



# An elusive search for regional flood frequency estimates in the River Nile basin

P. Nyeko-Ogiramoi<sup>1,2</sup>, P. Willems<sup>1</sup>, F. M. Mutua<sup>3</sup>, and S. A. Moges<sup>4</sup>

<sup>1</sup>Katholieke Universiteit Leuven, Hydraulics Section, Kasteelpark Arenberg 40, 3001 Leuven, Belgium

<sup>2</sup>Ministry of Water and Environment, Directorate of Water Development, Rural Water Department, P.O. Box 20026, Kampala, Uganda

<sup>3</sup>University of Nairobi, Department of Meteorology, P.O. Box 30197, 00100, Nairobi, Kenya

<sup>4</sup>Addis Ababa University, AAiT, Department of Civil Engineering, Climate, Hydrology and Water Resources Systems Chair, P.O. Box 385, Addis Ababa, Ethiopia

Correspondence to: P. Nyeko-Ogiramoi (nyeko\_ogiramoi@yahoo.co.uk)

Received: 17 January 2012 – Published in Hydrol. Earth Syst. Sci. Discuss.: 1 March 2012

Revised: 28 July 2012 – Accepted: 1 August 2012 – Published: 6 September 2012

**Abstract.** Estimation of peak flow quantiles in ungauged catchments is a challenge often faced by water professionals in many parts of the world. Approaches to address such problem exist, but widely used techniques such as flood frequency regionalisation is often not subjected to performance evaluation. In this study, the jack-knifing principle is used to assess the performance of the flood frequency regionalisation in the complex and data-scarce River Nile basin by examining the error (regionalisation error) between locally and regionally estimated peak flow quantiles for different return periods ( $Q_T$ ). Agglomerative hierarchical clustering based algorithms were used to search for regions with similar hydrological characteristics. Hydrological data employed were from 180 gauged catchments and several physical characteristics in order to regionalise 365 identified catchments. The Generalised Extreme Value (GEV) distribution, selected using  $L$ -moment based approach, was used to construct regional growth curves from which peak flow growth factors could be derived and mapped through interpolation. Inside each region, variations in at-site flood frequency distribution were modelled by regression of the mean annual maximum peak flow (MAF) versus catchment area. The results showed that the performance of the regionalisation is heavily dependent on the historical flow record length and the similarity of the hydrological characteristics inside the regions. The flood frequency regionalisation of the River Nile basin can be improved if sufficient flow data of longer record length of at least 40 yr become available.

## 1 Introduction

Estimation of the peak flow quantiles (usually referred to as design floods) is required in many civil and water engineering applications. Estimation of the peak flow quantiles in ungauged catchments is a challenge often faced by water professionals in many parts of the world mainly due to the absence of peak flow quantiles at ungauged catchments and insufficient record length of stream flow observations at other catchments. One prevalent approach for obtaining such estimates is the regionalisation method (e.g., Das and Cunnane, 2011; Sarhadi and Modarres, 2011; Bernardara et al., 2011; Nezhad et al., 2010; Micevski and Kuczera, 2009; Li et al., 2010; Özçelik and Benzedden, 2010; Ouarda and Shu, 2009; Ellouze and Abida, 2008; Aronica and Candela, 2007; Merz and Blöschl, 2005; Northrop, 2004; Kumar et al., 2003; Cunnane, 1988; Parida et al., 1998; Kjeldsen et al., 2001; Alexander, 1990; Schmidt and Schulze, 1997). Regionalisation applied in flood frequency analysis is the identification (delineation) of groups of catchments with similar hydrological and physical characteristics; hence the search for regions with similar flood frequency distributions. The similarity can be used to estimate (design) peak flows for given return periods at any location in the region. The information on similarity is inferred from the sample of available peak flow data at the gauged sites. It is obvious that the regional flood frequency estimates improve when that sample is larger and/or more representative of the whole population of site

peak flows in the studied region (Chebana and Ouarda, 2009; Lubomír, 2005; Northrop, 2004).

Regional flood frequency analysis using physical catchment characteristics have been shown to produce reliable results if the physical catchment characteristics, chosen objectively, influence the spatial variability of hydrological characteristics (Kjeldsen et al., 2001; Kachroo et al., 2000; Parida et al., 1998; Zrinji and Burn, 1996; Kim and Hawkins, 1993; LeBoutillier and Waylen, 1993; Cunnane, 1988; Wiltshire, 1986; Acreman and Sinclair, 1986; Mosley, 1981). One of the key steps in such analysis is the delineation of the study area into homogeneous regions. Despite increasing research, there is no consensus on a common objective method for delineating homogeneous regions for the purpose of flood frequency estimation.

### 1.1 Clustering

One of the prevalent approaches for delineating homogeneous regions is based on clustering techniques (Clarke, 2011; Guse et al., 2010; Everitt, 2001; Zrinji and Burn, 1996; Aldenderfer and Blashfield, 1984; Roger, 1980) and have been applied in many flood frequency analysis studies (e.g., Clarke, 2011; Ramachandra Rao and Srinivas, 2005; Kachroo et al., 2000; Kim and Hawkins, 1993; Acreman and Sinclair, 1986). However, application of the different clustering techniques to the same dataset, normally leads to different results (Lubomír, 2005). The relative performance of the clustering can be improved by application of an ensemble of clustering algorithms (Ramachandra Rao and Srinivas, 2005). Hierarchical Clustering (HC) is one of the widely used methods in hydrology (LeBoutillier and Waylen, 1993; Mosley, 1981) and consists of four different cluster algorithms. Once the homogeneous regions are delineated they can be tested for homogeneity (Wiltshire, 1986). Several approaches such as the one based on the homogeneity index ( $H_1$ ) (Hosking, 1994; Hosking and Wallis, 1993), the  $S_1$ -statistic based homogeneity test (Alila, 1990) and the graphical test (GT) (Mkhandi et al., 1996) can be used for that purpose. In case of acceptable homogeneity (similarity in the peak flow properties), a unique flood frequency distribution (also called growth curve) can be assumed for each region after scaling of the discharge values by the local specific mean peak flow. Scaled peak flow estimates (also called flood growth factors =  $Q_T/MAF$ ) for each region can be obtained from the growth curve for the desired return periods and can be converted to real life flood magnitudes (at specific site) by the local mean peak flow. Estimation of the local mean peak flow for the ungauged catchment is typically done using a regional regression model derived from the relationship between the mean peak flow and catchment characteristics. The relationship is derived from the data at the gauged stations (Ellouze and Abida, 2008; Merz and Blöschl, 2005; Castellarin et al., 2005; Wagener et al., 2007).

### 1.2 L-Moment

In the development of the regional growth curves, selection and calibration of the probability distribution that adequately fits the scaled peak flow data is required. The available approaches to support such tasks include the maximum likelihood method (e.g., Prescott and Walden, 1983), the regression in quantile plots (Willems et al., 2007; Merz and Blöschl, 2005) and the  $L$ -moment based method (Hosking, 1994, 1990; Hosking et al., 1985), hereafter referred to as  $L$ -moment method. The  $L$ -moment method makes use of the  $L$ -coefficient of variation ( $L$ -CV),  $L$ -skewness ( $L$ -CS) and  $L$ -kurtosis ( $L$ -CK), which are referred to as the second, third and fourth order  $L$ -moment ratios, respectively. The ratios help in selecting appropriate probability distributions and estimating their parameters. Hosking (1990) developed  $L$ -moment ratio diagram based method and a measure based on the  $L$ -CK ( $Z$ -statistic based) for selecting distributions that adequately fit the sample flow data. The distribution parameters are estimated based on  $L$ -moments and it has been noted by Hosking (1990) that  $L$ -moment estimators, compared with other estimators, like the ones in the maximum likelihood method, are reasonably more efficient. The advantages of  $L$ -moment estimators for distribution parameters are documented in Hosking et al. (1985), Hosking and Wallis (1987a, b, 1995) and Hosking (1990). Stedinger et al. (1993) provide models for the estimation of parameters of several distributions in terms of sample  $L$ -moments. More details on this have been discussed by Hosking (1986, 1990) and Vogel and Fennessey (1993).

### 1.3 Regionalisation performance

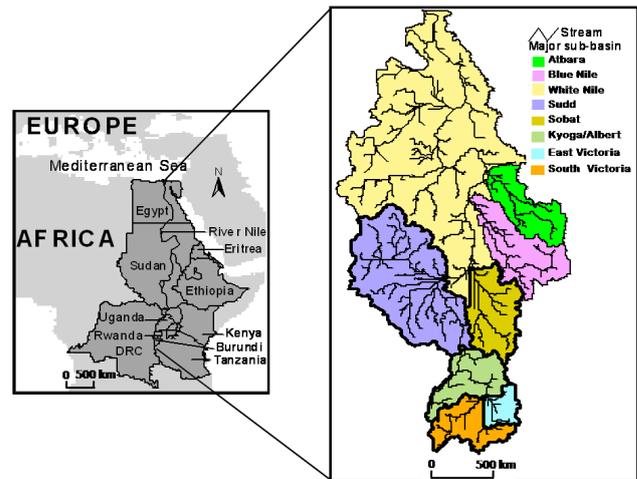
A prevalent approach for the assessment of the regionalisation performance is the jack-knifing technique (e.g., Merz and Blöschl, 2005; Sarhadi and Modarres, 2011). In the jack-knifing approach, a gauged catchment is assumed ungauged and the local (at-site) values of the peak flow quantiles for different return periods are estimated based on the data of the other gauged catchments in the same homogeneous region. The estimates formerly obtained are compared with the same values, but locally estimated and the difference is called, in this paper, regionalisation error and includes the error due to the at-site estimation. This process can be repeated for all the gauged catchments in a homogeneous region and for all the delineated homogeneous regions.

Some regional flood frequency analysis studies have been carried out for the River Nile basin by Kim and Kaluarachchi (2008), Willems et al. (2005), Abdo et al. (2005), but mainly focused on the sub-basin scale (Blue Nile, White Nile, etc.). In this paper, results are shown of a study attempting to regionalise the entire River Nile basin. The study took part in a larger project aimed at enhancing cooperation among the River Nile basin riparian countries in resolving research-based hydrological problems. The feasibility and

performance of the regionalisation of flood frequency distributions were analysed, taking into account the huge basin area, the strong differences in hydrological characteristics across the basin and the limited availability of data. Different HC techniques were applied. Best-fit probability distributions for the regionalisation were selected and calibrated using the *L*-moment approach. Estimates of the mean peak flow were for each delineated region obtained using regression analysis. The performance of the tested regionalisation approach was finally examined using the jack-knifing principle.

## 2 The Nile basin and data

The River Nile basin is situated between 8° S to 33° N and 20° E to 42° E covering an area of approximately 3 762 000 km<sup>2</sup> (Fig. 1). The climate is mainly tropical in the upstream parts of the basin and arid and semi-arid in the downstream parts. The elevation varies from less than 20 to 2150 m a.m.s.l. The mean annual rainfall varies from 1200 mm in the upstream parts to less than 10 mm in the downstream parts. The main rivers in the basin are: Victoria Nile, Albert Nile, White Nile, Blue Nile, Sobat, Atbara, and Main Nile, each with several tributaries (Fig. 1). The daily flow data, from a total of 227 flow gauging stations, and from which the annual maximum flow (AMF) data were derived, were obtained mainly from the River Nile basin Flow Regimes from International, Experimental and Network Data (FRIEND/Nile) project. The AMF data were analysed for data errors and trends and the data found with anomalies or trends screened out or detrended, respectively, using the respective methods described in Khaled (2008), Kundzewicz and Robson (2000) and Hosking and Wallis (1993). Data screening resulted in flow series of 180 gauged catchments with record lengths ranging from 4 to 116 yr. For these 180 stations, the AMF data were detrended for 12 stations. The detrending was done because changes in uptake of water, catchment land-use or river morphology might have caused the trends. These influences have to be removed given that regionalisation in flood frequency analysis mainly deals with regional differences in river flow levels and variability due to natural catchment runoff processes. Human influencing factors, however, change over time and should be dealt with separately. For each dataset, the mean of the AMF data (MAF) and the respective statistical properties of the coefficient of variation (CV), coefficient of skewness (CS), coefficient of kurtosis (CK), and the *L*-coefficients, were estimated. The mean annual rainfall (MAR) estimates were obtained from observed precipitation data for a total of 584 rainfall gauging stations with record length ranging from 5 to 99 yr. The MAR estimates were spatially interpolated for each sub-basin based on Thiessen polygon method (Linsley et al., 1949). The influence of the maximum annual rainfall on characterizing similar flood homogeneous region



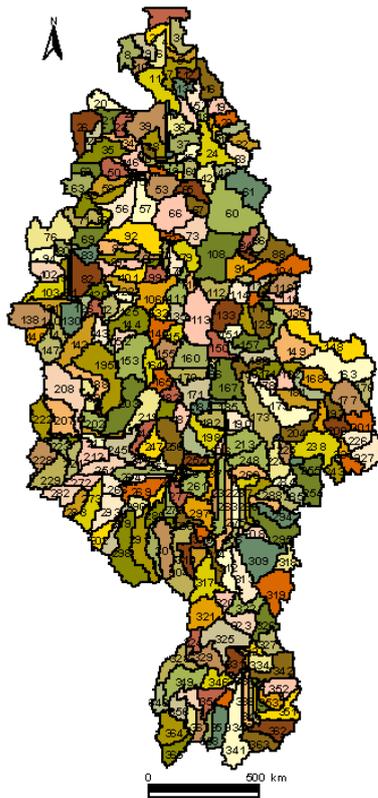
**Fig. 1.** Location of the River Nile in Africa, the countries in which the River Nile basin takes part, the major streams and major catchments of the River Nile.

was compared with that of MAR and the findings showed negligible difference. To support the digital delineation of catchments, a Digital Elevation Model (DEM) for the study area was built based on the 92 m grid resolution topographical data obtained from the Shuttle Radar Topography Mission website (<http://srtm.csi.cgiar.org/>; last access: September 2011). The watershed delineation was carried out using the AVSWATX extension (Di Luzio et al., 2001) for ArcView GIS (Geographical Information System) and resulted in an automatic 365 subbasins (catchments) (Fig. 2) the result of which was validated against previous studies. For each delineated catchment (Fig. 2), several physical catchment characteristics were extracted from the DEM data after watershed delineation and some were objectively (e.g., use of correlation coefficient) selected for cluster and regression analyses (Table 1).

## 3 Methodology and results

### 3.1 Cluster analysis

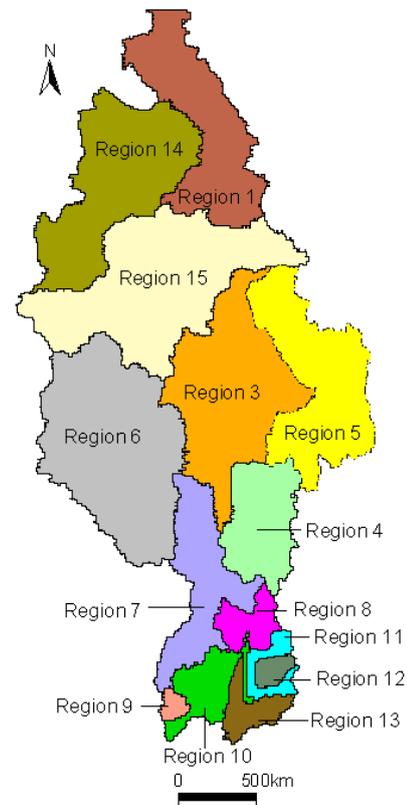
The data matrix consisting of the physical and hydrological catchment characteristics used in the clustering are indicated in Table 1. Omission of landuse and soil types in the clustering was established to be insignificant for this study. Each characteristic was standardised by its corresponding mean value before use for clustering to give each characteristic equal weight in the clustering (Kachroo et al., 2000). A two case approach was adopted in the clustering. In the first case, only the physical characteristics of the 365 automatic generated catchments (Fig. 2) were used to define similarity among them using four HC algorithms (William and Edelsbrunner, 1984). This was done to ensure that each of the ungauged and gauged catchments has equal representation in the clustering



**Fig. 2.** Sub-catchments of the River Nile basin delineated using ArcView GIS from which several catchment characteristics were extracted and used for cluster analysis.

process and to reflect the stability of the clusters. In the second case, both the physical and hydrological characteristics were used but only for the gauged catchments. In both cases, an initial optimal number of 30 sub-clusters were defined as a criterion for stopping the algorithms, starting from an initial condition where the whole basin is defined as a single homogeneous region. The outcomes, from the two cases were evaluated and compared until optimal clusters were achieved. Dendrograms for each of the four HC algorithms were derived and used to aid in the identification of groups of catchments with similar properties and in the assessment of the performance of the four HC algorithms.

Figure 3 shows the result of the delineation of the catchments into possible homogeneous regions. The delineation results for the river network are similar to the one developed for Africa by Sutcliffe and Parks (1999) and Karyabwite (2000). In the first case, where only the catchment physical characteristics were used for clustering, 30 regions were obtained (dendrograms not shown); while in the second case, where both the physical catchment characteristics and hydrological properties were used, 15 regions were delineated. In the second case, four regions (2 and 5; 1 and 14) obtained in the first case, were delineated into only two separate regions. This also indicated the level of stability of the identified



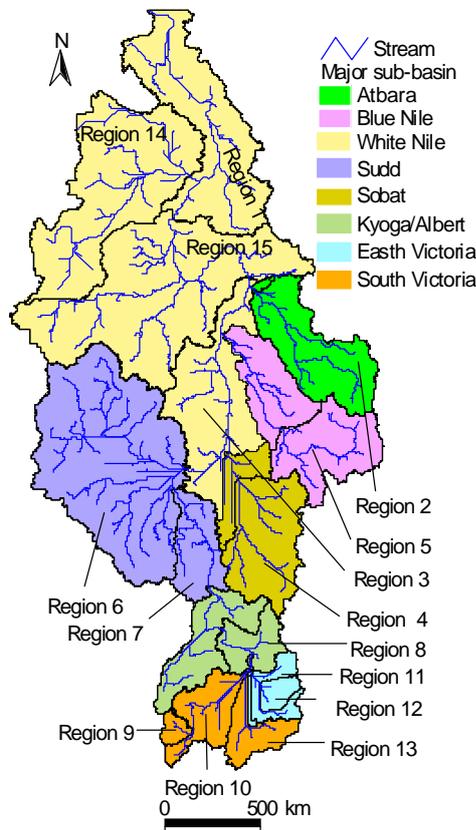
**Fig. 3.** Delineated regions of the River Nile basin showing different homogeneous regions and their spatial distribution.

clusters. Further analysis was made by comparing weighted regional values of the physical characteristics of regions 1 and 14, and regions 2 and 5. Regions 2 and 5 were found to be similar and, hence, merged; regions 1 and 14 were found different and kept separately. Region 14 was short of flow data having record length greater than 5 yr. Further inter-regional comparisons were made using only the weighted physical properties and it was found that region 14 is physically similar to region 3. Delineated regions (Fig. 3) approximately match and overlap over the major catchments (Fig. 4). This observation suggests that the results of the clustering are highly influenced by the proximity among the catchments and geographical characteristics although one catchment may belong to a group of similar catchments, which are not geographically connected. If the major catchments (Fig. 4) are assumed homogeneous, it approximates the results of the clustering. However, the homogeneous regions are not convincing because regions 1, 3 and 15 are highly influenced by regions 4, 5, 6 and other regions upstream. This causes the dependency of the flow data in regions 1, 3 and 15. If we take into account the total size of the River Nile basin ( $\sim 3\,762\,000\text{ km}^2$ ) and the number of delineated regions, it would mean that the average area of each homogeneous region is about  $268\,714\text{ km}^2$ . Using expert judgment, the average size of each region is, thus, too big to be considered in

**Table 1.** Physical and flow catchment characteristics considered paramount in influencing the magnitude of peak flows.

Vel <sup>a</sup>	Velocity [m s <sup>-1</sup> ]	MaxE	Subbasin maximum elevation [m]
MAR	Annual Areal rainfall [mm]	MeanE	Subbasin mean elevation [m]
MAF	Mean annaul flood [m <sup>3</sup> s <sup>-1</sup> ]	PointE	Elevation at flow station point [m]
CV	Coefficient of variation	Area	Watershed area [km <sup>2</sup> ]
L-CV	L-coefficient of variation	RhL <sup>c</sup>	Stream reach [m]
L-CS	L-coefficient of skewness	RhS	Stream reach slope [-]
L-CK	L-coefficient of kurtosis	RhW	Stream reach width [m]
Sol1	Basin slope	RhH <sub>min</sub>	Minimum elevation of the stream reach [m]
Len1 <sup>b</sup>	Stream reach length [m]	RhH <sub>max</sub>	Maximum elevation of the stream reach [m]

<sup>a</sup> We refer to MAF/catchment area as velocity. <sup>b</sup> Longest path within the catchment length. <sup>c</sup> Longest path within the catchment width.



**Fig. 4.** Map showing overlay of delineated regions onto major catchments, and DEM based delineated river system of the Nile basin.

anyway homogeneous, despite the fact that the average number of gauged catchments per region is about 7, which from the regionalisation point of view may be reasonable.

### 3.2 The L-moment method

#### 3.2.1 L-moments and the L-moment ratios

The first four sample L-moments,  $l_1$  to  $l_4$  (Hosking and Wallis, 1997; Greenwood et al., 1979) of the AMF were used

to obtain the L-moment ratios ( $L-CV$ ,  $L-CS$  and  $L-CK$ ) as follows:  $t_2 = l_2/l_1 = L-CV$ ,  $t_r = l_r/l_2$  ( $t_3 = L-CS$ ;  $t_4 = L-CK$ ) for  $r = 3, 4$ , respectively, independent of the flow units; where  $l_1 = b_0$ ,  $l_2 = 2b_1 - b_0$ ,  $l_3 = 6b_2 - 6b_1 + b_0$ ,  $l_4 = 20b_3 - 30b_2 + 12b_1 - b_0$  and

$$b_0 = \frac{1}{n} \sum_{j=1}^n X_j \tag{1}$$

$$b_r = \frac{1}{n} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} X_j \tag{2}$$

are unbiased sample estimators ( $b_r$ ) of Probability Weighted Moments (PWMs) and  $X_j$  ( $j = 1, 2, \dots, n$ ) is the ordered set of AMF values ( $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$ ). The L-moments represent the location, dispersion (scale) and shape of the data sample similar to the conventional moments. We used the L-moments and the L-moment ratios to (1) carry out regional heterogeneity tests, (2) select candidate probability distributions and best-fit probability distribution for the dataset, and (3) to calibrate the parameters of the candidate probability distributions.

#### 3.2.2 Homogeneity test

The methods based on  $H_1$ ,  $S_1$ -statistic and GT were used for testing the homogeneity of the delineated homogeneous regions. The details of each of the test methods can be found in the respective publication. A region is considered “acceptably homogeneous” (A) if  $H_1 < 1$ ; “probably homogeneous” (P) if  $1 \leq H_1 < 2$ ; and “definitely heterogeneous” (H) if  $H_1 \geq 2$ . The GT assumes that a group of catchments form a homogeneous region if the L-CV values are similar to those obtained from synthetic generated data of the assumed parent probability distribution (Mkhandi et al., 1996).

Table 2 shows the homogeneity test results for the three test methods. The first and the second columns contained the region IDs and the average record length. The number of gauged catchments in the region is presented in the third column and the implications of the  $H_1$ ,  $S_1$  and GT test results are contained in columns 5, 7 and 8, respectively. Columns 10

**Table 2.** The average record length, the number of gauged catchments, regional homogeneity test results, catchment characteristics used in the regression for the MAF, the NME in estimating the MAF and the NMSE of the CV of the AMF per region.

Region	Year	No	$H_1$	HR	$S_1$	SR	GTR	SBCH	NME	NMSE
1	21	9	-1.69	P	-2.74	P	A	MeanE	-0.023	0.000
3	48	22	-0.26	A	-0.06	A	P	Area	0.522	0.013
4	33	14	0.83	A	-0.05	A	H	MAR	0.796	0.017
5	32	15	-0.12	A	-0.06	A	A	MeanE	0.017	0.034
6	19	13	0.21	A	0.03	A	A	Area	0.382	0.059
7	36	15	0.21	A	0.04	A	A	Area	-0.062	0.032
8	39	15	-0.28	A	-0.30	A	A	Len1	0.013	0.031
9	32	7	-0.27	A	-0.18	A	A	Len1	-0.101	0.006
10	16	20	-0.23	A	-0.16	A	P	Area	2.500	0.055
11	40	18	-2.09	H	-0.60	A	A	Area	0.053	0.051
12	32	10	-0.50	A	-0.19	A	P	MAR	3.174	0.094
13	28	11	-0.32	A	-0.11	A	P	MAR	3.585	0.007
15	27	12	-1.190	P	-14.09	H	A	Len1	0.028	0.003

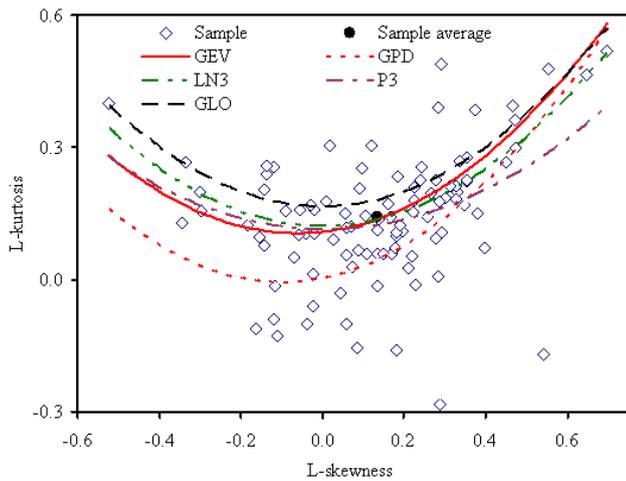
Year: average record length; No: number of gauged catchments in the region;  $H_1$ : homogeneity index; HR: interpretation of  $H_1$  results;  $S_1$ :  $S_1$ -statistic; SR: interpretation of  $S_1$ -statistic; GTR: interpretation of GT results; A: acceptably homogeneous; P: possibly homogeneous; H: definitely heterogeneous; SBCH: catchment characteristic with stronger relationship with MAF; NME: normalised mean error in MAF estimate. NMSE: normalised mean squared error in CV.

and 11 contain the values of normalised mean error (NME) in MAF estimate and the normalised mean squared error in CV (NMSE), respectively. It can be seen from Table 2 that the homogeneity test results significantly differ for regions 4, 11 and 15, and it is not clear why such differences. The GT method differs with the  $H_1$  and  $S_1$  based methods over eight regions.  $H_1$  differs with GT and  $S_1$  over region 11, which has the highest  $L$ -coefficient of variation. It is, however, observed that there is an advantage of GT over the two other methods because it allows identification of the outlier catchments that may not be part of the homogeneous region. For a region to be hydrologically homogeneous, the value of CV of the AMF should insignificantly vary from their mean even if the values of the corresponding catchment physical characteristics vary significantly. The NMSE error in CV values should, therefore, be zero for a region with perfect homogeneity. The NMSE (Table 2) indicates that regions 1, 9, 13 and 15 are hydrologically homogeneous though the test results indicate that regions 1, 13 and 15 are probably homogeneous, and region 15 is actually heterogeneous. The NME values in estimating the MAF for regions 3 and 4 are indications of heterogeneity of the regions. For regions 6 and 10 the higher values of NME is probably due to the higher percentage of gauged catchments with flow data of shorter record lengths. This observation cannot well be substantiated because regions 1 and 15, which consist of higher number of catchments with flow data of shorter record length, have lower values of NME. For regions 12 and 13, the higher values of NME is probably due to the lack of stronger correlations between the catchment characteristics and the MAF. Our general observation, on the analysis of the homogeneity results for this study, is that, for a large and complex River Nile basin, delineating regions which are both hydrologically and physically

homogeneous may not be possible unless more stations with longer records become available. Furthermore, if the number of gauged catchments in each region is not optimally representing the entire region (e.g., skewed to one side of the region), as may be for the case of regions 4 and 9 (not spatially shown), establishing whether or not a region is both hydrologically and physically homogeneous may also be difficult, even if the size of the regions are reasonably smaller. Defining extremely large region as a result of lack of data may not render the regionalisation as a valid option in estimating the AMF for an ungauged catchment, especially for practical applications. For the River Nile basin, achieving a compromise between physical and hydrological homogeneity; as well as between optimal data and regional homogeneity is indeed a difficult task. However, as noted by Cunnane (1988), a small departure (not definite how small) from the homogeneity range does not negate the benefit of regionalisation. Indeed, especially for the huge spatial scale of the River Nile basin, and although we cannot consider the regions as “homogeneous”, they still might be helpful in the regionalisation analysis.

### 3.2.3 Selection of the candidate probability distributions

The selection of the candidate parent distributions was based on the  $L$ -moment ratio diagram and the  $Z$ -statistic (Hosking and Wallis, 1993). Details of the computation of the  $Z$ -statistic are explained by Hosking and Wallis (1993). All distributions whose absolute  $Z$ -statistic values are less than 1.64 qualify for candidate probability distributions and the distribution with the lowest absolute  $Z$ -statistic value qualifies for the best-fit distribution. Five 3-parameter distributions,



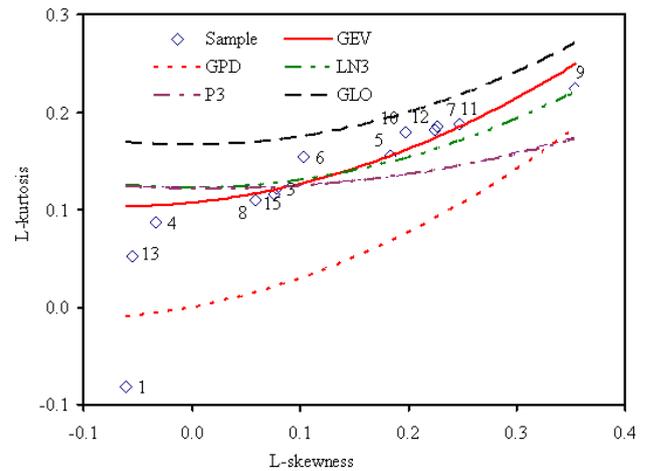
**Fig. 5.** *L*-moment ratio diagram for the at-site AMF data in the River Nile basin.

GEV: Generalised Extreme Value distribution; GLO: Generalised Logistic distribution; LN3: 3-parameter Log Normal distribution; GPD: Generalised Pareto Distribution; P3: Pearson type 3 distribution.

recommended by Yue and Wang (2004) for the study of extreme events, were considered. These are Generalised Extreme Value distribution (GEV), Generalised Pareto Distribution (GPD), Generalised logistic distribution (GLO), Log-normal 3-parameter distribution (LN3) and Pearson type 3 distribution (P3). Regional and at-site calibration of the distribution parameters were based on the *L*-moment parameter estimators (Hosking and Wallis, 1997, 1987). The *L*-moment ratio diagram constructed using the AMF was applied in the selection of the candidate distributions (Fig. 5).

### 3.2.4 Selection of the best-fit distribution and construction of regional growth curves

Given that a single growth curve was envisaged for each region, the same type of the distribution was selected per region. The selection of the regional distribution was based on the regional *L*-moments ratio diagram and the *Z*-statistics (Hosking and Wallis, 1993). Regional *L*-moments were obtained based on weighted sample points. The best-fit probability distribution was selected based on the observation of the regional sample point. This was done by identifying which selected distribution plots on or near the regional sample point (Fig. 6). The GEV came out as the best regional distribution for most regions. The scaled regional sample *L*-moments (*L*-CV, *L*-CS and *L*-CK) were used for the regional parameter estimation. The regional parameters of the GEV distribution are provided in Table 3. The growth curve model of the GEV, in function of the return period, *T*, is given in Eq. (3).



**Fig. 6.** Regional *L*-moment ratio diagram of the AMF for the River Nile basin.

GEV: Generalised Extreme Value distribution; GLO: Generalised Logistic distribution; LN3: 3-parameter Log Normal distribution; GPD: Generalised Pareto distribution; P3: Pearson type 3 distribution.

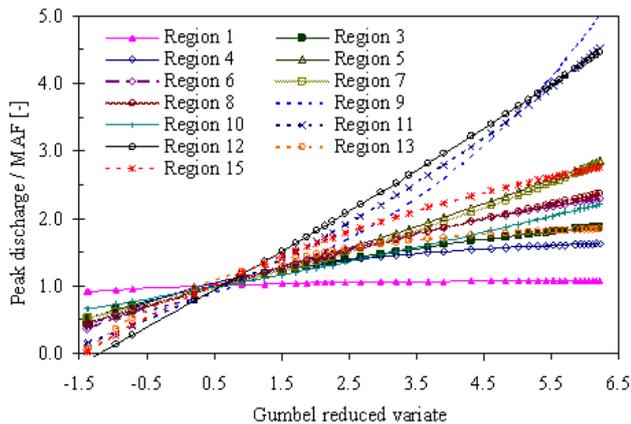
$$Gf_T = \xi + \left(\frac{\alpha}{k}\right) \left[1 - \left\{1 - \ln\left(1 - \frac{1}{T}\right)\right\}^k\right] \quad (3)$$

The  $\xi$ ,  $\alpha$ , and  $k$  are the distribution location, scale and the shape parameters, respectively. The regional curve was plotted versus the extreme value type one (EV1) or the Gumbel reduced variate, given by  $[-\ln(-\ln(1 - 1/T))]$ . If applicable in practical application, it is possible to estimate a peak flow quantile for a given return period by the use of the growth curve model given in Eq. (3). The selection of the GEV distribution is consistent with the conclusions by Willems et al. (2005) based on AMF data from 56 gauging sites in the River Nile basin. The shapes of the growth curves (Fig. 7), for most regions, indicate that the slope of the growth curve generally becomes larger with increase in return periods except for regions 1, 3, 4, 13 and 15. For regions 9, 11 and 12, the increase in the slope is very strong as the return period increases. In contrast, strong decreasing slopes are in regions 13 and 15. The shape of the growth curve for region 1 can be explained by its downstream location in the basin, which has a very gentle topographical slope; it is in the arid region and flow peaks are attenuated before reaching region 1. The identified physical evidence to explain the behaviour of AMF for region 4 is that the topographical slope is very gentle. Figure 7 and Table 3 indicate that the distribution's shape varies spatially, but for most regions the values are close to zero, indicating normal tail behaviour of the growth curve, except for regions 1, 4, 13 and 9.

**Table 3.** The average record length, the number of gauged catchments, regional homogeneity test results, catchment characteristics used in the regression for the MAF, the NME in estimating the MAF and the NMSE of the CV of the AMF per region.

Region	MAF	CV	$L$ -CV	$L$ -CS	$L$ -CK	$\zeta$	$\alpha$	$k$
1	2101	0.038	0.022	-0.062	0.052	0.989	0.041	0.410
3	4541	0.247	0.136	0.078	0.155	0.895	0.249	0.185
4	485	0.321	0.182	-0.033	0.087	0.940	0.299	0.554
5	4640	0.454	0.250	0.183	0.116	0.836	0.283	-0.019
6	357	0.491	0.255	0.103	0.181	0.861	0.346	0.212
7	1007	0.405	0.222	0.226	0.154	0.842	0.240	-0.083
8	428	0.499	0.298	0.059	0.185	0.931	0.498	0.681
9	242	0.449	0.253	0.354	0.224	0.755	0.267	-0.268
10	246	0.396	0.214	0.197	0.179	0.886	0.167	-0.099
11	102	0.664	0.345	0.247	0.188	0.686	0.373	-0.218
12	212	0.633	0.307	0.224	0.109	0.663	0.402	-0.217
13	132	0.427	0.244	-0.055	0.122	0.866	0.430	0.347
15	118	0.506	0.302	0.076	-0.081	0.760	0.458	0.060

MAF: regional MAF in  $\text{m}^3 \text{s}^{-1}$ ;  $\xi$ ,  $\alpha$  and  $k$  are the location, scale and shape parameters, respectively; all the regional moments and the distribution parameters shown have been scaled and are unitless.



**Fig. 7.** Regional growth curves developed using GEV distribution for the homogeneous regions in the River Nile basin.

### 3.3 Estimation of MAF using regression model

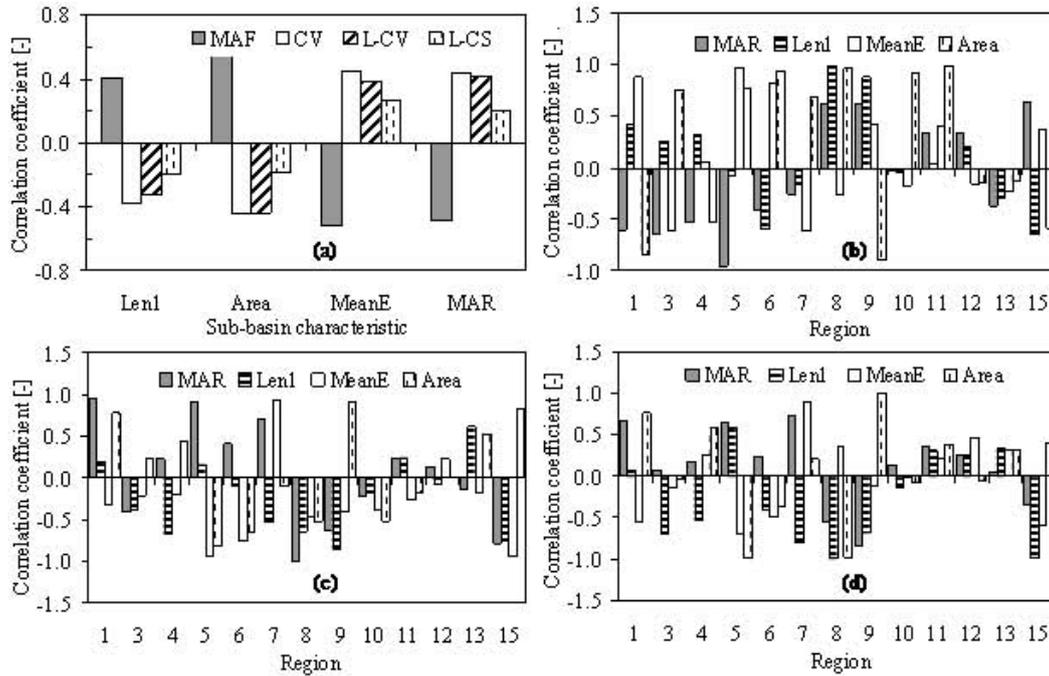
The growth curve provides estimates of scaled AMF quantiles,  $Q_T/\text{MAF}$ , for given recurrence intervals or return periods,  $T$ , for the different regions. Regional regression models were developed for estimation of the MAF in each homogeneous region. In the first step, we assessed the correlation of each of the available physical catchment characteristics with the at-site MAF values by measure of correlation coefficients, for the entire basin. The catchment characteristics with higher correlation coefficients were selected. In the second step, we repeated the correlation analysis per region, but assumed one gauged catchment as an ungauged catchment. We called this catchment the local catchment, and left out the value of MAF of that local catchment. We then developed a regression model for estimation of MAF as a function of the catchment characteristics having the highest

correlation coefficient with the MAF for that region in the form of Eq. (4).

$$\text{MAF} = aC_1^b C_2^c \dots \quad (4)$$

where,  $C_1, C_2, \dots$ , are the various catchment characteristics considered. The physical/geographical characteristics used in this study such as elevation, slopes and river morphology were thought to be of primary influence and were, therefore, considered in this study as compared to other physical characteristics such as climatic indices. We then used the regression model to estimate the value of the MAF for the local catchment.

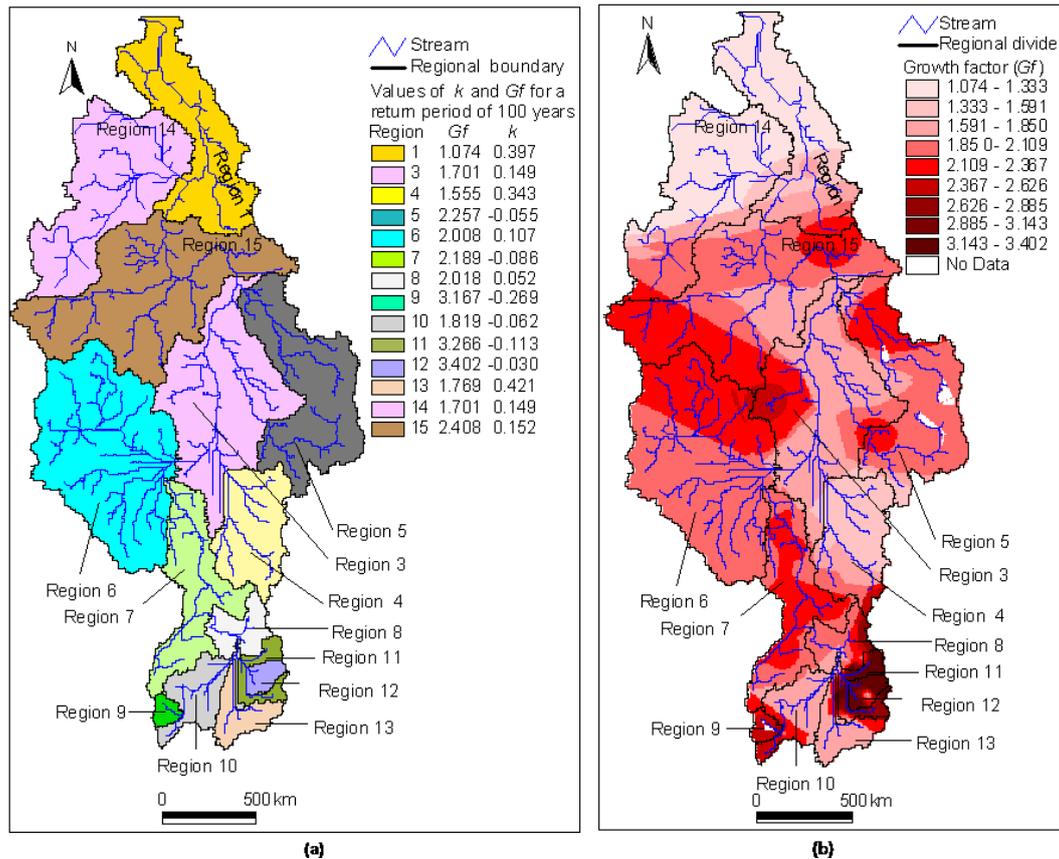
Plots of the correlation coefficient versus the Len1, Area, MeanE and MAR, for the entire data are shown in Fig. 8a. The values of the correlation coefficient vary significantly with these catchment characteristics; indicating that the behaviour of the MAF and also the AMF properties, is controlled differently by the different catchment characteristics. The slopes of the relationships between the catchment characteristics and the moments/ $L$ -moment ratios decrease with the order of the moments/ $L$ -moment ratios. The slopes are steeper for the ordinary moments than for the  $L$ -moments ratios. It is also observed that when the AMF moments are plotted versus the different catchment characteristics, the scatter of the data points around the mean value (or the mean squared error) is less for the higher moments than for the lower moments (not shown). Similarly, when the  $L$ -moment ratios are plotted versus the different catchment characteristics, and compared with the similar plots for the ordinary moments, the respective scatter of the data points (or the mean squared error) is higher for the ordinary moments than for the  $L$ -moment ratios (not shown). Higher slopes, reflecting higher values of correlation coefficient, means that the catchment characteristics are better estimators of ordinary



**Fig. 8.** Correlation coefficients between catchment characteristics and flood properties: (a) for the entire dataset of the basin (b) MAF values per region (c) CV values per region (d) L-CS values per region. MeanE: catchment mean elevation (H); Len1: catchment average stream reach length (L).

moments than estimators of *L*-moments or their ratios. The values of the correlation coefficient between the ordinary moments and the catchment characteristics (Fig. 8b, c), and between the *L*-moments or their ratios (e.g., Fig. 8d for *L*-CS) and the catchment characteristics, per region, vary significantly with the selected catchment characteristics. The catchment characteristic which highly influences flow property in one region is not necessarily the same in another region as shown in Fig. 8b–d and this is probably due to climatic differences (wet and dry) between the regions. Identification of the most influential physical catchment characteristics on flow properties was based on the use of correlation coefficient. The catchment characteristic with the highest value of correlation coefficient was related in a power law to the MAF to derive a regression model for estimation of the local MAF. A possible advantage of using a simple regression model, as compared to the multiple regression model, is that it eliminates the catchment characteristics which are not strongly related with the MAF; and which can significantly influence the accuracy of MAF estimates. The homogeneity test result may indicate departure from the homogeneity range (Table 2), but it is still possible to identify a good regional estimator for MAF as indicated by the values of the NME and NMSE for regions 1, 11 and 15 (Table 2). In regions where it is difficult to identify a reasonable estimator for MAF, such as for regions 12 and 13 in this study (Table 2), we included additional catchment characteristics in

the analysis. The estimated MAF was evaluated by comparing with the local MAF. The estimated MAF together with the growth curve model were used to estimate the AMF values for the return periods ranging from 2.3 to 500 yr. They were compared with the quantiles obtained from the distribution of the local catchment estimated during at-site calibration for the same return periods. This process was repeated, in turn, for each gauged station in the region and for all the delineated regions. This procedure allowed us to obtain an estimate of the regionalisation error. It is expected that the performance of the regional estimators diminishes when extremely large regions are delineated on the account of the increasing variance of the parameter estimates. When evaluating this variance, it is important to take into account the huge spatial scale of the Nile basin considered, as well as the strong climatic variations across the basin and the data limitations. For the same reasons, flood frequency analysis may not be as reliable as it is for other basins. The acceptability of the accuracy of the regional flood frequency analysis should be seen in light of this scale context. Despite the huge catchment area and the low density of the stream gauge network, water and civil engineering works require flood frequency estimates to be made at ungauged locations through regionalisation, with highest possible accuracy.



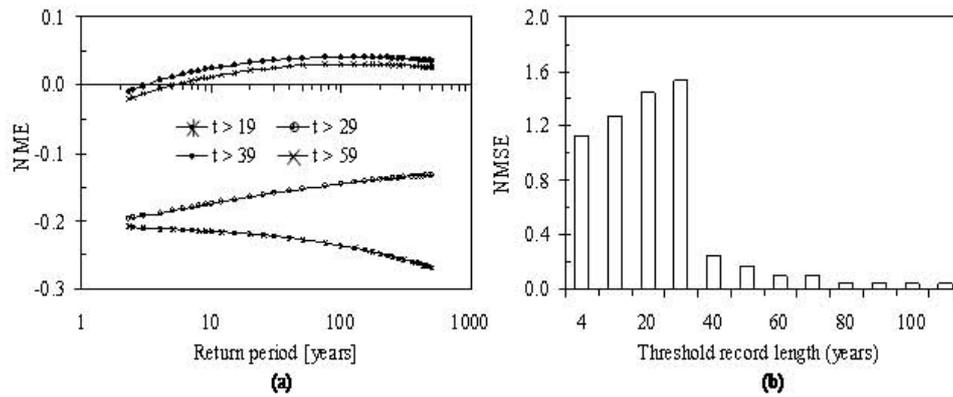
**Fig. 9.** Variation in regional growth factor for 100 yr return period ( $G_{f100}$ ) simulated using GEV distribution, and the regional shape parameter ( $k$ ) in the River Nile basin: (a) constant value of ( $G_{f100}$ ) in a region (b) value of  $G_{f100}$  varies over a region.

### 3.4 Mapping and comparison of local and regional growth factors

Two flood growth factor maps (Fig. 9) were produced based on regional growth factors for a 100 yr return period. The regional GEV parameters (Table 3) were used to calculate the value of the growth factor corresponding to a return period of 100 yr ( $G_{f100}$ ). The first map (Fig. 9a) was produced by representing each region with a constant value of  $G_{f100}$ . The second map (Fig. 9b) was produced by interpolating the value of the  $G_{f100}$  to produce a continuous map based on ordinary Kriging method. The two maps were compared in the context of the spatial variation of the value of the  $G_{f100}$  and the suitability for practical engineering application. The legend of Fig. 9a consists of both the regional values of  $G_{f100}$  and the shape parameter and Fig. 9b shows overlays of delineated regions and the legend consists of ranges of the values of  $G_{f100}$ . Considering the fact that it is very elusive to delineate a homogeneous region for the River Nile basin, because of the varying degree of the catchment characteristics among catchments within a given region, the map produced by interpolating regional growth factors may be more appropriate in practice. Instead of using  $G_{f100}$  value which is constant

over a region, Fig. 9b would take into consideration variability of  $G_{f100}$  over and across a given region. For most regions, the peak flows are expected to at least double their respective MAF values; meanwhile for region 1, the expected peak flow, for the same return period, will be close to the regional MAF value. Higher values are for the upper River Nile region (except for region 13), the Sudd and Sobat major catchments, and the Atbara catchments. This is explained by the higher frequency of heavy rain storms in the regions.

The map shown in Fig. 9b can be seen as a “regional design map” giving regional growth factor values for the specified design return period and a given river (stream) location, based on the regional growth curves. They can be transferred to design flow values ( $Q_T$ ) by multiplying by the MAF estimate at that location. The latter requires regional MAF estimates, which have been discussed in previous sections. Practical application of this is found in the design of culverts, bridges for roads, rail communications, hydraulic structures for irrigation, reservoir spillways, as well as for flood risk assessment and flood management in the River basin. The shapes of the growth curves (Fig. 7) are largely affected by the values of  $k$  of the distribution. The regional values of  $k$  are in the range  $-0.269$  to  $0.421$ . The higher values of the



**Fig. 10.** NME (a) in estimating AMF for the regionalisation, based on *L*-moment method for the threshold record length (*t*) ranges: asterisks (*t* > 19: *t* = 20 to 116); circle (*t* > 29: *t* = 30 to 116); full circle (*t* > 39: *t* = 40 to 116); crosses (*t* > 59: *t* = 60 to 116). NMSE (b) in estimating the MAF using the jack-knifing principle for record length ranging from 4 to 116 yr. GEV distribution and *L*-moments (and their ratios) were used.

growth factors are related to lower values of *k* and are found in regions 5, 6, 7, 8, 9, 10, 11, 12 and 15. Higher values of growth factors and negative/lower values of *k* mean that there is high variability of AMF in the region; they correspond to heavy tail behaviour of the extreme value distribution. Values of *k* greater than zero correspond to a light tail behaviour, which means that extremes do not rise strongly with increasing return period. The latter might be due to the flooding influences, which bend down the tail of the distribution, as observed for the cases of regions 13 and 15. The trends in Fig. 9a are clearly affected by tropical humid and subtropical arid or semi-arid conditions. The values of *k* are generally lower (in most cases negative) and growth factors higher for the tropical humid area in the upstream parts of the River Nile basin around Albert, Kyoga and Victoria Lakes.

### 3.5 Regionalisation error

Merz and Blöschl (2005) indicated that for any modelling application, the regionalisation error consists of a systematic component, or bias, and a random error component. They further state that the bias is a measure of whether a regionalisation method tends to overestimate or underestimate flood quantiles in all the catchments considered. Non-negligible bias is an indication of poor model structure or inappropriate assumptions. In practical applications, biases, if known, can be removed from the estimates using a bias removal technique. The random error is a measure of the scatter of the regionalised values centred at the local values. Random errors are related to how much information a method can extract from the data and can be removed from the estimates. We use the Normalised Mean Error (NME) as a measure of the bias between regional and local estimates (Eq. 5a).

$$NME = \frac{\sum_{i=1}^n (Q_i^s - Q_i^o)}{\sum_{i=1}^n Q_i^o} \quad (5a)$$

where  $Q_i^s$  and  $Q_i^o$ , respectively, are the regionalised and the local values of the estimates of station *i* out of *n* stations. The Normalised Standard Deviation Error (NSDE) was used as a measure of the random error (Eq. 5b).

$$NSDE = \frac{\sqrt{\sum_{i=1}^n \left\{ (Q_i^s - Q_i^o) - n^{-1} \sum_{i=1}^n (Q_i^s - Q_i^o) \right\}^2}}{\sum_{i=1}^n Q_i^o} \quad (5b)$$

NME can be positive or negative, while NSDE is always non-negative. The performance of the regionalisation is considered perfect when both the values of NME and NSDE are zero.

The regionalisation error is a function of both the NME and NSDE, and we used the Root Mean Square Error (RMSE), given in Eq. (6), as a measure of the total regionalisation error. It was expected that the most homogeneous regions and the catchments with longer record lengths would produce the lowest error values.

$$RMSE = \left[ NME^2 + NSDE^2 \right]^{0.5} \quad (6)$$

We used the plot of the NME versus the return period, and the NMSE versus the record length to assess the variation of the error with the return period ranging from 2.3 to 500 yr and to explicitly reveal the effect of the record length in estimating the AMF. This was done by screening out data with record

lengths less than 20, 30, 40 and 50 yr from the error analysis one at a time. We found that for shorter return periods the random error will be smaller and represents the regionalisation error alone, while for longer return periods both the bias and the random error are likely to be important. Figure 10 shows the general performance of the regionalisation by measure of biases in the values of the AMF and the local MAF. It also illustrates how the record length affects the results of the regionalisation. Figure 10a shows the plot of NME versus return period with increasing threshold record length,  $t$ , from 20 to 60 yr. Figure 10b shows the bar chart of NMSE versus the threshold record length. The effect of record length reduces as higher thresholds are selected. For the value of the threshold record length of 20 ( $t > 19$ ) years (Fig. 10a), the values of the NME increase significantly with increase in return period. The values of NME reduce when the value of the threshold record length of 30 yr is selected. In this case, the NME values reduce with increase in return period. The difference in the values of the NME becomes smaller when the threshold records of 40 yr or more are selected (Fig. 10a:  $t > 39$  and  $t > 59$ ) and no longer change significantly. The NMSE is also affected by the values of the record length of the data used for the regionalisation. Most of the catchments used have data with record length ranging from 20 to 70 yr. Only one catchment has a record length longer than 100 yr. It is clear from Fig. 10b that catchments with record lengths less than 40 yr may not be good for regionalisation, because of the high error they introduce in the regionalisation, although a reasonable amount of regional information can still be extracted from such data. If such data are used, a correction factor may be established and applied to the resulting MAF in case of practical application provided the interest is in the estimate of the AMF frequency estimation for return period longer than the record length observed. The correction factor can be thought of as a ratio between the estimated and the true value of the MAF for the considered catchment and can be obtained by analysing the values of the NMSE. The factor is close to one as the average record length of the regional data used increases. The peak flow quantiles are, thus, often underestimated if shorter record lengths are used. As the average record length of the regional data increases, the estimated MAF is close to the true value, provided the region under consideration is both hydrologically and physically homogeneous. This compromise is very difficult to achieve in regionalisation, especially for the complex River Nile basin where flood data are limited.

#### 4 Conclusions

In this study, we used agglomerative hierarchical cluster algorithms to search for homogeneous regions in the complex River Nile basin and regionalised 365 identified catchments into groups of “homogeneous” regions. 180 flow data were used; about 40 % of which have flow record length greater

than 30 yr (regions with similar characteristics used for the regional flood frequency analysis). Several catchment physical characteristics were digitally extracted and used in the clustering process and the regression modelling for estimation of the MAF. Using the  $L$ -moment based method; the GEV distribution was selected as the overall best-fit distribution for the data and was used to construct regional growth curves for the estimation of the peak flow quantiles for selected return periods, for all the regions. The performance of the regionalisation was examined by analysing homogeneity test results and the error between locally and regionally estimated AMF values for the different return periods using the jack-knifing principle.

The hierarchical clustering algorithms, applied in this study, were reasonably efficient in the identification of catchments with similar hydrological and physical characteristics, but are not objective enough in establishing the optimal number of clusters. The three different hydrological homogeneity test methods applied in the study lead to different homogeneity test results for a number of regions signifying complexity in the flood frequency regionalisation. The selection of the GEV as the best-fit distribution is consistent with the conclusions by Willems et al. (2005) based on flow data from 56 gauging sites in the River Nile basin. Catchment physical characteristics were found to be better estimators of lower moments/ $L$ -moments than higher moments/ $L$ -moments. The performance of the regionalisation is strongly dependent on the record length of the AMF data used and the physical catchment characteristics used in the regression technique.

The performance of the regionalisation, however, can be improved if the flow data used have record length longer than 40 yr provided the catchments are considered to fall in a homogeneous region. The compromise between availability of flow data with longer record length and delineating homogeneous region is very difficult to achieve in regionalisation, especially for the complex and highly ungauged River Nile basin where flow data are limited, both in availability and record length and where the types of climate vary from humid in the upstream to arid in the downstream. Such limitations will continue to constrain, and eventually affect the applicability of the results of the regionalisation of the River Nile catchments at basin scale. However, if sufficient data would become available, the effect of record length and regional size may be eliminated. In order to make better conclusion on the physical homogeneity of the delineated regions, a homogeneity test method that incorporates the values of the physical catchment characteristics may be required. Nevertheless, the reliability of the regionalisation results would still have to be examined in the context of both physical and hydrological homogeneity. Physical homogeneity may be difficult to establish because of the several physical catchment characteristics (whose values may be altered by human activities) that have significant influence on the basin’s response to hydrology. One more essence on the applicability of the regionalisation results is that the use of the

flood growth factor map, derived by spatially interpolating the regional flood growth factor, for the different return periods, may be more appealing and reasonable compared to growth curves.

The performance of the regionalisation approach applied to the River Nile at basin scale and where the AMF data are used is considered satisfactory. Nevertheless, the applicability of the results of the regionalisation for engineering practice would require updating and possible comparison with the use of other methods such as peak over threshold or partial duration series data if continuous flow series would become available. In addition, a basin scale study may also be necessary to investigate the rating curves used by each country or water authority to validate the accuracy of the upper quantiles predicted based on the rating curves of the hydrometric stations. Improvements on the limitation of data would definitely improve the accuracy of the growth curves or the growth factor maps and eventually the regionalisation performance. Overall, we believe this study has highlighted both the significance of availability of longer record length data and the importance of regionalisation in extremely data limited River Nile basin. We hope this study will invoke scientific debate and methodological innovation in the River Nile basin and elsewhere in similar challenging river basins for better representation of similar region and enhance applicability in design study.

*Acknowledgements.* This research was conducted at Katholieke Universiteit Leuven and financially supported by the Flemish Interuniversity Council (VLIR) of Belgium. The work was linked to the research activities in the FRIEND/Nile Project funded by the Flemish Government of Belgium through the Flanders-UNESCO Science Trust Fund cooperation and executed by UNESCO Cairo Office. The authors are indebted to UNESCO Cairo Office, the FRIEND/Nile Project management team, the overall coordinator, the thematic coordinators, the theme researchers and the implementing institutes in the Nile countries for the successful execution and smooth implementation of the project, and for facilitating the implementation of the FRIEND/Nile activities.

Edited by: K. Bishop

## References

- Abdo, G., Sonbol, M., and Willems, P.: Flood Frequency Analysis of the Eastern Nile Rivers, International Conference of UNESCO Flanders FIT FRIEND/Nile Project – “Towards a better cooperation”, Sharm El-Sheikh, Egypt, 12–14 November 2005.
- Acreman, M. C. and Sinclair, C. D.: Classification of drainage basin according to their physical characteristics: an application for flood frequency analysis in Scotland, *J. Hydrol.*, 84, 365–380, 1986.
- Aldenderfer, M. S. and Blashfield R. K.: Cluster analysis, Beverly Hills, Sage Press, CA, 1984.
- Alexander, W. J. R.: Flood hydrology of Southern Africa, South African National Committee for Large Dam report, prepared on behalf of SANCOLD, Pretoria, RSA, 3–74, 1990.
- Alila, Y.: A hierarchical approach for the regionalisation of precipitation annual maxima in Canada, *J. Geophys. Res.*, 104, 645–655, 1990.
- Aronica, G. T. and Candela, A.: Derivation of flood frequency curves in poorly gauged Mediterranean catchments using a simple stochastic hydrological rainfall-runoff model, *J. Hydrol.*, 347, 132–142, 2007.
- Bernardara, P., Andreewsky, M., and Benoit, M.: Application of regional frequency analysis to the estimation of extreme storm surges, *J. Geophys. Res.*, 116, C02008, doi:10.1029/2010JC006229, 2011.
- Castellarin, A., Vogel, R. M., and Matalas, N. C.: Probabilistic behavior of a regional envelope curve, *Water Resour. Res.*, 41, W06018, doi:10.1029/2004WR003042, 2005.
- Chebana, F. and Ouarda, T. B. M. J.: Index flood-based multivariate regional frequency analysis, *Water Resour. Res.*, 45, W10435, doi:10.1029/2008WR007490, 2009.
- Clarke, R. T.: Classification procedures in the context of PUB: ways forward?, *Hydrol. Earth Syst. Sci. Discuss.*, 8, 855–867, doi:10.5194/hessd-8-855-2011, 2011.
- Cunnane, C.: Methods and merits of regional flood frequency analysis, *J. Hydrol.*, 100, 269–290, 1988.
- Das, S. and Cunnane, C.: Examination of homogeneity of selected Irish pooling groups, *Hydrol. Earth Syst. Sci.*, 15, 819–830, doi:10.5194/hess-15-819-2011, 2011.
- Di Luzio, M., Srinivasan, R., and Arnold, J.: AVSWAT ArcView interface for SWAT2000 user’s guide, Blackland Research Center, Texas Agricultural Experiment Station, 720 East Blackland Road Temple, Texas, 2001.
- Ellouze, M. and Abida, H.: Regional Flood Frequency Analysis in Tunisia: Identification of Regional Distributions, *Water Res. Manage.*, 22, 943–957, 2008.
- Everitt, B. S., Landau, S., and Leese, M.: Cluster Analysis, 4th Edn., Hodder Arnold, London, 2001.
- Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J. R.: Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form, *Water Resour. Res.*, 15, 1049–1054, 1979.
- Guse, B., Thielen, A. H., Castellarin, A., and Merz, B.: Deriving probabilistic regional envelope curves with two pooling methods, *J. Hydrol.*, 380, 14–26, 2010.
- Hosking, J. R. M.: The theory of probability weighted moments, Research Report RC12210, IBM Thomas J. Watson Research Center, New York, 1986.
- Hosking, J. R. M.: *L*-moments analysis and estimation of distributions using linear combination of order statistics, *J. R. Stat. Soc.*, 52, 105–124, 1990.
- Hosking, J. R. M.: The four parameter Kappa distribution, *IBM J. Res. Dev.*, 38, 251–258, 1994.
- Hosking, J. R. M. and Wallis, J. R.: Regional frequency analysis: an approach based on *L* moments, Cambridge, UK, Cambridge University Press, 1987a.
- Hosking, J. R. M. and Wallis, J. R.: Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics*, 29, 339–349, 1987b.

- Hosking, J. R. M. and Wallis, J. R.: Some statistics useful in regional frequency analysis, *Water Resour. Res.*, 29, 271–281, 1993.
- Hosking, J. R. M. and Wallis, J. R.: A comparison of unbiased and plotting-position estimators of  $L$ -moments, *Water Resour. Res.*, 31, 2019–2025, 1995.
- Hosking, J. R. M. and Wallis, J. R.: Regional flood frequency analysis: an Approach based on  $L$ -moments, Cambridge University Press, New York, 1997.
- Hosking, J. R., M., Wallis, J. R., and Wood, F.: Estimation of the generalized extreme-value distribution by the method of probability weighted moments, *Technometrics*, 27, 251–261, 1985.
- Kachroo, R. K., Mkhanda, S. H., and Parida, B. P.: Flood frequency analysis of Southern Africa – Part 1: delineation of homogeneous regions, *Hydrol. Sci.*, 45, 437–447, 2000.
- Karyabwite, D. R.: Water sharing in the River Nile valley: using GIS/remote sensing for the sustainable use of natural resources, UNEP/DEWA/GRID Project Report GNV011, Geneva, 2000.
- Khaled, H. H.: Trend detection in hydrologic data: The Mann-Kendall trend test under the scaling hypothesis, *J. Hydrol.*, 349, 350–363, 2008.
- Kim, K. and Hawkins, R. H.: Classification of environmental hydrologic behaviours in the north eastern United States, *Water Resour. Bull.*, 29, 449–459, 1993.
- Kim, U. and Kaluarachchi, J. J.: Application of parameter estimation and regionalisation methodologies to ungauged basins of the Upper Blue Nile River Basin, Ethiopia, *J. Hydrol.*, 362, 39–56, 2008.
- Kjeldsen, T. R., Smithers, J. C., and Schulze, R. E.: Flood frequency analysis at ungauged sites in the KwaZulu-Natal Province of South Africa, *Water Resources Commission SA*, 27, 315–324, 2001.
- Kumar, R., Chatterjee, C., Kumar, S., Lohani, A. K., and Singh, R. D.: Development of regional flood frequency relationships using  $L$ -moments for Middle Ganga Plains Subzone 1(f) of India, *Water Resour. Manag.*, 17, 243–257, 2003.
- Kundzewicz, Z. W. and Robson, A.: Detecting trend and other changes in hydrological data, World Climate Programme Data and Monitoring (WCDMP-45, WMO/TD 1013), World Meteorological Organization, Geneva, Switzerland, 2000.
- Linsley, R. K., Kholer, M. A., and Paulhus, J. L. H.: Applied Hydrology, McGraw Hill, New York, 1949.
- Lubomír, S.: Identification of homogeneous regional classes for flood frequency analysis in the light of regional taxonomy, *Hydrol. Sci.*, 50, 1105–1118, 2005.
- LeBoutillier, D. W. and Waylen, P. R.: Regional variations in flow-duration curves for rivers in British Columbia, Canada, *Phys. Geogr.*, 14, 359–378, 1993.
- Li, M., Shao, Q., Zhang, L., and Chiew, F. H. S.: A new regionalisation approach and its application to predict flow duration curve in ungauged basins, *J. Hydrol.*, 389, 137–145, 2010.
- Merz, R. and Blöschl, G.: Flood frequency regionalisation – spatial proximity vs. catchment attributes, *J. Hydrol.*, 302, 283–306, 2005.
- Micevski, T. and Kuczera, G.: Combining site and regional flood information using a Bayesian Monte Carlo approach, *Water Resour. Res.*, 45, W04405, doi:10.1029/2008WR007173, 2009.
- Mkhanda, S., Kachroo, R., and Gunasekara, T. A. G.: Flood frequency analysis of Southern Africa – Part 2: Identification of regional distributions, *Hydrolog. Sci. J.*, 45, 449–463, 1996.
- Mosley, M. P.: Delimitation of New Zealand hydrologic regions, *J. Hydrol.*, 49, 173–192, 1981.
- Nezhad, K., Chokmani, K., Ouarda, T. B. M. J., Barbet, M., and Bruneau, P.: Regional flood frequency analysis using residual kriging in physiographical space, *Hydrol. Process.*, 24, 2045–2055, doi:10.1002/hyp.7631, 2010.
- Northrop, P. J.: Likelihood-based approaches to flood frequency estimation, *J. Hydrol.*, 292, 96–113, 2004.
- Özçelik, C. and Benzedden, E.: Regionalization approaches for the periodic parameters of monthly flows: a case study of Ceyhan and Seyhan River basins, *Hydrol. Process.*, 24, 3251–3269, doi:10.1002/hyp.7752, 2010.
- Ouarda, T. B. M. J. and Shu, C.: Regional low-flow frequency analysis using single and ensemble artificial neural networks, *Water Resour. Res.*, 45, W11428, doi:10.1029/2008WR007196, 2009.
- Parida, B. P., Kachroo, R. K., and Shrestha, D. B.: Regional flood frequency analysis of Mahi-Sabarmati basin (Sub-zone 3a) using index flood procedure with  $L$ -moments, *Water Resour. Manag.*, 12, 1–12, 1998.
- Prescott, P. and Walden, A. T.: Maximum likelihood estimation of the parameters of the three parameter generalized extreme-value distribution from censored samples, *J. Stat. Comput. Sim.*, 16, 241–250, doi:10.1080/00949658308810625, 1983.
- Ramachandra Rao, A. and Srinivas, V. V.: Regionalization of watersheds by hybrid-cluster analysis, *J. Hydrol.*, 318, 37–56, 2005.
- Roger, K. B.: The growth of cluster analysis: Tryon, Ward, and Johnson, *Multivar. Behav. Res.*, 15, 459–458, 1980.
- Sarhadi, A. and Modarres, R.: Flood seasonality-based regionalisation methods: a data-based comparison, *Hydrol. Process.*, 25, 3613–3624, doi:10.1002/hyp.8088, 2011.
- Schmidt, E. J. and Schulze, R. E.: Flood volume and peak discharge from small catchments in southern Africa based on the SCS technique, Technology Transfer Report TT/3/87, Water Research Commission, Pretoria, 1997.
- Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E.: Frequency analysis of extreme events, in: Handbook of hydrology, edited by: Maidment, D. R., Chapter 8, 66 pp., McGraw-Hill, Inc., New York, 1993.
- Sutcliffe, J. V. and Parks, Y. P.: The hydrology of the Nile, The International Water Management Institute, Colombo, Sri Lanka, Gibb Water, Reading, UKIAHS Special Publication, 5, ISBN 1-901502-75-9, 1999.
- Vogel, R. M. and Fennessey, N. M.:  $L$ -moment diagrams should replace product moment diagrams, *Water Resour. Res.*, 29, 1745–1752, 1993.
- Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment Classification and Hydrologic Similarity, *Geogr. Compass*, 1, 901–931, doi:10.1111/j.1749-8198.2007.00039.x, 2007.
- Willems, P., Sonbol, M., Gamal, S., Mkhanda, A., Opere, A., Tedesse, L., Abdel Motaleb, M., Farid, S., Zaki, A., and Al-Weshah, R.: Regional flood frequency analysis in the Nile basin. International Conference of UNESCO Flanders FIT FRIEND/Nile Project-“Towards a better cooperation”, Sharm El-Sheikh, Egypt, 12–14 November 2005, CD-ROM Proceedings, 18, 2005.
- Willems, P., Guillou, A., and Beirlant, J.: Bias correction in hydrologic GPD based extreme value analysis by means of a slowly varying function, *J. Hydrol.*, 338, 221–236, 2007.

- William, H. E. D. and Edelsbrunner, H.: Efficient algorithms for agglomerative hierarchical clustering methods, *J. Classif.*, 1, 7–24, 1984.
- Wiltshire, S. E.: Regional flood frequency analysis – Part 1: Homogeneity statistics, *Hydrolog. Sci. J.*, 31, 321–333, 1986.
- Yue, S. and Wang, C. Y.: Possible regional probability distribution type of Canadian annual stream flow by *L*-moments, *Water Resour. Manag.*, 18, 425–438, 2004.
- Zrinji, Z. and Burn, D. H.: Regional flood frequency with hierarchical region of influence, *J. Water Res. Pl.-ASCE*, 122, 245–252, 1996.