |
| | Meta | Area | 0.8 |
| | Acri | Area | 0.7 |
| | $P$ | Long | −0.7 |
| | $T$ | Long | −0.7 |
| | El | Long | 0.8 |

**Table 3.** Non-significant watershed variables identified from 4 components for both PCA and PLS and, the relation of groups of explanatory variables to explain response variables (flow parameters), as summarized from PLS analysis[a].

| | Non significant variables | | | Group of variables | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | PCA | PLS_total/ high flow | PLS_low flow | Total/high flow regime | | | Low flow regime | | |
| | | | | Climate and topography | Geology and soil | Land use | Climate and topography | Geology and soil | Land use |
| | S/L stone | $T$ | El | $+P$ | +Ali | +GL | $+P$ | +Luvi | +WL |
| | Meta | Slope | S/L stone | +Slope | +Tu/Ba | +DWF | −Slope | +Allu | +SGL |
| | SGL | Allu | Meta | +El | −Acri | −DDF | −Lat | +Wet | +DWF |
| | BL | S/L stone | GL | $-T$ | −Cambi | −OWL | −Long | +Areno | −BL |
| | Lepto | Meta | DDF | −ET | −Allu | −WL | −El | +S/L stone | −RF |
| | Cambi | Acri | SGL | −Area | −Rego | −CU | | −Ali | −GL |
| | Acri | OWL | Areno | −Lat | −EutV | | | −Tu/Ba | −OWL |
| | | WL | Acri | −Long | | | | −EutV | |
| | | DDF | Lepto | | | | | −Lepto | |
| | | SGL | Cambi | | | | | | |
| | | RF | Ali | | | | | | |
| | | Areno | Rego | | | | | | |
| | | Luvi | Lat | | | | | | |
| | | Wet | | | | | | | |
| $R^2$ | 0.6 | 0.6 | 0.6 | 0.3 | 0.5 | 0.3 | 0.3 | 0.3 | 0.3 |

[a] "+" sign indicated positive relationship between hydrologic and watershed variables, while "–" indicates inverse relationship.

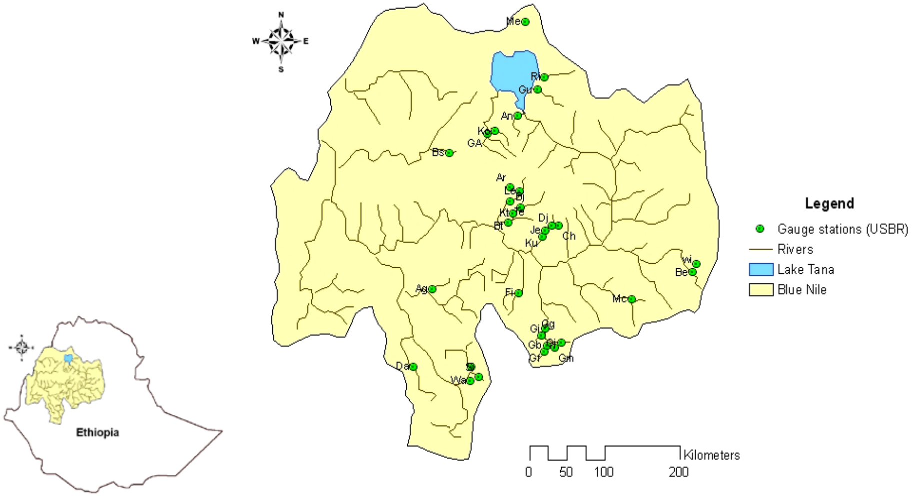

**Fig. 1.** Location of USBR hydrological stations in Blue Nile basin included in this study.

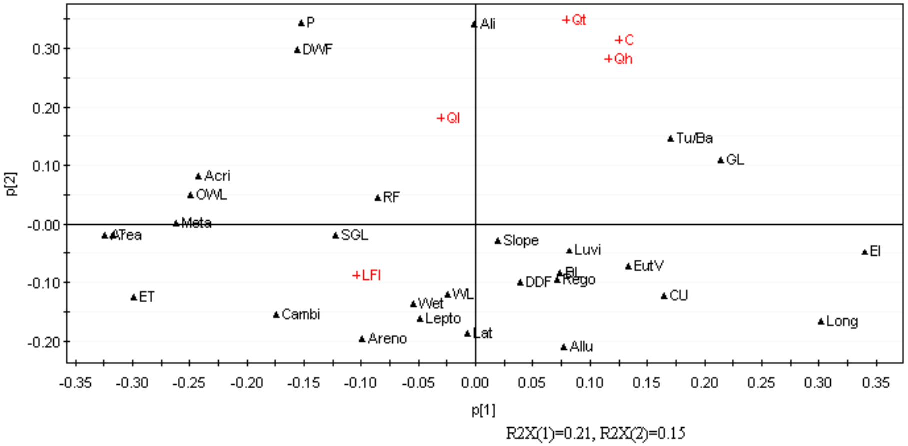

**Fig. 2.** PCA plot with selection of major watershed variables.

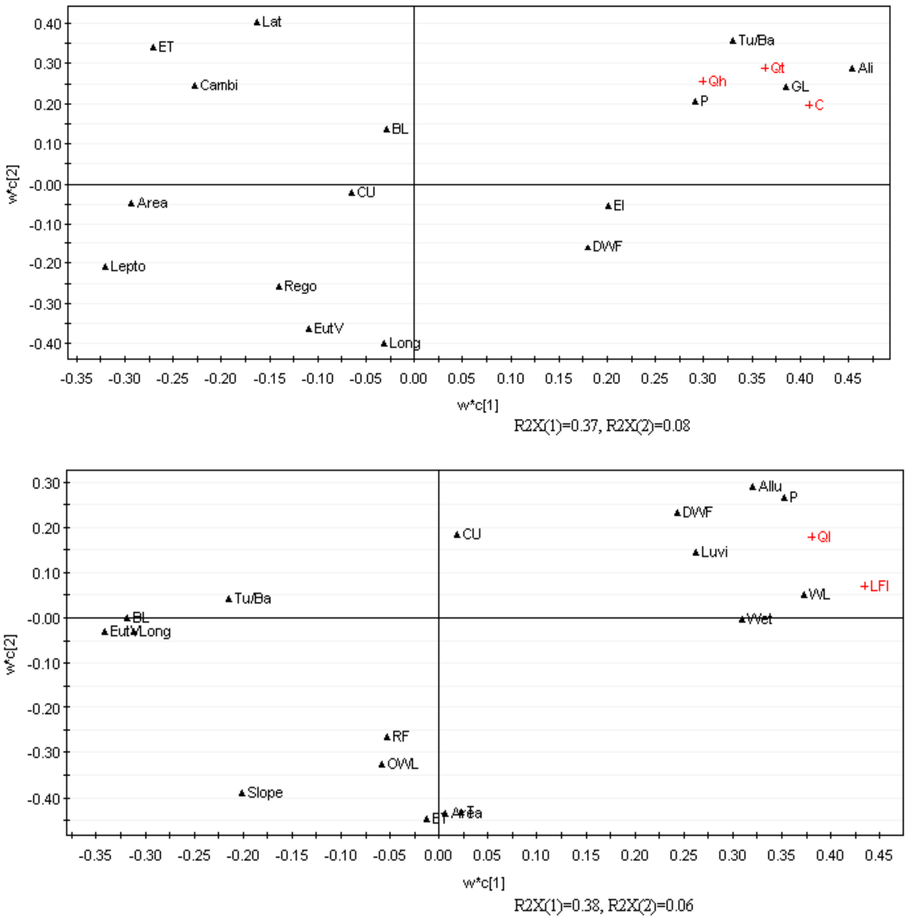

**Fig. 3.** PLS loading plots for total/high flow group (the first pane) and low flow group (second panel) with significant watershed variables.