Journal of Hydrology (2007) **347**, 243–259

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jhydrol

Global streamflows — Part 1: Characteristics of annual streamflows

Thomas A. McMahon [a,*], Richard M. Vogel [b], Murray C. Peel [a], Geoffrey G.S. Pegram [c]

[a] *Department of Civil and Environmental Engineering, The University of Melbourne, Victoria 3010, Australia*
[b] *Department of Civil and Environmental Engineering, Tufts University, Medford, MA, USA*
[c] *Civil Engineering Programme, University of KwaZulu-Natal Durban, South Africa*

Received 15 February 2007; received in revised form 20 July 2007; accepted 3 September 2007

KEYWORDS
Global hydrology;
Global streamflow data;
Global streamflow characteristics;
Global rivers

Summary This is the first of three related papers that summarizes streamflow characteristics of a set of 1221 global rivers. The rivers are well distributed world-wide, are unimpacted by upstream reservoirs or diversions for the period of data collection and have at least 10 years of continuous monthly and annual streamflow data. The following key features of annual flows are examined: mean, variability and skewness, distribution type (Gamma or Lognormal), flow percentiles and dependence. High and low frequency persistence is examined through the Empirical Mode Decomposition technique. Low flow run length, magnitude and severity are also explored, where severity is based on Extended Deficit Analysis. It has been observed elsewhere that there are large differences in hydrologic characteristics between Australia and southern Africa in contrast to the rest of the world. This issue is tested further in this paper. The range of analyses and results presented herein also form a suite of empirical evidence that future unified theories of hydrology at the catchment scale must be able to adequately describe.

Introduction

This is the first of a series of three papers that describes the hydrologic characteristics of a unique global data set, which consists of historical annual and monthly streamflow records of 1221 rivers. With the likelihood of climate change having an impact on future streamflow, we develop and discuss key characteristics of both unregulated rivers (this paper) and hypothetically regulated ones (paper 2, McMahon et al., 2007a). We also examine the variations in hydrologic characteristics on a country and on a climatic basis in paper 3 (McMahon et al., 2007b). In addition, the global data set enables us to examine, on a world-wide basis, climate change

* Corresponding author. Tel.: +61 3 8344 7731; fax: +61 3 8344 6215.
E-mail addresses: t.mcmahon@civenv.unimelb.edu.au (T.A. McMahon), richard.vogel@tufts.edu (R.M. Vogel), mpeel@civenv.unimelb.edu.au (M.C. Peel), pegram@ukzn.ac.za (G.G.S. Pegram).

0022-1694/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2007.09.002

impacts on gauged and, by inference, ungauged catchments. Further, it allows us to explore broad differences in regional hydrologic features and to make world-wide hydrologic comparisons of surface water resources.

Although paper 3 (McMahon et al., 2007b) of this series concentrates largely on continental, regional and climatic differences in hydrologic characteristics of the annual streamflow time series, we have deliberately included in this paper observed differences between hydrological responses in Australia—southern Africa (ASA) and the rest of the world (RoW).

We had four objectives in carrying out this project. Firstly, there have been a number of papers highlighting the hydrologic differences between ASA and RoW and discussing the reasons for the differences (McMahon et al., 1992; Peel et al., 2001; Peel et al., 2002; Peel et al., 2004a). With this enhanced data set, we wish to review these results and evaluate them more fully. Moreover, we wanted to extend the analysis to include reservoir capacity estimates and also reservoir performance at continental, country and regional climatic scales (using the Köppen classification) (Köppen, 1936). As far as we can determine, analysis relating to reservoir capacity and performance at these large geographic scales has not yet been addressed. The second objective was to develop empirical relationships between hydrologic characteristics (for both the natural flows and hypothetically regulated rivers) based on the global data set and relevant independent parameters. Such relationships would be useful to water managers concerned with exploring surface streamflow characteristics by continent, region, country or climate zones. Furthermore, the relationships could be used to estimate surface streamflow characteristics of ungauged rivers. The third objective of this project is to make available to modellers of global runoff, streamflow characteristics that are based on measured values which can be used for comparison with model estimates. Furthermore, the analysis provides a global characterization of hydrologic characteristics which grid-based models must be able to reproduce. The fourth objective relates to the search for a unified theory of hydrology at the catchment scale (Sivapalan, 2005). In presenting the annual and monthly characteristics of streamflow globally, we provide a range of empirical evidence that future theories of hydrology at the catchment scale must be able to adequately describe. This ties in with adopting a top-down approach to our analysis which, in the context of our global data set of streamflows, involves learning how surface streamflow characteristics vary across a range of scales from small catchments to continents and within a range of climate zones. Using the hypothesis of top-down modelling namely that ''... at the whole-catchment scale much of the complexity needed to model hydrological response at finer scales is unnecessary...'' (Post et al., 2005), we explore and develop relationships that will be of assistance to both water managers and hydrologic modellers.

Following this introduction we review the key literature relevant to this paper. Details of the historical streamflow data are discussed in Section 'Annual streamflow data'. 'What probability distribution function?' is examined in Section 'What probability distribution function?' followed by a discussion of the key statistical characteristics in Section 'Mean, variability and skewness'. In Section 'Flow percentiles' we present an analysis of flow percentiles which is followed in Section 'Dependence' by an examination of auto-correlation in annual data. Section 'Persistence characteristics' examines persistence through Empirical Mode Decomposition (EMD). The final section describes analyses of low flow run length and magnitude, then severity using Extended Deficit Analysis (EDA). Some conclusions are presented in Section 'Summary and conclusions'.

Literature

As far as we are aware there is no other series of papers that has captured, assessed and reported on such a range of surface hydrologic characteristics world-wide as is attempted here. Most papers and reports that addressed global runoff dealt mainly with mean or median flows and in only a few cases with variability, skewness and/or auto-correlation.

One of the earliest attempts to compute global mean runoff was by Ayushinskaya et al. (1977) which was published about the same time as a number of other atlases of the world water balance (Korzun et al., 1974; Van Der Leedon, 1975; UNESCO, 1977), the most comprehensive being the report by Baumgartner and Reichel (1975) and by UNESCO (1978). These papers and reports concentrated on the mean annual water balance although several discussed streamflow variability. A recent report by Shiklomanov and Rodda (2003) presents an up-to-date analysis of world water resources and, inter alia, includes information of streamflow variability and skewness for selected rivers, countries and continents.

Kalinin (1971) carried out the first detailed analyses of river flows world-wide, but only 8 of his 137 catchments were from the southern hemisphere. McMahon (1975, 1977, 1979) carried out the first analyses of southern hemisphere rivers along the lines adopted by Kalinin and described some interesting deviations from Kalinin's conclusions. These analyses were extended to a world-wide data set by McMahon and his colleagues over the next two decades (see for example, McMahon et al., 1992; Peel et al., 2001; Peel et al., 2004a).

An extensive report of continental comparisons of streamflow characteristics by McMahon et al. (1992) utilizes the world-wide data set of annual flows and peak discharges noted in the previous paragraph. The report describes the variability, skewness, distribution types and auto-correlation of both annual flows and annual peak discharges. In addition, it examines the time series structure of annual flows. Finally, it explores inter-continental comparisons along with those relating to seasonal flow regime stratification (as defined by Haines et al. (1988)) and to climate using the Köppen classification (Köppen, 1936). Of the many conclusions in McMahon et al. (1992), the key one relates to the hypothesis that there are major differences in terms of runoff characteristics between the northern and southern hemispheres; however, that hypothesis could not be confirmed. What was found, however, was that the annual characteristics of streamflow for Australia and southern Africa are different from those found for streams in the rest of the world. This observation has been examined by Peel et al. (2001, 2004a) who concluded that the variability of annual runoff for temperate Australia and temperate and arid southern Africa was higher than for other continents

across similar climatic zones. They further concluded that for temperate areas the distribution of evergreen and deciduous vegetation was a potential cause of the observed variability differences, along with differences in the variability of annual precipitation, percentage of forested catchment area, seasonality of precipitation and mean annual daily temperature range.

Peel et al. (2004b, 2005) analyzed a slightly larger data set than that used herein. They found that, for annual streamflow, run lengths of years below the median were similar across all continents except for the Sahel region in northern Africa where a distinct bias towards longer lengths was evident. However, in terms of relative magnitude of the deficits below the median flow, Australia and southern Africa exhibited values that were much larger than elsewhere. These issues are reviewed in more detail herein and in the third paper in this series.

Both Yevdjevich (1963) and Kalinan (1971) suggested that the annual time series of world rivers were normally distributed, a feature not found by McMahon's (1975, 1977, 1979) analysis. Markovic (1965) studied five distributions (Normal, Lognormal 2 and 3 parameter, Gamma 2 and 3 parameter) and concluded that for rivers in western United States and southwestern Canada, all five probability distribution functions describe the annual flows satisfactorily, though more recently, Vogel and Wilson (1996) found that the Gamma 2 distribution was the only two parameter distribution which was satisfactory for US rivers. This aspect is explored in Section 'What probability distribution function?', because the shape of the tail of the probability distribution describing flow volumes has a critical influence on the reliability of reservoir systems.

Annual streamflow data

The global data set consists of 1221 unregulated rivers with 10 years or more of continuous historical annual and monthly flows. The water year was the basis of the annual flow volumes. Fig. 1 shows the locations of the rivers and it is noted that they are reasonably distributed globally, although there are some regions (for example, arid regions — Mediterranean North Africa, the Middle East, South Western Africa, central Australia and tropical regions — central America, Indonesia, non-coastal Brazil, Peru, Bolivia, Ecuador) with little or no data.

The data set was initially collated by the first author during the 1980s and the details were reported several years later (McMahon et al., 1992), with subsequent additions and revisions to the data set since that time (Peel et al., 2001, 2004a). Considerable effort has gone into ensuring the data are free of errors, are not impacted by major water withdrawals from the streams and the flow values are not affected by upstream reservoirs. However, except for obvious transcription errors and those related to catchment areas and location, measurement errors particularly relating to the adequacy of rating curves have not been addressed. Such a task is well beyond a project of this magnitude.

Fig. 2 shows for the global data set a plot of length of continuous record and catchment area. The catchment areas vary from 115 to 65 200 km^2 (10th percentile to 90th percentile) with a median area of 1670 km^2 and with lengths from 15 to 58 years (10th percentile to 90th percentile) and a median of 29 years. Because of the difficulty in collating monthly and annual streamflow data without omitting parts of the records, we were unable to utilize a common period in our analyses. This limitation may affect some results and needs to be kept in mind in their interpretation. In the main, the period from 1950 to 1980 is best represented in the data set.

As a further introduction to the data, Table 1 sets out some key hydrologic features of the rivers. The rivers cover a very wide range of hydrologic regimes, with mean annual runoff varying from a minimum of 0.373 mm to a maximum of 5370 mm. Great variability is also observed in the annual coefficient of variation (C_v) from 0.062 to 2.97. In Table 1 the statistics for the coefficient of skewness (γ) and lag-one serial correlation coefficient (auto-correlation) (ρ) of annual flows are based on both a minimum of 10 and 25 years of annual flow records respectively, the latter shown in parentheses. Again, large variations in these parameters is evident with skewness varying from −2.2 to 6.1 and auto-correlation ranging from −0.48 to 0.90 for 25 years or more data.

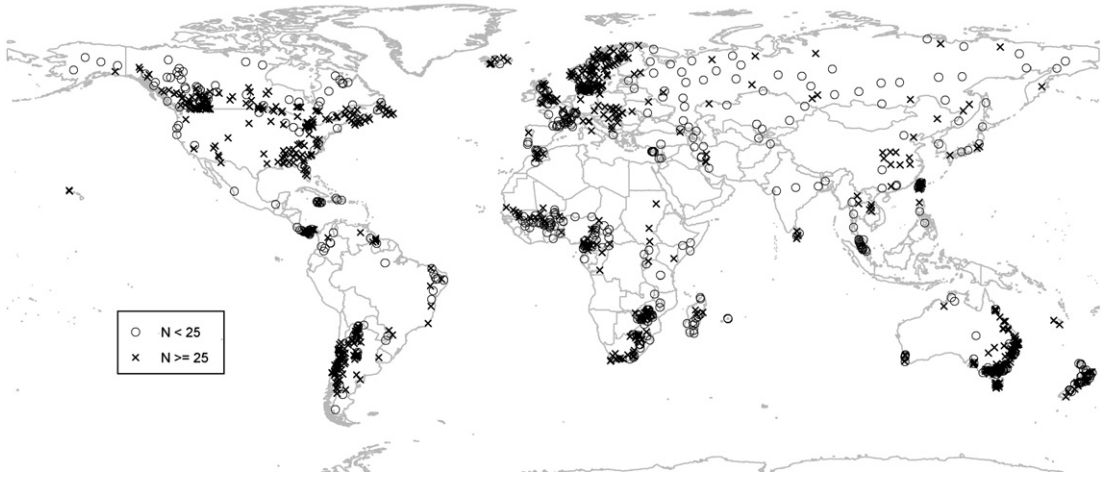

Figure 1 Location of streamflow stations (1221) with 10 or more years of continuous streamflow data (plate carrée projection).

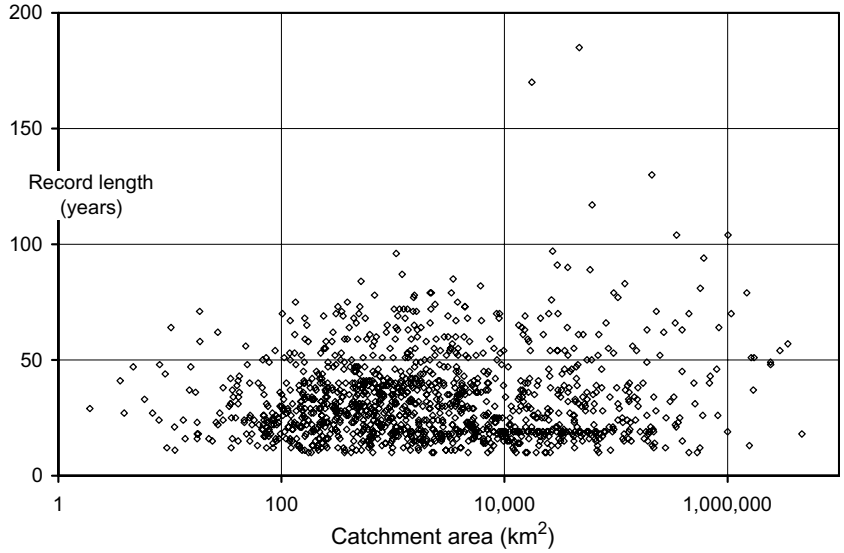

Figure 2 Length of continuous record versus catchment area.

Table 1 The key hydrologic features of the 1221 rivers used in this paper

	Area (km^2)	N (years)	μ (mm)	C_v	γ	ρ
Median	1610	29	339	0.314	0.552 (0.561)[a]	0.103 (0.119)
Maximum	4,640,000	185	5370	2.97	6.14 (6.14)	0.936 (0.895)
Minimum	0.360	10	0.373	0.0619	−2.22 (−2.22)	−0.756 (−0.482)

Area is catchment area (km^2), N is the number of years of continuous annual or monthly data, μ is the mean annual runoff (mm), C_v is the coefficient of variation of annual flows, γ is the coefficient of skewness of annual flows and ρ is lag-one serial correlation coefficient.
[a] Values in parentheses are computed using 25 or more years of historical data (729 rivers).

What probability distribution function?

A perennial challenge in hydrology is to choose an appropriate probability distribution function (pdf) to describe streamflow data. Over the past fifty or so years many papers and reports have dealt with this question in relation to annual streamflows — Yevdjevich (1963), Markovic (1965), Löf and Hardison (1966), Kalinin (1971), McMahon et al. (1992), Vogel and Wilson (1996) and others. We do not plan to traverse details in this paper but suffice to say that for Europe and portions of North America the normal pdf is often adequate but the 2-parameter lognormal (LN2) and Gamma are more generally applicable. The normal distribution was found to provide a poor approximation to the distribution of annual flows in ASA.

To explore this issue further, we use *L*-moment diagrams to assess whether the global annual flows tend towards a Lognormal (LN2 and LN3) or a Gamma (or Pearson III) distribution. To do this we developed theoretical relationships between *L*-C_v and C_v and *L*-Skewness and the coefficient of skewness (γ) for these distributions, which are described in the Appendix.

Based on the theoretical analysis in the Appendix, traditional *L*-moment diagrams for annual average streamflows based on the global dataset of 1221 rivers are provided in Fig. 3. What one observes from these two figures is that it is nearly impossible to distinguish whether the observations arise from either a Lognormal or a Gamma pdf. As an innovation, in Fig. 4 we show relationships between (a) C_v and *L*-C_v and (b) Skewness and *L*-Skewness for the same data set. In each of these figures it is quite apparent that a Gamma pdf provides a much better fit to the observations than the LN2.

It is well known that ordinary product moments are biased downward (Vogel and Fennessey, 1993) which is why the estimated values of both C_v and Skewness tend to be a bit lower than the theoretical curves in Figs. 4a and b, respectively. However, even given this bias, Fig. 4 enables us to distinguish the tail behaviour between the Lognormal and Gamma probability distribution functions more effectively than through the use of traditional *L*-moment diagrams as shown in Fig. 3. It is possible that the downward bias in estimates of both C_v and Skewness still impedes our ability to discern which distribution is best, though there is some evidence here that a Gamma distribution function fits the global historical annual streamflows better than a Lognormal distribution function. This result is consistent with other recent investigations using *L*-moment diagrams and very large datasets (Vogel and Wilson, 1996). Fig. 4 indicates that a promising avenue for future research involves

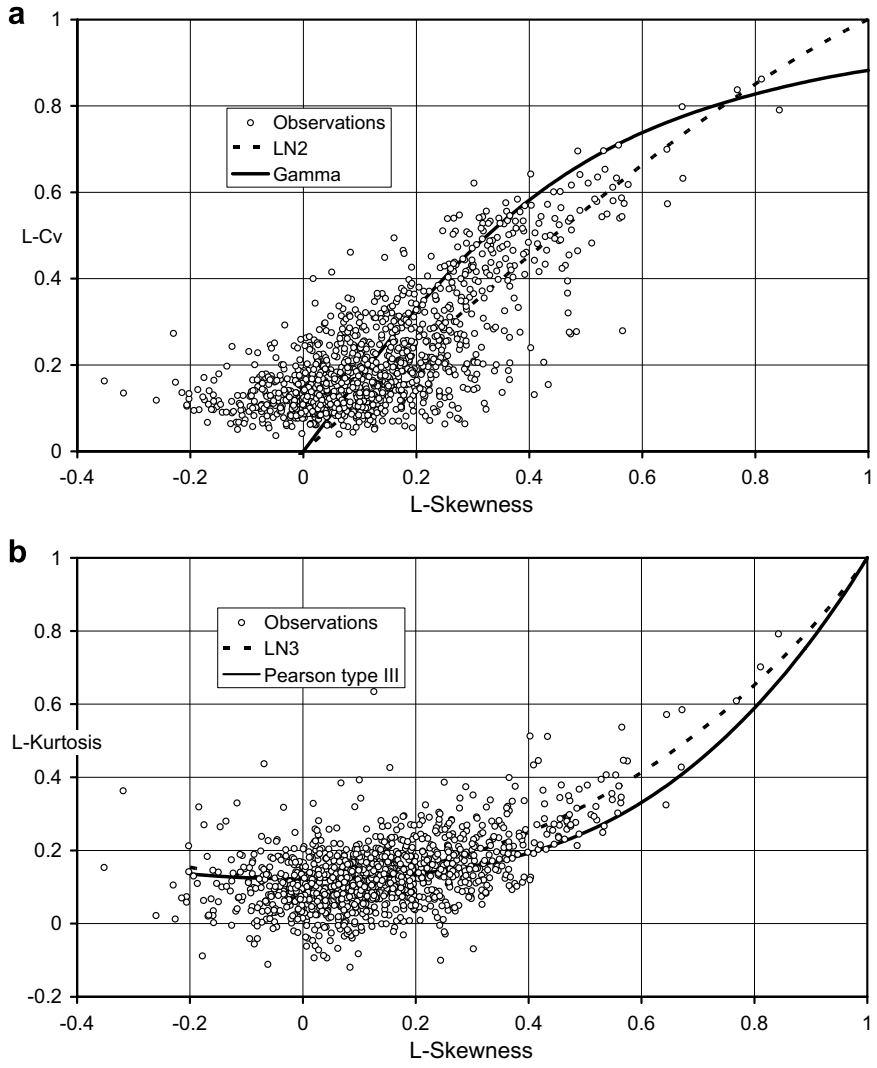

Figure 3 Relationships between L-C_v, L-Skewness and L-Kurtosis: (a) L-C_v versus L-Skewness, (b) L-Kurtosis versus L-Skewness.

plotting two different unbiased measures of C_v, skewness, or perhaps kurtosis against each other to discriminate tail behaviour.

Mean, variability and skewness

Fig. 5 shows the relationship between mean annual flow and the standard deviation of annual flows for the 1221 global rivers. We note the separation of the two curves representing ASA and RoW. For example, for a river with a mean annual flow (μ) of 1000×10^6 m^3 the ratio of the standard deviation (σ) of flow of such a river in ASA would be $1.78 \times \sigma_{RoW}$ where σ_{RoW} is the standard deviation of a typical river in the RoW. The global equation defining the relationship between σ and μ was fit using weighted least squares (WLS) regression which resulted in

$$\sigma = 0.752\mu^{0.876} \quad (1)$$

$$n = 1221,\ R^2 = 0.940,\ \text{se} = +72\%, -42\%$$

where μ is the mean annual streamflow (10^6 m^3), σ is the standard deviation of the annual flows (10^6 m^3), n is the number of rivers in analysis, R^2 is the square of the correlation coefficient and se is the standard error of estimate. The weightings in the WLS procedure used here and elsewhere in this paper are historical record lengths.

We have also related μ to catchment area for ASA and RoW by the following empirical relationships which were obtained using WLS regression:

$$\mu_{ASA} = 1.013\ \text{Area}^{0.727} \quad (2)$$

$$n = 263,\ R^2 = 0.60,\ \text{se} = +290\%,\ -74\%,\ p =< 0.001$$

$$\mu_{RoW} = 1.526\ \text{Area}^{0.818} \quad (3)$$

$$n = 958,\ R^2 = 0.78,\ \text{se} = +199\%,\ -67\%,\ p =< 0.001$$

where the subscripts ASA and RoW refer to Australia—southern Africa and rest of the world respectively, Area represents the drainage area in km^2 and p is the probability of the value of R^2 occurring by chance. Substituting a typical range of catchment areas (10 and 100,000 km^2) into Eqs. (2) and (3), rivers in the RoW yield between about 100% and 300% more annual flow than rivers in ASA.

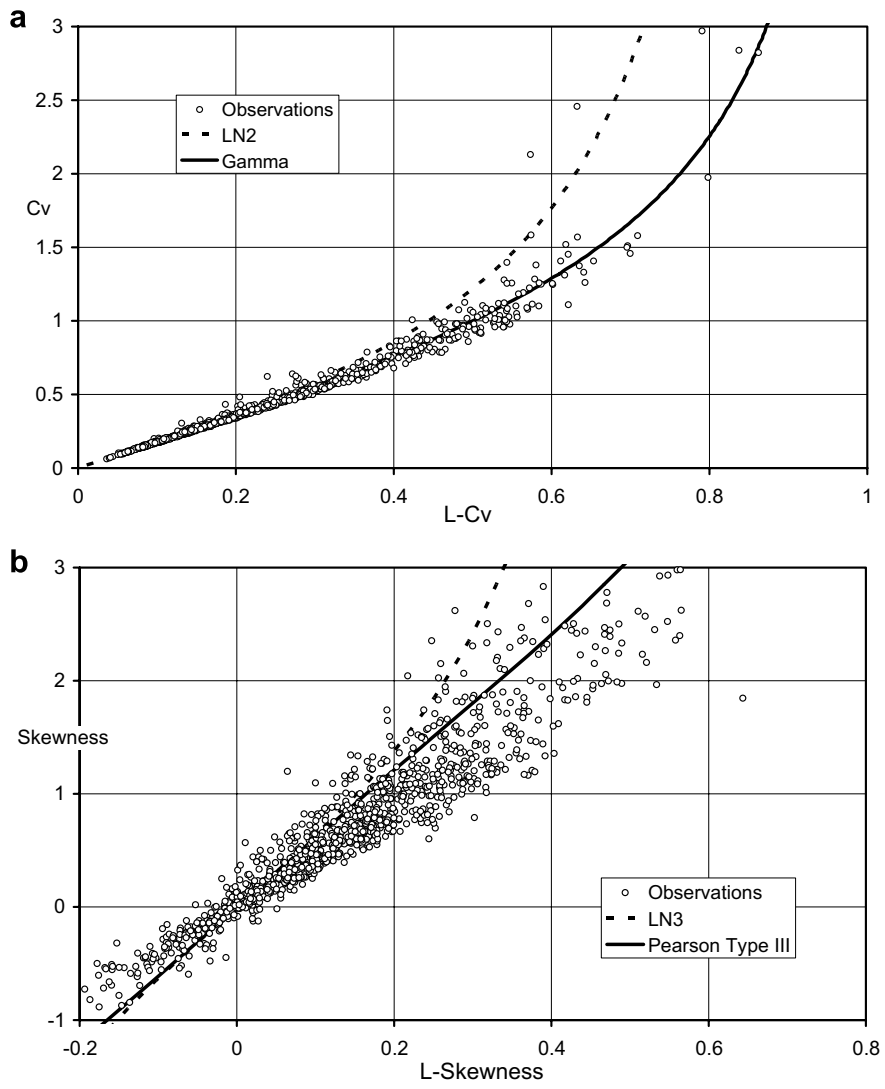

Figure 4 Product-moment versus *L*-moments: (a) Product-moment C_v versus L-C_v, (b) Coefficient of skewness versus *L*-Skewness.

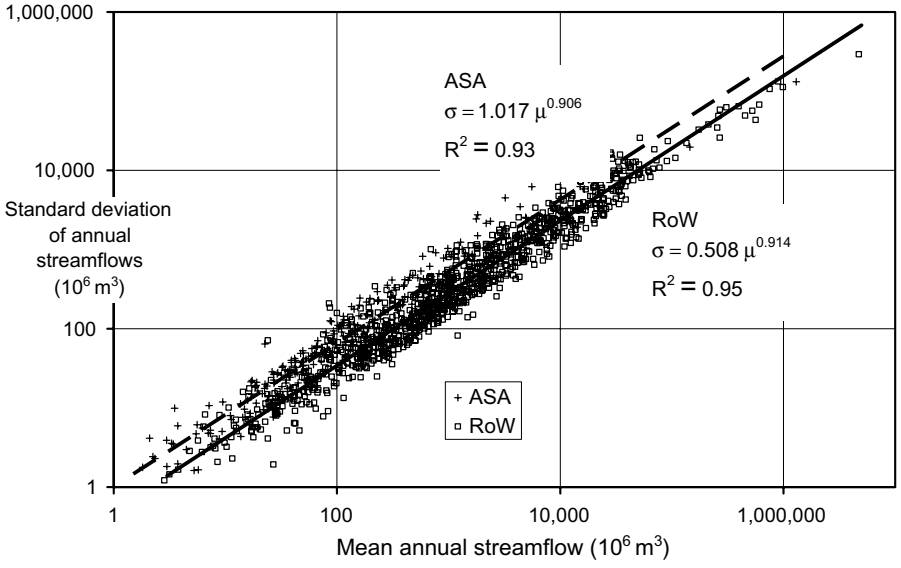

Figure 5 Standard deviation of annual flows versus mean annual flow. Solid line = RoW, broken line = ASA.

The next two figures explore the relationship between the C_v of annual flows and mean annual runoff (MAR) in mm (Fig. 6) and catchment area (Fig. 7). Similar figures have been produced elsewhere (see for example, McMahon et al., 1992) but are reproduced here for completeness. In Fig. 6, C_v decreases with an increase in MAR (which can be considered a surrogate for mean rainfall). The global curve fit using WLS is

$$C_v = 1.828\,\mathrm{MAR}^{-0.299} \tag{4}$$

where C_v is the coefficient of variation of annual streamflows and MAR is the mean annual runoff (mm). No statistics are provided here as the values would be spurious, given the mean is included in both sides of the equation.

The generalised relationship of Kalinin (1971) was also applied to the data set and the resulting curve superimposed on Fig. 6 is

$$C_v = \sqrt{5.82\,\mathrm{MAR}^{-1} + 0.206} \tag{5}$$

and it clearly provides a poor fit to the data, particularly for high mean annual runoff where it overestimates annual C_v.

The relationship between C_v and catchment area is given in Fig. 7. Based on the entire global data set, there is little association between the two variables but again the difference between the ASA region and the RoW is clear. The equation fit using WLS with all the data is

$$C_v = 0.447\,\mathrm{Area}^{-0.0384} \tag{6}$$
$$n = 1221,\ R^2 = 0.024,\ \mathrm{se} = +85\%,\ -46\%$$

where Area is catchment area (km^2). The generalised relationship between C_v and catchment area proposed by Kalinin (1971) is also applied to the data set and is clearly different from Eq. (6) (see Fig. 7).

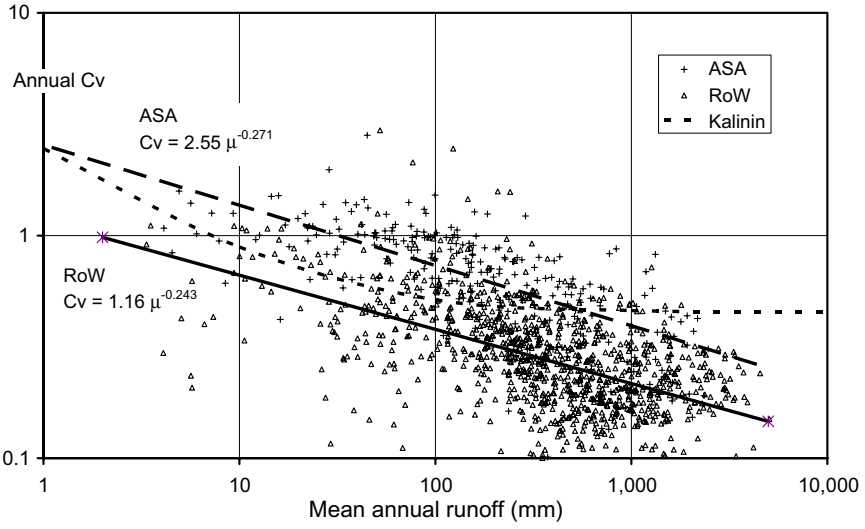

Figure 6 Coefficient of variation of annual flows versus mean annual runoff. Solid line = RoW, broken line = ASA.

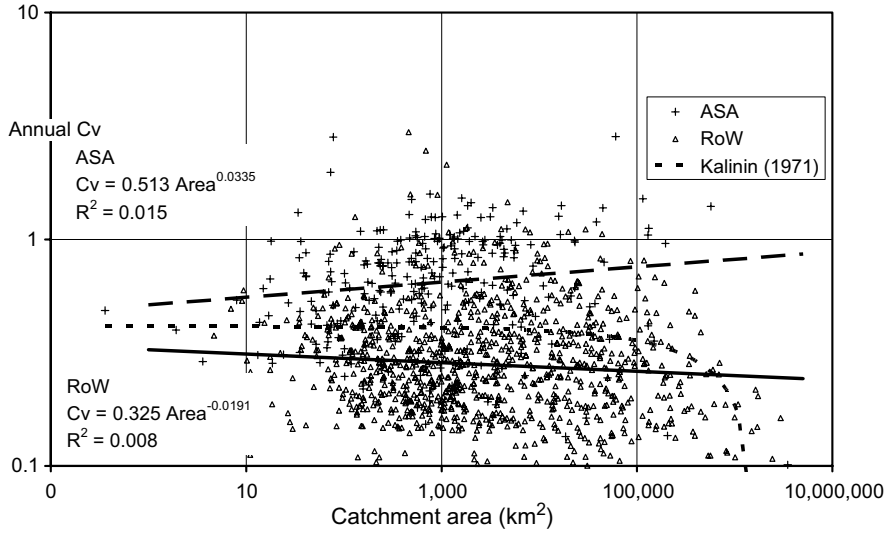

Figure 7 Coefficient of variation of annual flows versus catchment area. Solid line = RoW, broken line = ASA.

In this context we note that large C_v's imply larger reservoir storages or lower system yields. For example, the following equation based on the Gould–Dincer Gamma method (see McMahon and Adeloye, 2005) for estimating required over-year storage capacity can be used to explore this further.

$$S = \frac{z_p^2}{4(1-\alpha)} C_v^2 \mu \frac{1+\rho}{1-\rho} \quad (7)$$

where S is the required reservoir capacity to meet a draft of α (target draft divided by mean annual flow), z_p is the standardized Gamma variate at 100p% probability of non-exceedance, i.e. probability of not being able to meet the draft, C_v is the coefficient of variation of annual inflows, μ is the mean annual inflow, and ρ is the annual lag-one serial correlation coefficient.

From Eq. (7) we can compare two rivers, regions or continents by taking the ratio of the two estimated reservoir capacities, assuming the same draft and failure conditions in systems A and B, leading to

$$\frac{S_A}{S_B} = \left(\frac{C_{v_A}}{C_{v_B}}\right)^2 \frac{\mu_A}{\mu_B} \frac{1+\rho_A}{1-\rho_A} \frac{1-\rho_B}{1+\rho_B} \quad (8)$$

To compare required reservoir capacities in ASA with those in the RoW, a common catchment size of 10,000 km^2 is adopted. From Eqs. (2) and (3), $\mu_{ASA} = 0.312\,\mu_{RoW}$ where inflows μ are per unit catchment area. Also we note that $\rho_{ASA} \approx \rho_{RoW}$ (see Fig. 9), thus

$$\frac{S_{ASA}}{S_{RoW}} = 0.312 \left(\frac{C_{v_{ASA}}}{C_{v_{RoW}}}\right)^2 \quad (9)$$

Noting that $C_{v_{ASA}} \approx 2.6\ C_{v_{RoW}}$, (from Fig. 7 for area = 10,000 km^2), then for these specific conditions, reservoirs necessary to meet over-year storage requirements in the ASA continental region need to be a little more than twice as large as those in the RoW for the same level of regulation (i.e. for the same target draft and reliability), before adjusting for net evaporation losses, which would exacerbate the difference.

Flow percentiles

In Fig. 8, the empirical 10th and 90th exceedance percentiles of annual flow values of each record are plotted against the mean annual runoff. The differences between the two continental areas are evident. A plot relating the percentile values to catchment areas (not included here) exhibits similar differences. The 10th percentile values in Fig. 8 for ASA are 22% larger than the equivalent values for a catchment with MAR of 500 mm located in the RoW. For the 90th percentile flows ASA flows are typically 28% smaller than the RoW percentiles, again for a catchment with 500 mm of annual runoff. These differences relate directly to the relatively larger variance observed for ASA rivers compared to those for the RoW (Figs. 5–7).

Dependence

In examining the auto-correlation of annual streamflows of the global rivers, only values that are statistically significantly different from zero are considered. Fig. 9 shows a plot of estimates of the statistically significant lag-one serial correlation ρ against catchment area. Of the 1221 rivers in the global data set, 249 values were significantly different from zero at the 5% level, that is, about four times as many as would be expected by chance. Eleven of the rivers exhibited statistically significant negative auto-correlations. The average magnitude of the statistically significant auto-correlations (absolute values) is 0.441. The differences in average absolute values between ASA and RoW are negligible, 0.424 and 0.445 respectively.

A relationship between statistically significant positive ρ values and the independent variables record length N and catchment area (km^2) was fit using WLS regression resulting in

$$\rho = 0.832 N^{-0.281}\ \text{Area}^{0.0344} \quad (10)$$
$$(n = 249,\ R^2 = 0.22)$$

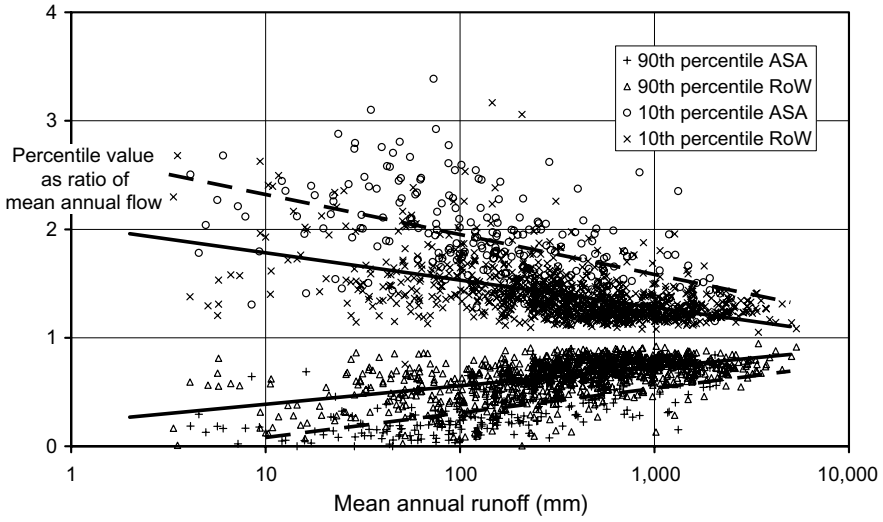

Figure 8 10th and 90th percentile values of annual flows (as ratio of mean annual flow) versus mean annual runoff (mm). Solid line = RoW, broken line = ASA.

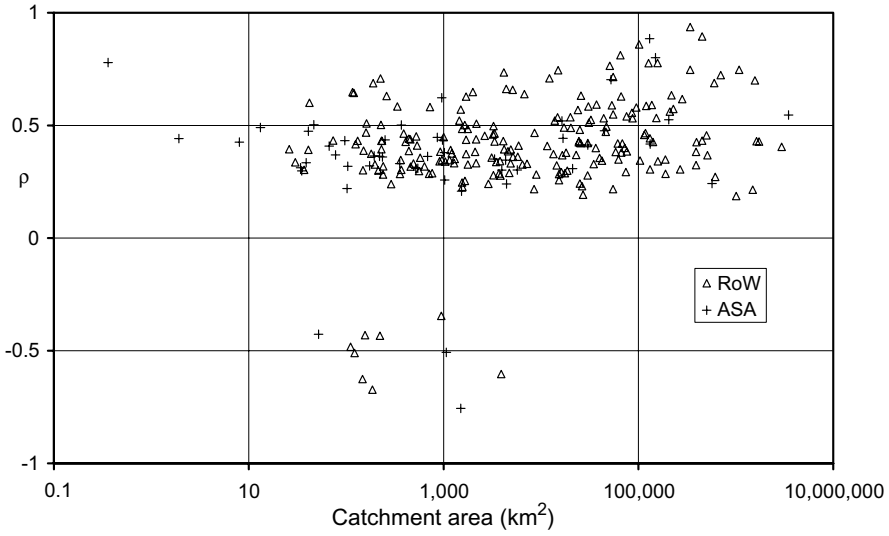

Figure 9 Statistically significant lag-one serial correlation of annual flows versus catchment area.

This overall relationship may be statistically significant although the residuals are not normally distributed. Nevertheless, it suggests that lag-one serial correlation is inversely related to record length and positively related to area. Approximately 90% of the variance in the relationship is accounted for by the record length term. The positive relationship of ρ with area, although weak and not statistically significant, is consistent with our understanding that high positive auto-correlation maybe related to water carryover in catchments from year to year. Yevdjevich (1964) argues that dependence is not only a function of carryover storage (see also Klemes, 1970) but is also affected by inconsistencies and inhomogeneities in the data. However, in Yevdjevich's annual runoff data set, estimates of annual ρ increased with increases in record length, although he noted that this conclusion should be accepted very cautiously. In summary, our analysis suggests that ρ is inversely related to record length and that it is positively, but weakly, related to catchment area. This issue is discussed further in Section 'Persistence characteristics'.

Also observed in Fig. 9 are values of ρ for some rivers that appear to be excessively high. After closer examination of these values and reviewing their spatial location we note that rivers with high values of ρ generally: (1) are in regions where permanent snow cover is likely to be contributing to the streamflow (e.g., Iceland, northern Canada, Chile/Argentina); (2) have natural lakes within the catchment (e.g., tropical East Africa); (3) experience a significant trend in streamflow during the period of record (e.g., Sahel); or (4) have a short record length (sampling variability).

The 11 rivers with significant negative auto-correlations are widely distributed and show no clear spatial pattern. A review of their flow characteristics and other potentially explanatory variables including climate, as defined by the Köppen classification, identified a common link as being length of data. Fig. 10 shows the relationship between the high negative ρ values and N fit using WLS regression. The high correlation ($R^2 = 0.754$) between the two variables would suggest that generally the high negative ρ values are mainly the result of sample size and sample variability

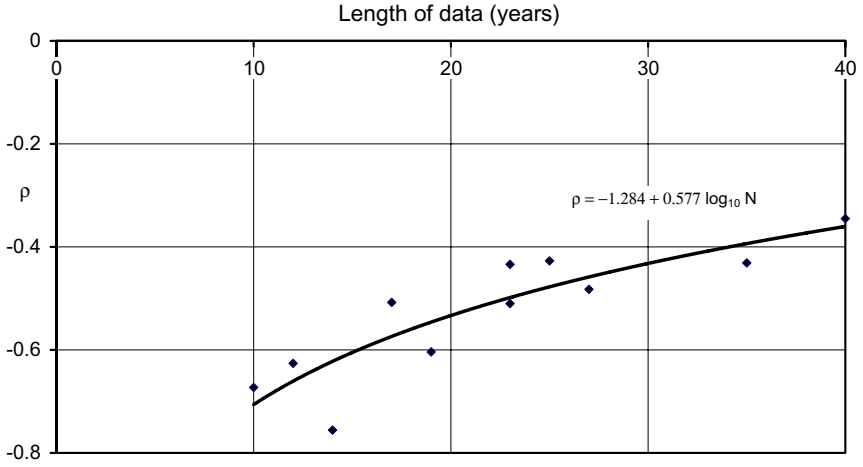

Figure 10 Relationship between the statistically significant negatively lag-one serial correlations and record length.

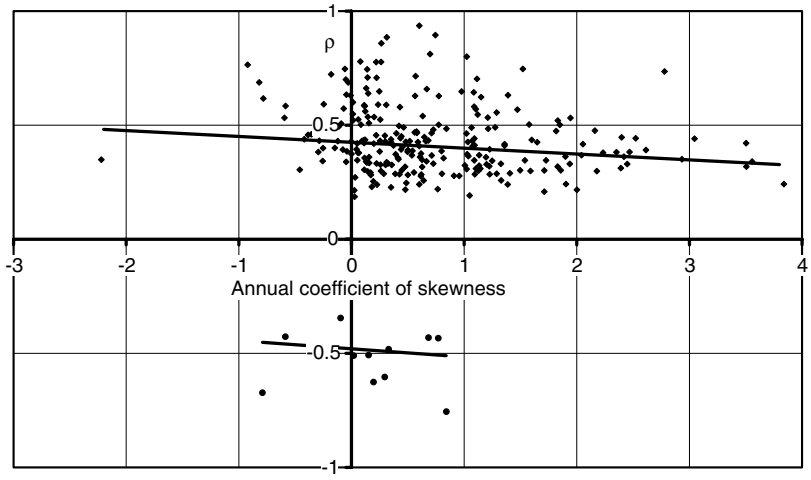

Figure 11 Statistically significant positive and negative lag-one serial correlation of annual flows versus coefficient of skewness of annual flows.

rather than some other factor. This explanation is plausible given that, unlike positive correlations, there is no physical explanation why large negative correlations should occur at the annual level.

Fig. 11, a plot of statistically significant positive and negative ρ against the coefficient of skewness of annual flows, shows that auto-correlation is not correlated with skewness (weighted least squares analysis R^2 for $+\rho = 0.018$ and $-\rho = 0.000$). Based on a Monte Carlo modelling study and a review of the eight most negatively skewed series of annual streamflow from 140 rivers compiled by Yevdjevich (1963), Klemes (1970) postulated that negatively skewed annual streamflows are likely to occur in catchments with large carry-over water storage and, hence, high positive auto-correlations. Fig. 11 would suggest otherwise especially as many of the large positive auto-correlations are associated with positive skewness.

Persistence characteristics

Inter-annual characteristics, i.e. high frequency features, were considered in the previous section where lag-one serial correlations of annual flows were discussed. In this section we examine low frequency characteristics by using Empirical Mode Decomposition (Huang et al., 1998) to examine the proportion of variance in the historical annual time series of flows from the global data set which is due to low frequency fluctuations. However, to ensure there are sufficient data to define the characteristics we have restricted the analysis to 595 rivers that have at least 30 years of historical data.

EMD is an adaptive form of time-series decomposition for non-linear and non-stationary data, and therefore appropriate for time-series of annual streamflows. It decomposes a time series into a set of independent intrinsic mode functions (IMFs) and a residual component. If the IMFs and residual are summed together they form the original time series. Depending on the nature of the time-series being studied, the IMFs and the residual component may exhibit linear or non-linear behaviour (amplitude and frequency modulation).

Application of EMD analysis is illustrated in Fig. 12 using the Zambesi River at Victoria Falls. The figure shows the historical time series, three IMFs and the residual. We chose this example as it is a difficult river to model yet the residual shows the trend very clearly. EMD analysis allows one to determine the average period for each IMF and the variance associated with each IMF and the residual. These values are shown in Table 2. Because of the major increase in the Zambesi flows ($\sim$70% after the 1940s and 1950s) we observe a very high proportion of variance, approximately 47%, accounted for in the residual. We note that the concept of ''period'' is misleading in the EMD analysis of non-stationary time series. Each IMF is likely to exhibit both amplitude and frequency modulation. We therefore use the term ''average period'' with this caveat.

Table 3 is a summary of the results of applying EMD to a subset of 595 global rivers with 30 or more years of annual flows. A minimum of 30 years was chosen to ensure that there was sufficient data for the shorter record lengths to adequately define at least the first two IMFs. The table shows that approximately 43% of the variance is accounted for by IMF1 which has an average period of 3.12 years. IMF2 accounts for a further 19% of the variance, thus the intra-decadal component defined here as the IMFs with periods less than 10 years accounts for 60% of the variance. The remaining variance is contained in the inter-decadal component, of which a significant proportion of the variance is within the residual or trend component. Results of our intra- and inter-decadal component analysis are consistent with Dettinger and Diaz (2000) who, using a high pass filter with a seven year cutoff, found that 61% of the variability of annual streamflow was due to interannual (<7 years frequency) fluctuations.

Fig. 13 shows that the ratio of variance due to inter-decadal (low frequencies on the annual time scale) fluctuations relative to total variance is directly correlated to lag-one serial correlation (based on WLS $R^2 = 0.45$, $p < 0.001$). As the proportion of total variance due to inter-decadal fluctuations decreases, the lag-one serial correlation decreases and tends toward zero (random). Conversely, the lag-one serial correlation increases as the proportion of total vari-

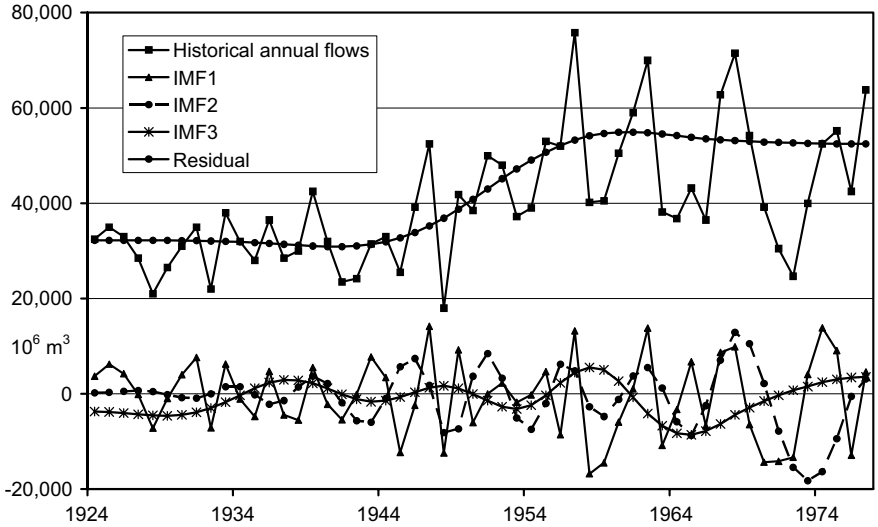

Figure 12 EMD analysis of annual streamflow (1924—1977) for Zambesi River at Victoria Falls showing the observed data, the three IMFs and the residual trend.

Table 2 Average period and variance of IMFs for Zambesi River at Victoria Falls

	Average period (years)	Variance (as ratio of sample variance)	Variance as % of ∑IMFs + residual
IMF1	3.6	0.384	31
IMF2	6.8	0.217	17
IMF3	15.4	0.067	5
Residual		0.579	47
∑IMFs + residual		1.247	100

Table 3 Mean average period and mean variance for rivers from global data set with 30 years or more of annual streamflows

Item (number of rivers)	Average period (years)	Variance (as ratio of sample variance)	Variance as % of ∑IMFs + residual
IMF1 (595)	3.12	0.592	43.3
IMF2 (594)	7.13	0.256	18.7
IMF3 (426)	15.5	0.162	11.9
IMF4 (85)	29.6	0.109	8.0
IMF5 (5)	53.8	0.038	2.8
Residual		0.209	15.3
∑IMFs + residual		1.366	100.0

ance due to inter-decadal fluctuations and the trend component increases. For the plotted data in Fig. 13, there is a negligible difference between the ASA rivers and those for the RoW (the separate relationships are not shown in the figure).

It was observed earlier that auto-correlation is only weakly correlated with catchment area, but as observed in Fig. 13, auto-correlation is strongly correlated with the ratio of inter-decadal fluctuations to total variance ($R^2 = 0.45$, $p < 0.001$). Thus large values of ρ are associated with large values of the ratio which, in turn, implies that a large trend component or strong inter-decadal fluctuation is associated with high positive auto-correlations. The cross-correlations in Table 4 among these variables for 120 global rivers (those with 30 or more annual flows and statistically significant positive auto-correlations) further support these observations.

Low flow run length, magnitude and severity

Low flows in the context of this paper relate to values below the median of a historical series. This discussion is based on an analysis of the annual time-series for the 1221 global rivers. We have purposely not used the term drought to refer to low flow sequences, as it has been argued elsewhere (McMahon and Finlayson, 2003) that drought is a construct of society, so we will continue in this paper to use the term low flows for a series of streamflows below the median record value. For each river we computed the frequency of run lengths below the median, as well as the run magnitude and run severity.

An illustration of run length is presented for the Zambesi River at Victoria Falls in Fig. 14. Fig. 14a shows the time series and Fig. 14b plots the resulting frequency of run lengths equal to or below the median for the same river. In understanding run lengths, the shape of the run length frequency distribution is important and can be described by a skewness metric as (Peel et al., 2004b)

$$g = \frac{\sum_{i=1}^{L} f_i \ell_i^3}{N-1} \tag{11}$$

where g is a metric of skewness of the run length frequency diagram, f_i is the frequency of run length ℓ_i, L is the longest

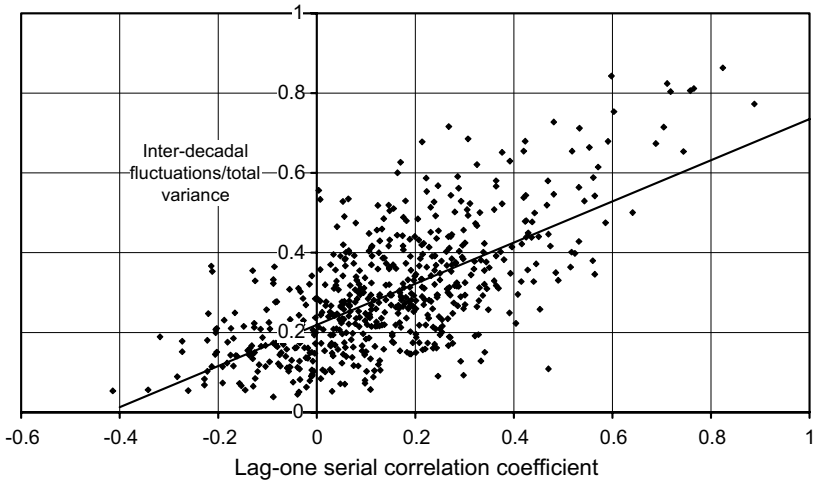

Figure 13 Ratio of inter-decadal variance to total variance calculated using EMD analysis versus annual lag-one serial correlation coefficient of the record.

Table 4 Cross-correlation matrix (as R^2) of pairs of variables: auto-correlation, ratio of intra-decadal fluctuations to total variance, record length and catchment area

	Ratio	*N*	Area
ρ	0.199	0.067	0.025
Ratio		0.011	0.004
N			0.0001

observed run length and *N* is the length of historical data. For the Zambesi River annual flows, $g = 68.0$.

To assess whether a record contains run lengths that are longer or shorter than expected we follow the procedure of Peel et al. (2004b) in which the run-frequency relationship in the historical series is compared with that for a first-order linear autoregressive, AR(1), model. The AR(1) model is a useful comparator as 91% of world rivers were found (in the analysis of 720 rivers by McMahon et al. (1992)) to be either white noise (77%) or AR(1) (16%). Furthermore, AR(1) does not capture long-term persistence in the form of decadal or multi-decadal fluctuations in runoff records (Thyer and Kuczera, 2000).

To test whether the run length behaviour at each station was different from an AR(1) model the run length skewness (g) of the frequency of run lengths was computed for each station along with the 90% confidence interval of g, based on the record length and lag-one serial correlation at each station following Peel et al. (2004b). A total of 170 stations (14% of the global data set) have g values significantly different, at the 10% level of significance, from an AR(1) model, which is similar to the 13% found in Peel et al. (2004b) using a slightly larger and earlier form of the current dataset. Globally and for ASA and RoW the proportion of stations exhibiting non-AR(1) behaviour is not statistically significantly different from the 10% expected. Like Peel et al. (2004b), the proportion of stations exhibiting longer runs below the median (9%) than expected from an AR(1) model was greater than the proportion of stations exhibiting shorter runs below the median (5%). This analysis supports the hypothesis that the AR(1) model adequately represents the range of lag-one serial correlations observed worldwide.

To compute run magnitude we use a simple metric defined as follows (Peel et al., 2005):

$$\mathrm{Rm} = \frac{\sum_{i=1}^{n_j} \frac{M_{ij}}{n_j}}{\tilde{q}} \qquad (12)$$

where Rm is the relative magnitude and is equal to the sum of the deficits below the median of the historical annual streamflows for each run length divided by the run length, n_j is the number of runs of length j, M_{ij} is the ith run magnitude of run length j where $i = 1,2, \ldots, n_j$ and $\tilde{q}$ is the sample median.

Fig. 15 shows the relative magnitude Rm of streamflow deficits versus run length, averaged separately for the ASA and RoW data sets. Two features are clear from the figure. First, for both sets of data, the relative magnitudes increase slowly with run length. This means that for reservoirs with long critical periods, storages will need to be relatively larger than those with shorter critical periods. Second, the relative magnitude of ASA rivers is approximately double the rivers in the RoW. This feature is consistent with our earlier observation that the annual variability of ASA rivers is more than double the variability of RoW rivers.

Run severity is the third characteristic we examine in this section. Here, severity of a streamflow record is defined as the product of run length and magnitude. A neat measure of this characteristic is the flow deficit relative to the mean for a given recurrence interval. It is a specific property of a time series and is equivalent to estimating the size of a reservoir to meet a specific target draft. Briefly, Extended Deficit Analysis estimates the deficit for a given recurrence interval of providing a given draft at a given level of reliability (Pegram, 2000). Details of the technique which is set out in McMahon and Adeloye (2005) were slightly modified; details of this modification are given in McMahon et al. (2007c).

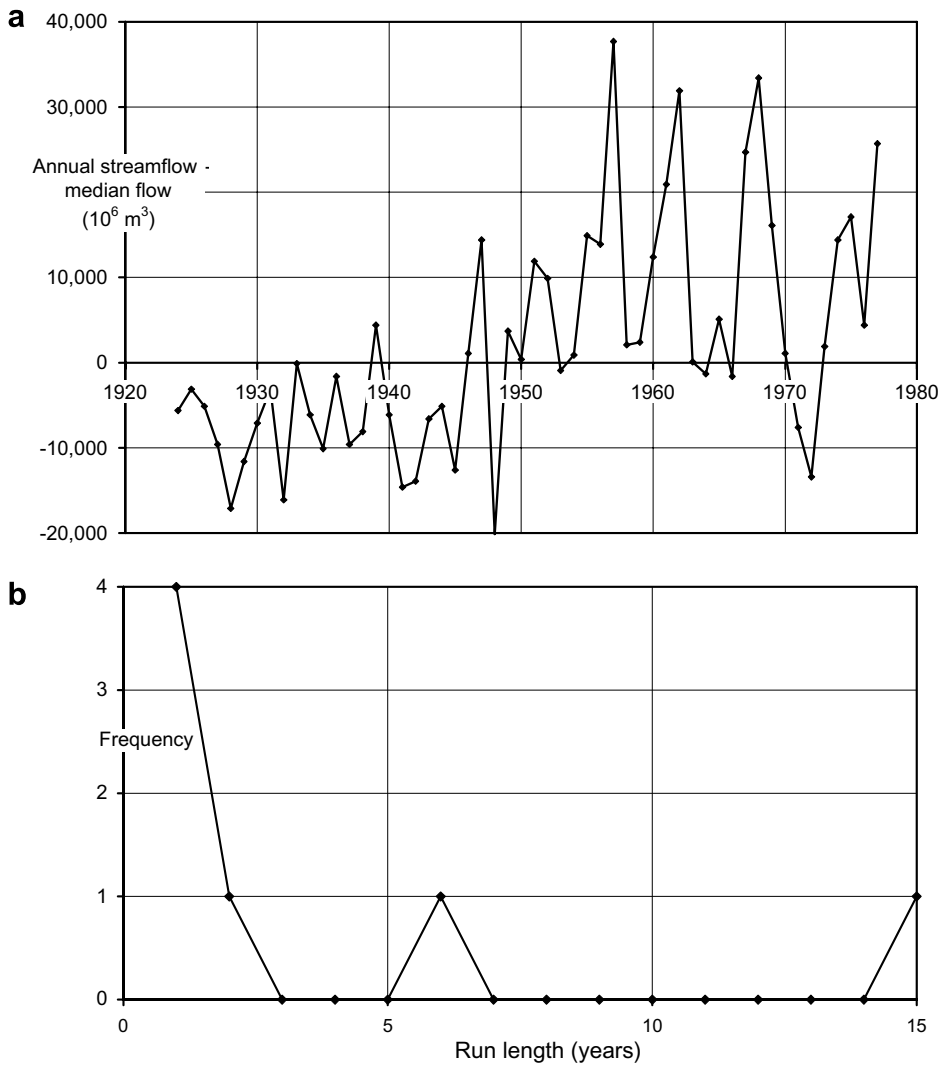

Figure 14 (a) Time series of annual streamflow relative to the median for Zambesi River at Victoria Falls. (b) Frequency distribution of run lengths equal to or below median (N = 54 years) for Zambesi River at Victoria Falls.

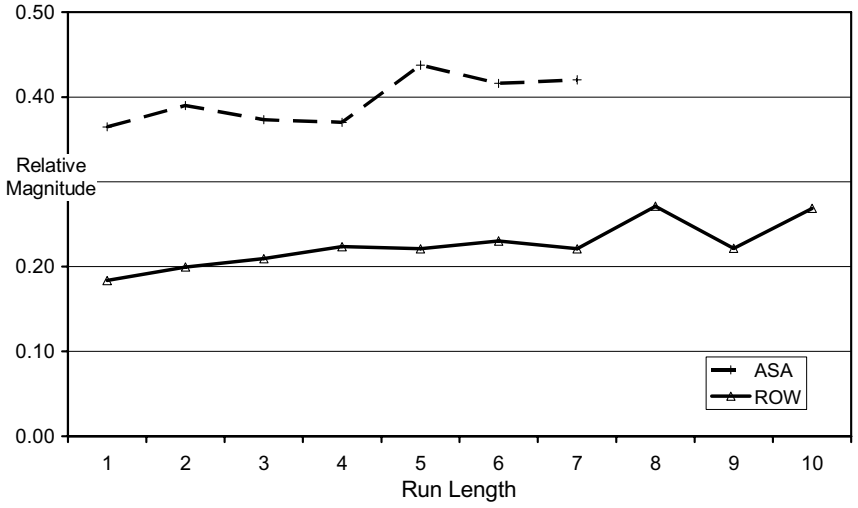

Figure 15 Relative magnitude versus run length for Australia—southern Africa and rest of world. (N = 1221, with minimum of 10 rivers per run length).

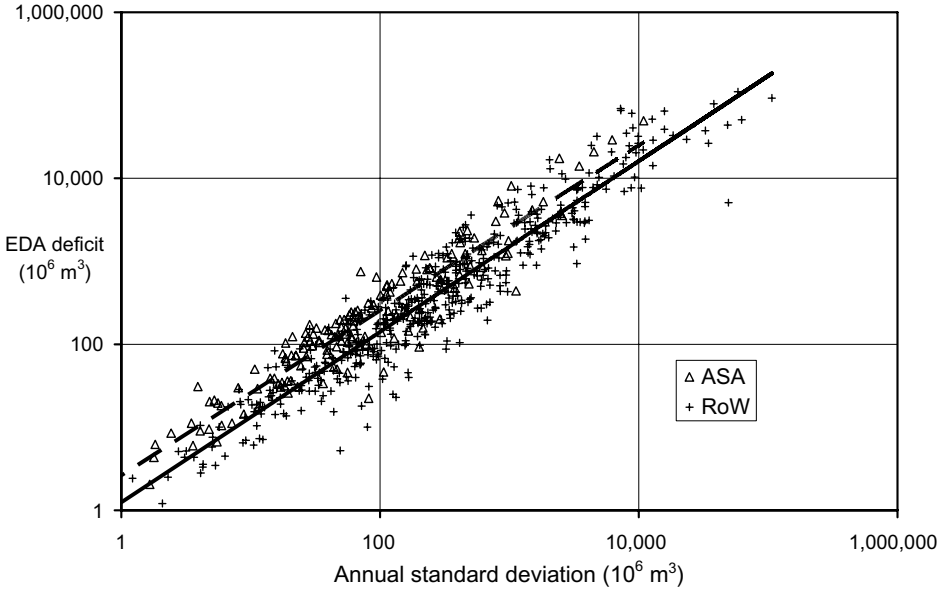

Figure 16 Extended Deficit values (estimates of storage required to provide a relative target draft of 75% with 99% annual reliability) versus standard deviation of annual flows. Solid line = RoW, broken line = ASA.

The importance of the standard deviation of flows in such an analysis is illustrated dramatically in Fig. 16 where deficits computed by EDA (for 99% reliability and 75% draft ratio) are plotted as a function of the standard deviation of annual flow. The equation representing this relationship fit using WLS is

$$\text{Def} = 1.746\sigma^{0.986} \quad (13)$$

$$n = 641,\ R^2 = 0.892,\ \text{se} = +106\%,\ -51\%, p =< 0.001$$

where Def is the deficit (10^6 m^3) and σ is the standard deviation of annual streamflows (10^6 m^3).

We note once again the clear difference between the rivers of Australia–southern Africa and those for the rest of the world. For an annual $\sigma = 1000 \times 10^6$ m^3, the deficit in ASA is 67% larger than the equivalent deficit for RoW rivers.

In summary, there were no differences in terms of run length between ASA and ROW, while there were significant differences in terms of run magnitude and severity. Peel et al. (2005) showed that run magnitude and, consequently, run severity are strongly related to interannual variability (C_v). The ASA and RoW difference in run magnitude and severity observed here is the result of differences in the interannual variability of streamflows observed earlier.

Summary and conclusions

As a result of the analyses described in this paper we offer the following points.

1. Although there are many reports dealing with water balance studies of large catchments worldwide including their distribution by country and climate, there are few reports or papers that address both the regulated and the unregulated characteristics of annual streamflows of global rivers. This is Part 1 of three papers that considers these issues.
2. The data set of annual and monthly flows for 1221 rivers, adopted in this project, is unique in that considerable effort has been made to ensure the data are free of major errors and the streamflows are not impacted by major water withdrawals nor by reservoirs upstream.
3. Based on a theoretical analysis of the relationship between product-moments and *L*-moments associated with the measures of C_v and skewness, it was shown that the Gamma probability distribution function (pdf) appears to fit the annual streamflows better than the Lognormal pdf for the 1221 rivers. This result is consistent with other recent studies using large datasets from very heterogeneous regions (Vogel and Wilson, 1996).
4. Useful global relationships were established among the following variables: mean annual flow, mean annual runoff, standard deviation of annual flows, coefficient of variation of annual flows and catchment area. The differences in annual streamflow characteristics between Australia–southern Africa and the rest of the world are highlighted in the analysis.
5. Based on the approximate Gould–Dincer Gamma reservoir capacity yield model which considers only over-year storage, we found that reservoirs in Australia–southern Africa need to be roughly twice the capacity of those in the rest of the world assuming a common catchment size (10,000 km^2) and the same draft and failure conditions.
6. Our analysis suggests that ρ is inversely related to record length and that it is positively, but weakly, related to catchment area.

7. It is postulated that the statistically significant negative auto-correlation observed in 11 of the 1221 rivers in the data set is the result of sample size and sampling variability rather than some physical explanation. The magnitude of these values was strongly negatively correlated with record length.
8. Low frequency oscillations in the annual flow records were studied by Empirical Mode Decomposition. Our analysis showed that 60% of the variability in annual streamflow is accounted for by intra-decadal variability. The remaining variance is contained in the inter-decadal component, of which a significant proportion of the variance (more than one-third) is within the residual or trend component. It was also observed that as the proportion of total variance due to intra-decadal (high frequencies on the annual time scale) fluctuations increases, the lag-one serial correlation decreases and tends toward zero (random).
9. For 14% of rivers, the run length behaviour was different, at the 10% level of significance, to that expected from an AR(1) process and there was no difference between Australia—southern Africa and the rest of the world.
10. Our analysis of the relative magnitude of streamflow deficits shows that relative magnitude increases with run length and that the relative magnitudes of Australia—southern Africa rivers are approximately double that for rivers in the rest of the world.
11. Run severity which is defined as the product of run length and magnitude was estimated using Extended Deficit Analysis. For an equivalent hypothetical river and given streamflow variability, the deficit in flow in Australia—southern Africa is approximately 67% more than for a similar river from the rest of the world.

Acknowledgement

The authors are grateful to the Department of Civil and Environmental Engineering, the University of Melbourne and the Australian Research Council Grant DP0449685 for financially supporting this research. Our original streamflow data set was enhanced by additional data from the Global Runoff Data Centre (GRDC) in Koblenz, Germany. Streamflow data for Taiwan and New Zealand were also provided by Dr. Tom Piechota of the University of Nevada, Las Vegas. Professor Ernesto Brown of the Universidad de Chile, Santiago kindly made available Chilean streamflows.

We are also grateful to Dr. Senlin Zhou of the Murray-Darling Basin Commission who completed early drafts of the computer programs used in the analysis. Two anonymous reviewers provided very helpful comments.

Appendix A. Frequency distribution of annual streamflows

The following describes the methodology for plotting theoretical relations between L-C_v and C_v and between L-Skewness and Skewness for the Lognormal and Gamma probability distribution functions (pdfs). Here the annual flows are denoted X and their logarithms are denoted $Y = \ln(X)$

Lognormal case

L-C_v versus C_v

Stedinger et al. (1993, Eq. 18.2.13) report the first and second L-moments for the Lognormal pdf as a function of the distribution parameters

$$\lambda_1 = \exp\left(\mu_y + \frac{\sigma_y^2}{2}\right) \tag{A1}$$

$$\lambda_2 = \exp\left(\mu_y + \frac{\sigma_y^2}{2}\right) erf\left(\frac{\sigma_y}{2}\right) = 2\exp\left(\mu_y + \frac{\sigma_y^2}{2}\right)\left[\Phi\left(\frac{\sigma_y}{\sqrt{2}}\right) - \frac{1}{2}\right] \tag{A2}$$

where μ_y and σ_y^2 are the mean and variance of the logarithms of the annual flows.

$erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)dt$ and $\Phi(x)$ is the cumulative probability density function of a standardized normal variate. For a Lognormal variable the coefficient of variation of X, denoted C_v, is given by

$$C_v = \sqrt{\exp(\sigma_y^2) - 1} \tag{A3}$$

Now, we obtain the L-C_v as a function of the ordinary product moment C_v by combining Eqs. (A1), (A2) and (A3) with the definition $L\text{-}C_v = \lambda_2/\lambda_1$,

$$L\text{-}C_v = erf\left[\frac{\sqrt{\ln(1+C_v^2)}}{2}\right] = 2\left[\Phi\left(\frac{\sqrt{\ln(1+C_v^2)}}{\sqrt{2}}\right) - \frac{1}{2}\right] \tag{A4}$$

L-Skew versus Skewness

There are no simple expressions for L-moment ratios above L-C_v for the Lognormal pdf, however, Hosking and Wallis (1997) provide the following approximation which has accuracy better than 2×10^{-7} for $|k| \leqslant 4$, corresponding to $|L\text{-Skew}| \leqslant 0.99$

$$L\text{-Skew} = \frac{A_0 + A_1k^2 + A_2k^4 + A_3k^6}{1 + B_1k^2 + B_2k^4 + B_3k^6} \tag{A5}$$

where $k = -\sigma_y$. The values of the coefficient A_i and B_i in Eq. (A5) are reproduced in Table A1.

The ordinary product moment skewness γ for a 2-parameter Lognormal variable is given by

$$\gamma = 3C_v + C_v^3 \tag{A6}$$

The ordinary skewness γ in Eq. (A6) can be related to the L-Skewness in Eq. (A5) using the fact that

$$-k = \sqrt{\ln(1+C_v^2)} \tag{A7}$$

The relationship between L-Skewness and γ is obtained by the simultaneous solution of Eqs. (A5), (A6) and (A7) using a numerical root-finding algorithm.

Table A1 Coefficients of the approximation in Eq. (A5), from Hosking and Wallis (1997, page 199)

Coefficient	Value
A_0	4.8860251×10^{-1}
A_1	4.4493076×10^{-3}
A_2	8.8027039×10^{-4}
A_3	1.1507084×10^{-6}
B_1	6.4662924×10^{-2}
B_2	3.3090406×10^{-3}
B_3	7.4290680×10^{-5}

Gamma case

L-C_v versus C_v

Stedinger et al. (1993) report the first and second *L*-moments for the Pearson type III pdf as a function of its parameters as

$$\lambda_1 = \zeta + \frac{\alpha}{\beta} \tag{A8}$$

$$\lambda_2 = \frac{\Gamma(\alpha + 0.5)}{\beta\Gamma(\alpha)\sqrt{\pi}} \tag{A9}$$

where $\Gamma()$ is the Gamma function. A Gamma pdf is simply a two parameter form of the Pearson type III pdf, when the lower bound parameter $\zeta = 0$. Stedinger et al. (1993) also report that the ordinary product moment skew is twice the coefficient of variation and is related to the model parameter α as follows

$$\gamma = 2C_v = \frac{2}{\sqrt{\alpha}} \tag{A10}$$

Combining Eq. (8) with $\zeta = 0$, Eqs. (9) and (10) with the definition of *L*-C_v we obtain a relationship between *L*-C_v and ordinary C_v:

$$L\text{-}C_v = \frac{C_v^2 \cdot \Gamma\left(\frac{1}{C_v^2} + \frac{1}{2}\right)}{\Gamma\left(\frac{1}{C_v^2}\right)\sqrt{\pi}} \tag{A11}$$

L-Skew versus Skewness

Hosking and Wallis (1997, Eqn. A.84) report an expression for the *L*-Skewness of a Pearson type III variable as

$$L\text{-Skew} = 6I_{1/3}(\alpha, 2\alpha) - 3 \tag{A12}$$

where $I_x(p,q)$ denotes the Incomplete Beta function defined as

$$I_x(p,q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int t^{p-1}(1-t)^{q-1}\mathrm{d}t \tag{A13}$$

and $0 > \alpha > \infty$. Combining Eq. (A12) with (A10) yields the relationship between *L*-Skewness and Skewness for a Gamma pdf:

$$L\text{-Skew} = 6I_{1/3}\left(\frac{4}{\gamma^2}, \frac{8}{\gamma^2}\right) - 3 \tag{A14}$$

References

Ayushinskaya, N.M., Voskresenskiy, K.P., Grigorkina, T.Y., Korzell, A.G., Markova, O.L., Rybkina, A.Y., Sokolov, A.A., 1977. Global river runoff. Soviet Hydrology Papers Selected Papers 16 (2), 127–131.

Baumgartner, A., Reichel, E., 1975. The world water balanceMean Annual Global, Continental and Marine Precipitation, Evaporation and Runoff. Elsevier, Amsterdam.

Dettinger, M.D., Diaz, H.F., 2000. Global characteristics of stream flow seasonality and variability. Journal of Hydrometeorology 1, 289–310.

Haines, A.T., Finlayson, B.L., McMahon, T.A., 1988. A global classification of river regimes. Applied Geography 8, 255–272.

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis: an Approach Based on *L*-Moments. Cambridge University Press.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society London A 454 (1971), 903–995.

Kalinin, G.P., 1971. Global Hydrology. Israel program of Scientific Translations, Jerusalem.

Klemes, V., 1970. Negatively skewed distribution of runoff, International Association of Scientific Hydrology, Publication 96, 219-236.

Köppen, W., 1936. Das geographisca System der Klimate. In: Köppen, W., Geiger, G. (Eds.), Handbuch der Klimatologie. 1. C. Gebr, Borntraeger, pp. 1–44.

Korzun, V.I., Sokolov, A.A., Budyko, M.I., Voskresenhsky, K.P., Kalinin, G.P., Konoplyantsev, A.A., Korotkevich, E.S. and Lvovich, M.I., 1974. *Atlas of World Water Balance* (annex to monograph World Water Balance and Water Resources of the Earth) USSR Committee for IHD, Hydrometeorological Publishing House, Moscow.

Löf, G.O.G., Hardison, C.H., 1966. Storage requirements for water in the United States. Water Resources Research 2 (3), 323–354.

Markovic, R.D. 1965. Probability functions of best fit to distributions of annual precipitation and runoff. Colorado State University, Hydrology Paper 8, Colorado State University.

McMahon, T.A., 1975. Variability, persistence and yield of Australian streams. Hydrology Symposium, Institution of Engineers, Australia, National Conference Publication 75 (3), 107–111.

McMahon, T.A., 1977. Some statistical characteristics of annual streamflows in northern Australia. Hydrology Symposium, The Hydrology of Northern Australia, Institution of Engineers, Australia, National Conference Publication 77 (5), 131–135.

McMahon, T.A., 1979. Hydrological characteristics of arid zones. International Association of Hydrological Sciences Publication 128, 105–123.

McMahon, T.A., Adeloye, A., 2005. Water Resources Yield. Water Resources Publication, LLC, Colorado, USA.

McMahon, T.A., Finlayson, B.L., 2003. Droughts and anti-droughts: the low flow hydrology of Australian rivers. Freshwater Biology 48, 1147–1160.

McMahon, T.A., Finlayson, B.L., Haines, A., Srikanthan, R., 1992. Global runoff: continental comparisons of annual flows and peak discharges. CATENA VERLAG, Germany.

McMahon, T.A., Vogel, R.M., Pegram, G.G.S., Peel, M.C., Etkin, D., 2007a. Global streamflows – Part 2 Reservoir storage-yield performance. Journal of Hydrology 347 (3–4), 260–271.

McMahon, T.A., Peel, M.C., Vogel, R.M., Pegram, G.G.S., 2007b. Global streamflows – Part 3 Country and climate zone characteristics based on global historical streamflow data. Journal of Hydrology 347 (3–4), 272–291.

McMahon, T.A., Pegram, G.G.S., Vogel, R.M., Peel, M.C., 2007c. Revisiting reservoir storage-yield relationships using a global streamflow database. Advances in Water Resources 30, 1858—1872.

Peel, M.C., McMahon, T.A., Finlayson, B.L., Watson, F.G.R., 2001. Identification and explanation of continental differences in the variability of annual runoff. Journal of Hydrology 250, 224—240.

Peel, M., McMahon, T.A., Finlayson, B.F., Watson, F., 2002. Implications of the relationship between catchment vegetation type and annual runoff variability. Hydrological Processes 16, 2995—3002.

Peel, M.C., McMahon, T.A., Finlayson, B.L., 2004a. Continental differences in the variability of annual runoff — update and reassessment. Journal of Hydrology 295, 185—197.

Peel, M.C., Pegram, G.G.S., McMahon, T.A., 2004b. Global analysis of runs of annual precipitation and runoff equal to or below the median: run length. International Journal of Climatology 24, 807—822.

Peel, M.C., Pegram, G.G.S., McMahon, T.A., 2005. Global analysis of runs of annual precipitation and runoff equal to or below the median: run magnitude and severity. International Journal of Climatology 25, 549—568.

Pegram, G.G.S., 2000. Extended deficit analysis of Bloemhof and Vaal Dam inflows during the period (1920 to 1994). Report to the department of water affairs and forestry for the Vaal River continuous study, DWAF, Pretoria, South Africa.

Post, D.A., Littlewood, I.G., Croke, B.F., 2005. New directions for Top-Down modelling: Introducing the PUB Top-Down Modelling Working Group. International Association of Hydrological Sciences 301, 125—133.

Shiklomanov, I.A., Rodda, J.C., 2003. World Water Resources at the Beginning of the 21st Century. Cambridge University Press, New York.

Sivapalan, M., 2005. Pattern, process and function: elements of a unified theory of hydrology at the catchment scale. In: Anderson, M.G. (Ed.), Encyclopedia of Hydrological Sciences, vol. 1. John Wiley & Sons, pp. 193—219 (Chapter 13).

Stedinger, J.R., Vogel, R.M., and Foufoula-Georgiou, E., 1993. Frequency Analysis of Extreme Events, Chapter 18, Handbook of Hydrology, M^{c}Graw-Hill Book Company, David R. Maidment, Editor-in-Chief, 18.1—18.66.

Thyer, M., Kuczera, G., 2000. Modeling long-term persistence in hydroclimatic time series using a hidden state Markov model. Water Resources Research 36 (11), 3301—3310.

UNESCO, 1977. Atlas of World Water Balance. UNESCO.

UNESCO 1978. World water balance and water resources of the earth. Studies and Reports in Hydrology 25.

Van Der Leedon, F., 1975. Water Resources of the World: Selected Statistics. Water Information Center Inc., New York.

Vogel, R.M., Fennessey, N.M., 1993. *l*-moment diagrams should replace product-moment diagrams. Water Resources Research 29 (6), 1745—1752.

Vogel, R.M., Wilson, I., 1996. Probability distribution of annual maximum mean and minimum streamflows in the United States. Journal of hydrologic Engineering 1 (2), 69—76.

Yevdjevich, V.M., 1963. Fluctuations of wet and dry years. Part 1. Research data assembly and mathematical models. Colorado State University, Hydrology Paper 1, Colorado State University.

Yevdjevich, V.M. 1964. Fluctuations of wet and dry years. Part 2. Analysis by serial correlation. Colorado State University, Hydrology Paper 4, Colorado State University.