Stochastic assessment of Phien generalized reservoir storage–yield–probability models using global runoff data records

CrossMark

Adebayo J. Adeloye [*], Bankaru-Swamy Soundharajan, Jagarkhin N. Musto, Chuthamat Chiamsathit

Institute for Infrastructure and Environment, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

SUMMARY

This study has carried out an assessment of Phien generalised storage–yield–probability (S–Y–P) models using recorded runoff data of six global rivers that were carefully selected such that they satisfy the criteria specified for the models. Using stochastic hydrology, 2000 replicates of the historic records were generated and used to drive the sequent peak algorithm (SPA) for estimating capacity of hypothetical reservoirs at the respective sites. The resulting ensembles of reservoir capacity estimates were then analysed to determine the mean, standard deviation and quantiles, which were then compared with corresponding estimates produced by the Phien models. The results showed that Phien models produced a mix of significant under- and over-predictions of the mean and standard deviation of capacity, with the under-prediction situations occurring as the level of development reduces. On the other hand, consistent over-prediction was obtained for full regulation for all the rivers analysed. The biases in the reservoir capacity quantiles were equally high, implying that the limitations of the Phien models affect the entire distribution function of reservoir capacity. Due to very high values of these errors, it is recommended that the Phien relationships should be avoided for reservoir planning.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The determination of reservoir capacity to meet the demand with an acceptable level of satisfaction (or reliability) is an age-old problem and there currently exist many techniques for accomplishing this task as documented by McMahon and Adeloye (2005), Adeloye (2012). Most of these techniques such as behaviour simulation, the mass curve and the sequent peak algorithm are sequential, involving the analysis of time-series runoff data at the reservoir site (Adeloye, 2012). Such sequential techniques are generally preferred because they automatically take into account the effect of runoff characteristics-mean, coefficient of variation (CV = mean/standard deviation), serial dependence, skewness-on capacity estimates. However, sequential methods are sometimes infeasible, e.g. when the site is ungauged or the available record is too short that using them will result in significant uncertainty in the estimated capacity (Adeloye, 1990, 1996), or unwarranted, e.g. during preliminary analysis to screen potential reservoir sites. For these situations, generalised storage–yield–probability (S–Y–P) relationships offer a way out and there are numerous examples of such applications as recently reviewed by Kuria and Vogel (2015).

Generalised relationships relate the storage capacity to the demand and runoff summary statistics, most of which can be indirectly estimated from easily measureable catchment characteristics (see Adeloye et al., 2003), thus making them applicable to ungauged or poorly gauged catchments.

Several generalised storage–yield relationships have been reported in the literature including Vogel and Stedinger (1987), Burchberger and Maidement (1989), Bayazit and Bulu (1991), Adeloye et al. (2003), Bayazit and Onoz (2000), Phien (1993), Silva and Portela, 2012, Kuria and Vogel (2015) and McMahon et al. (2007a). Apart from few exceptions that used recorded data (e.g. McMahon et al., 2007a; Adeloye et al., 2003), a common feature of most of the existing generalised relationships is that they have been developed using runoff data sampled stochastically, i.e. a distribution hypothesis of the runoff is first assumed, then plausible statistics of the runoff (mean, CV, and serial dependence) are assumed and used to generate large replicates of runoff data for typical record lengths commonly encountered in practice. Capacity estimates are then obtained by routing the runoff record through a hypothetical reservoir using a suitable reservoir planning technique such as the sequent peak algorithm, SPA (see Loucks et al., 1981). The resulting capacity estimates are then summarised in terms of the mean and standard deviation of reservoir capacity, two of the three statistical parameters required to fully specify

* Corresponding author. Tel.: +44 131 451 8236.
  E-mail address: a.j.adeloye@hw.ac.uk (A.J. Adeloye).

the 3-parameter log-normal distribution often assumed for reservoir capacity (Vogel and Stedinger, 1987; Bayazit and Bulu, 1991).

To produce the generalised models, the mean capacity and standard deviation of capacity are independently related to the demand and runoff characteristics usually through regression. The resulting calibrated regression relationships can then provide estimates of the mean, standard deviation and ultimately quantiles of reservoir capacity without the need for Monte Carlo simulation. This would be a welcome development if estimates of the quantiles and other reservoir statistics provided by these generalised models compare favourably with similar estimates obtained using Monte Carlo simulation analyses with observed runoff data. However, very few attempts have been made to assess the performance of the existing generalised models using observed runoff data and because of this, there is very little guidance on the bias of the commonly used generalised methods. However, having this information is necessary if the models are to engender confidence in their adoption for reservoir capacity planning analyses.

McMahon et al. (2007b) analysed a large number of global runoff records to assess the performance of some of the existing generalised relationships, including Phien's, and is one of only few studies as far as we are aware that have undertaken this task. However, the study was limited in that it used single historic records and was thus limited to the mean capacity estimate, which is not sufficient for fully specifying the probability distribution function of capacity and hence obtaining the quantile estimates. Nonetheless, they found significant biases with both the Vogel–Stedinger (V–S) and Phien models in predicting the mean of capacity. Additionally, their treatment of the Phien models did not fully satisfy the criteria specified in the original development of the models, especially that relating to the distribution hypothesis for the annual runoff which was that the annual runoff should have a gamma distribution. For example, although they tried to establish that some of the data records followed the gamma distribution

using L-moments diagrams, such a "global" goodness-of-fit approach falls short of establishing that each of the records behaved as gamma. Finally, McMahon et al. (2007b) only examined one of the four Phien models (the one which will be referred to later on in this paper as the generic-model); there is therefore no guidance on the remaining three models.

The assessment carried out by Adeloye et al. (2010) focused on the Vogel–Stedinger (V–S) generalised storage–yield model within a Monte Carlo framework and using runoff of three global rivers. They found that the V–S model significantly over-estimates the reservoir capacity especially at high demands where the bias can be as much as 140%. The V–S model assumes that the annual runoff exhibits a normal/log-normal distribution; consequently it is not straightforward to infer the bias of generalised models of gamma-inflow-fed reservoirs from the V–S situation.

As implied above, the Phien (1993) models are unique in that unlike other generalised approaches that assume that the annual runoff is normally or log-normally distributed, the runoff is assumed to be distributed as gamma in the development of the Phien models. As noted by McMahon et al. (2007c,d), many world rivers cannot justifiably be modelled using the normal distribution and the gamma is a more plausible distribution hypothesis to use in such situations. It is therefore important that a complete assessment of the Phien models is carried out and reported.

The aim of this work therefore is to carry out independent assessment of the Phien (1993) models using runoff records that meet the gamma distribution hypothesis. The associated tasks include:

i. Assembling rivers annual runoff data records that exhibit gamma distribution and estimating their summary statistics. As will be seen later, six such records were assembled and used in the study.

Table 1
Details of Phien (1993) models.

Model number	m (see Eq. (8))	Mean of scaled capacity (μ_v)	Standard deviation of scaled capacity (σ_v)
1	0.0 (full regulation)	$0.97n^{0.59}\left[\frac{(1+\rho)}{(1-\rho)}\right]^{0.42}$	$0.69n^{0.55}\left[\frac{(1+\rho)}{(1-\rho)}\right]^{0.52}$
2	0.25	$0.67n^{0.41}\left[\frac{(1+\rho)}{(1-\rho)}\right]^{0.5}$	$0.85n^{0.16}\left[\frac{(1+\rho)}{(1-\rho)}\right]^{0.89}$
3	0.5	$0.18n^{0.4}\left[\frac{(1+\rho)}{(1-\rho)}\right]^{0.67}$	$0.24n^{0.07}\left[\frac{(1+\rho)}{(1-\rho)}\right]^{1.15}$
4	Generic (m-independent)	$1.467n^{0.466}\left[\frac{(1+\rho)}{(1-\rho)}\right]^{0.531}\left[\frac{(1-m)}{(1+m)}\right]^{2.047}$	$1.787n^{0.243}\left[\frac{(1+\rho)}{(1-\rho)}\right]^{0.855}\left[\frac{(1-m)}{(1+m)}\right]^{2.198}$

Table 2
Details of the catchments used.

	Rivers					
	Homochitto	Mareetsane	Namsen	Nariel	Onkaparinga	Beas
Country	USA	South Africa	Norway	Australia	Australia	India
Gauging station	Eddiceton	Neverset	Bertnem	Upper Nariel	Clarendon Weir	Pong Dam
Latitude	31.50	−26.12	64.47	−36.45	−35.12	32.1
Longitude	−90.78	25.28	12.07	147.83	138.63	77.14
Catchment area (km^2)	466.2	566	5177	252	445	12,561
Record length (years)	46	37	23	38	69	15
Mean annual flow, μ_Q ($\times 10^6$ m^3)	238.164	3.378	7674.957	139.859	81.468	8485.173
Serial correlation coefficient (ρ)	0.057	0.112	0.032	−0.042	−0.01	0.323
CV	0.395	1.012	0.185	0.491	0.684	0.225
95% CONF. of skew = 2CV	[0.104, 1.476]	[1.264, 2.783]	[−0.574, 1.313]	[0.232, 1.732]	[0.801, 1.933]	[−0.687, 1.587]
Sample skew, Υ_Q	0.6926	1.8680	0.3908	0.8983	1.1836	−0.0933

Table 3
Statistics of the observed and simulated runoff data records (the simulated statistics are the mean of 2000 replicates).

River	μ_Q ($\times 10^6$ m^3)		CV		Skewness (Υ_Q)		Serial correlation coefficient (ρ)	
	Simulated	Observed	Simulated	Observed	Simulated	Observed	Simulated	Observed
Homochitto	237.99	238.16	0.3941	0.395	0.6093	0.693	0.0312	0.057
Mareetsane	3.3814	3.38	0.9655	1.012	1.4926	1.868	0.0979	0.112
Namsen	7737.3	7674.96	0.1809	0.185	0.2215	0.391	−0.0354	0.032
Nariel	140.12	139.86	0.4863	0.491	0.7565	0.898	−0.0683	−0.042
Onkaparinga	82.4946	81.47	0.6813	0.684	1.0848	1.184	−0.0359	−0.010
Beas	8316.8	8485.17	0.2175	0.225	−0.1050	−0.093	0.1735	0.323

Table 4
Results of Kolmogorov–Smirnov goodness-of-fit test for Gamma distribution.

River	Test statistic	Critical value			Remarks
		5%	2%	1%	
Homochitto	0.062	0.196	0.219	0.235	No evidence to reject the
Mareetsane	0.146	0.218	0.244	0.262	null hypothesis of gamma
Namsen	0.105	0.275	0.307	0.330	distribution at the 5%
Nariel	0.093	0.215	0.241	0.258	level
Onkaparinga	0.099	0.161	0.180	0.193	
Beas	0.154	0.338	0.377	0.404	

ii. Generating, for each of the historic records, 2000 replicates (length of each replicate = historic record length) using the lag-1 autoregressive stochastic model and assess the performance of the model in preserving the statistics of the historic records.

iii. Routeing each stochastic replicate through hypothetical reservoirs using the SPA and hence obtaining the population of reservoir capacity for various assumed demands (as ratio of the historic mean annual runoff).

iv. Analysing the population of SPA capacities to determine the mean and standard deviation of the distribution of capacity and hence reservoir capacity quantiles.

v. Using Phien models on the historic data to estimate the mean and standard deviation of the distribution of reservoir capacity and hence reservoir capacity quantiles.

vi. Comparing the results in (iv) and (v) and making recommendations.

In the following Section, further details about the Phien models will be given. This will be followed by the methodology, including the consideration that went into the selection of the runoff records analysed. The results and discussions will then be presented, followed by the main conclusions of the study.

2. Phien models

2.1. Assumptions

Phien models are based on certain assumptions, namely that the annual flows are distributed as gamma, that the serial dependence (ρ) is in the interval [0, 0.5], that the record length (n years) is in the interval [20, 50], and that the standardised net inflow parameter m (see Eq. (8)) is in the interval [0, 1.0]. While some of these are straightforward, the distribution hypothesis and the net inflow parameter require further explanations.

Table 5
m-specific Phien models and (**observed**) statistics of scaled reservoir capacity, V.

River	m = 0.0		m = 0.25		m = 0.5	
	μ_v	σ_v	μ_v	σ_v	μ_v	σ_v
Homochitto	9.74 (**7.92**)	6.01 (**2.22**)	3.41 (**4.20**)	1.74 (**1.59**)	0.90 (**2.44**)	0.36 (**0.92**)
Mareetsane	8.97 (**6.79**)	5.65 (**2.25**)	3.29 (**3.13**)	1.85 (**1.24**)	0.89 (**1.35**)	0.40 (**0.57**)
Namsen	6.33 (**5.04**)	4.00 (**1.35**)	2.50 (**2.95**)	1.48 (**1.14**)	0.66 (**1.85**)	0.32 (**0.79**)
Nariel	8.01 (**6.29**)	4.88 (**1.53**)	2.86 (**3.16**)	1.41 (**1.04**)	0.73 (**1.82**)	0.28 (**0.70**)
Onkaparinga	11.69 (**9.15**)	7.01 (**2.57**)	3.76 (**3.87**)	1.64 (**1.28**)	0.97 (**2.01**)	0.32 (**0.64**)
Beas	6.35 (**4.13**)	4.33 (**1.45**)	2.84 (**2.78**)	2.38 (**1.27**)	0.83 (**1.92**)	0.63 (**1.01**)

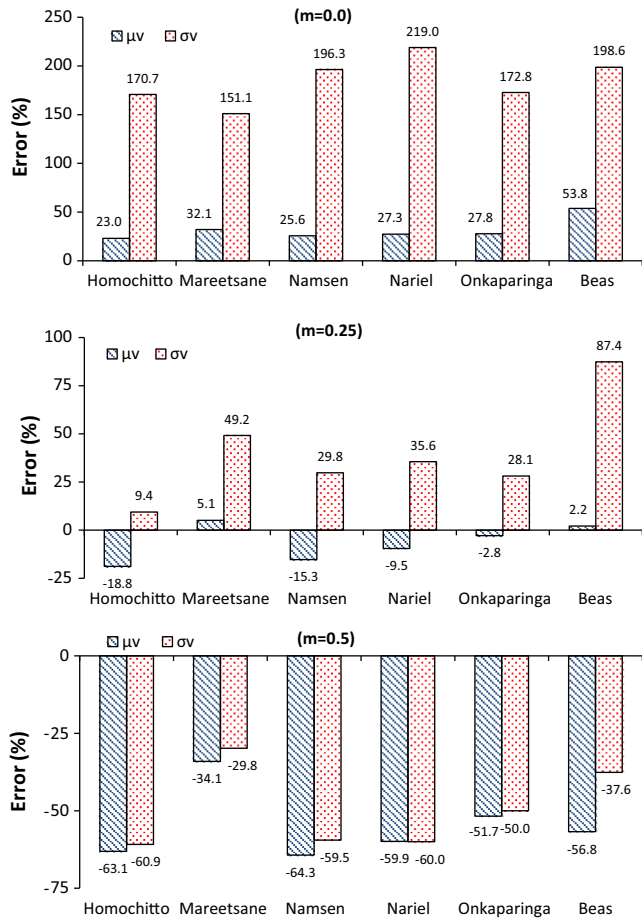The bold values indicate the observed values of the relevant statistics.

Fig. 1. Error (%) in mean and standard deviation of reservoir capacity (*m*-specific Phien models).

The probability density function (pdf) of the gamma distribution is:

$$f(x) = \frac{x^{c-1} e^{-x/\beta}}{\beta^c \Gamma(c)}, c > 0 \tag{1}$$

where $\Gamma(c)$ is the complete gamma function, β is the scale parameter and c is the shape parameter. Because the pdf in Eq. (1) has 2 parameters (β and c), this form of the gamma distribution is known as the 2-parameter gamma. The mean, variance and skewness of the distribution are given by (NERC, 1975):

$$\text{Mean} = E[x] = \mu = \beta c \tag{2}$$

$$\text{Variance} = E[x - E(x)]^2 = \sigma^2 = \beta^2 c \tag{3}$$

$$\text{Skewness} = \gamma = 2/c^{0.5} \tag{4}$$

Eqs. (2)–(4) can be manipulated further to reveal the unique relationship between the coefficient of variation (CV) and the skewness of the 2-parameter gamma function, i.e.:

$$\text{Skewness} = \gamma = 2CV \tag{5}$$

In other words, the skewness γ is twice the CV for the 2-parameter gamma distribution. While standard goodness of-fit-tests such as the Chi-squared and Kolmogorov–Smirnov (see NERC, 1975) can be carried out to ascertain the compliance with the gamma distribution, this theoretical "skew-CV" relationship of the gamma function can be used to rapidly test whether a given time series data record is distributed as gamma or not. To do this, the sample skew estimate is compared with the 95% CONF interval for the skew = 2CV. Assuming that the distribution of the skew is normal, the 95% CONF becomes:

$$95\% \text{ CONF} = [2CV - 1.96\sigma_g; 2CV + 1.96\sigma_g] \tag{6}$$

where σ_g, is the standard error of the gamma skew, whose indicative estimate is (Matalas and Benson, 1968):

$$\sigma_g = \left[\frac{6n(n-1)}{(n-2)(n+1)(n+3)}\right]^{0.5} \tag{7}$$

Both the approximate and formal Kolmogorov–Smirnov goodness-of-fit tests will be implemented in this study to establish the gamma distribution compliance for the data records.

The drift (m) was varied between 0 and 1 in Phien experiments. The drift integrates the demand and the coefficient of variation of annual runoff through (McMahon and Adeloye, 2005):

$$m = \frac{(1 - \alpha)}{CV} \tag{8}$$

where α is the demand ratio (=D/μ_Q), D is the volumetric demand, μ_Q is the mean annual runoff, and all other symbols are as defined previously. The use of m (rather than the CV) to characterise runoff variability makes it possible to analyse any runoff record with the Phien method, provided the demand ratio α under consideration results in an m value within the range of 0–1. The lower and upper bounds of m thus correspond to full regulation (i.e. $\alpha = 1$) and $\alpha = 1-$CV, respectively. The latter situation will limit the maximum value of the CV to 1, i.e. zero demand for which the capacity will be zero. Consequently, if CV > 1 situations must be considered, then the maximum m must be less than unity. Although this was not made explicitly clear in Phien's analyses, the fact that it was found that the mean of reservoir capacity was zero for $m = 0.75$ can easily be explained by the fact that cases for which the CV was above 1 must have featured in the analyses. For example, a CV of 1.33 will result in a demand ratio $\alpha = 0.0$ for $m = 0.75$, i.e. no demand and hence no need for a reservoir. The higher the m value adopted, the lower will be the CV that can result in nil demand and nil capacity. These difficulties, although not made clear by Phien (1993), must have forced Phien (1993) to restrict the maximum value of m eventually considered to 0.5.

Table 6
Generic Phien model and (**observed**) statistics of scaled reservoir capacity, *V*.

River	*m* = 0.0		*m* = 0.25		*m* = 0.5		*m* = 0.7	
	μ_v	σ_v	μ_v	σ_v	μ_v	σ_v	μ_v	σ_v
Homochitto	9.28 (**7.92**)	5.00 (**2.22**)	3.26 (**4.20**)	1.63 (**1.59**)	0.98 (**2.44**)	0.45 (**0.92**)	0.27 (**1.64**)	0.11 (**0.61**)
Mareetsane	8.89 (**6.79**)	5.21 (**2.25**)	3.12 (**3.13**)	1.69 (**1.24**)	0.94 (**1.35**)	0.47 (**0.57**)	0.26 (**0.58**)	0.11 (**0.24**)
Namsen	6.54 (**5.04**)	4.04 (**1.35**)	2.30 (**2.95**)	1.31 (**1.14**)	0.69 (**1.85**)	0.36 (**0.79**)	0.19 (**1.34**)	0.09 (**0.59**)
Nariel	7.64 (**6.29**)	4.03 (**1.53**)	2.69 (**3.16**)	1.31 (**1.04**)	0.81 (**1.82**)	0.36 (**0.70**)	0.22 (**1.23**)	0.09 (**0.45**)
Onkaparinga	10.44 (**9.15**)	4.91 (**2.57**)	3.67 (**3.87**)	1.60 (**1.28**)	1.1 (**2.01**)	0.44 (**0.64**)	0.30 (**1.22**)	0.11 (**0.39**)
Beas	7.39 (**4.13**)	6.12 (**1.45**)	2.60 (**2.78**)	1.99 (**1.27**)	0.78 (**1.92**)	0.55 (**1.01**)	0.21 (**1.44**)	0.14 (**0.86**)

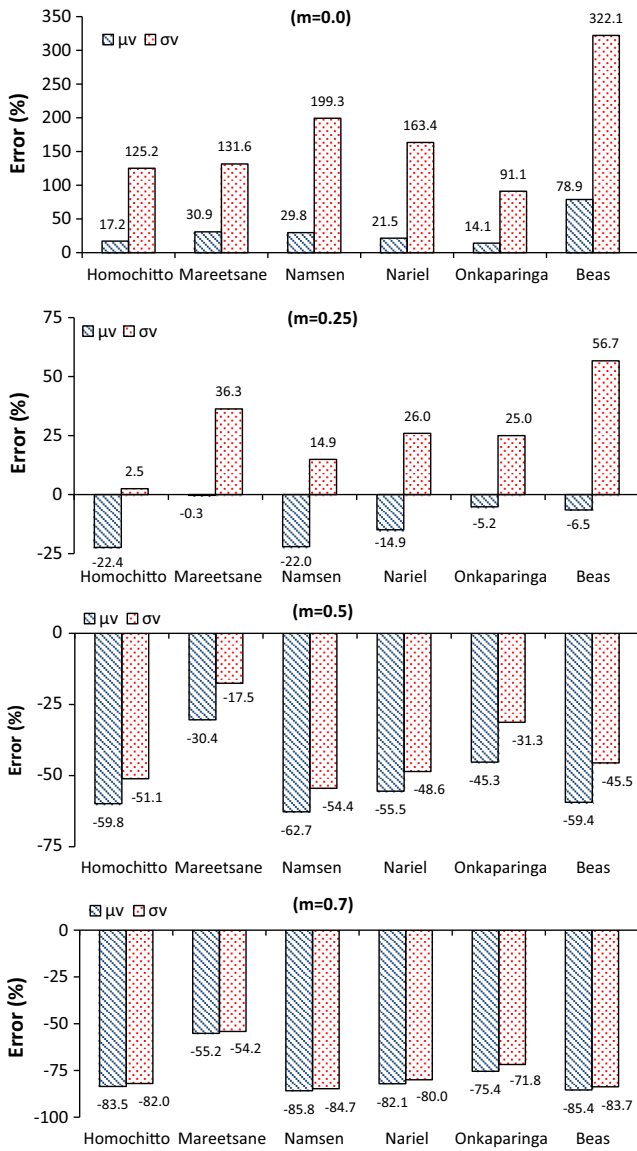The bold values indicate the observed values of the relevant statistics.

Fig. 2. Error (%) in mean and standard deviation of reservoir capacity (generic Phien model).

2.2. Models for mean and standard deviation of reservoir capacity

Phien (1993) generated 2000 replicates of annual runoff records with record lengths (n) varying between 20 and 50 years, and ρ in the interval [0, 0.5]. Once generated, each replicate was routed through hypothetical reservoirs using the SPA thus resulting in 2000 capacity estimates for each combination of m, n and ρ. The resulting "population" of capacity estimates was then analysed to determine its mean and standard deviation. Finally, the mean and standard deviation were independently related to the m, ρ and n via regression to develop generalized models. Four such models were developed by Phien, with three of these being for specific m values, while the fourth is generic and could in principle be applied to any m satisfying $0 \leqslant m \leqslant 0.5$. These models are summarized in Table 1.

In Table 1, the mean μ_V and standard deviation σ_V relate to the scaled capacity $V = K_a/\sigma_Q$, where K_a is the volumetric capacity (obtained via the SPA–see Section 3.1) and σ_Q is the standard deviation of the annual runoff. Consequently, V is dimensionless. Because all the four models include n, the record length (years), they cannot be applied at ungauged sites, which would limit the

use of the models to rapid evaluations for preliminary reservoir design studies at gauged sites.

2.3. Probability distribution and quantiles of reservoir capacity

The ultimate use of the generalized models is the provision of quantiles of reservoir capacity. The distribution of reservoir capacity has been established to be the three parameter distribution (see Vogel and Stedinger, 1987). The 3-parameter log-normal distribution has 3 parameters: the mean, standard deviation and the lower limit. In particular for reservoirs fed by annual inflows that are normally distributed, Bayazit and Bulu (1991) showed that the standardized storage capacity, V^* (see Eq. (9)) has a three parameter log-normal distribution with the following fixed parameters: location (or lower limit) = -2.0; mean = 0.5816; and standard deviation = 0.4724, where:

$$V^* = \frac{V - \mu_v}{\sigma_v} \tag{9}$$

Goodness-of-fit tests reported by Phien (1993) also confirmed that the 3-parameter log-normal distribution (with the above fixed parameters) can be used to describe the standardized capacity obtained for rivers fed by annual flows with the gamma distribution. The resulting quantiles for V^* thus become:

$$V_P^* = -2.0 + \exp(0.5816 + 0.4724 Z_p) \tag{10}$$

where Z_p is the standard normal variate for P (%) cumulative probability which can be approximated using (Stedinger et al., 1993):

$$z_p = \frac{(0.01P)^{0.135} - (1 - 0.01P)^{0.135}}{0.1975} \tag{11}$$

With V_P^* known, the corresponding quantile for the scaled capacity V_P can be obtained by re-arranging Eq. (9), i.e.:

$$V_P = \mu_V + \sigma_V V_P^* \tag{12}$$

3. Methodology

3.1. Stochastic data generation

This work relies on generating several realisations of the at-site historic runoff data record for each of the rivers analysed. Thus, similar to the approach by Phien (1993), annual runoff was assumed to follow the lag-1 autoregressive model (Fiering and Jackson, 1971):

$$Q_{t+1} = \mu_Q(1 - \rho) + \rho Q_t + \sigma_Q z_g (1 - \rho^2)^{0.5} \tag{13}$$

where Q_{t+1} and Q_t are the annual runoff for years $t + 1$ and t respectively; μ_Q = mean of Q; z_g is the as-gamma standard variate, i.e. with a mean of zero and variance of unity; and all other symbols are as defined previously.

The variate z_g imparts the gamma distribution to the generated data; however, rather than generate this standard gamma variate directly, it was approximated by transforming from standard normal variate using the Wilson–Hilferty expression (Wilson and Hilferty, 1931):

$$z_g = \frac{2}{\gamma_Q} \left[\left(1 + \frac{\gamma_Q}{6} \left(z_n - \frac{\gamma_Q}{6} \right) \right)^3 - 1 \right] \tag{14}$$

where γ_Q is the skew coefficient of annual runoff and z_n is the equivalent standard normal variate. If the lag-1 serial correlation coefficient is significant, then the skew coefficient must be corrected for the effect of serial dependence using (Thomas and Burden, 1963):
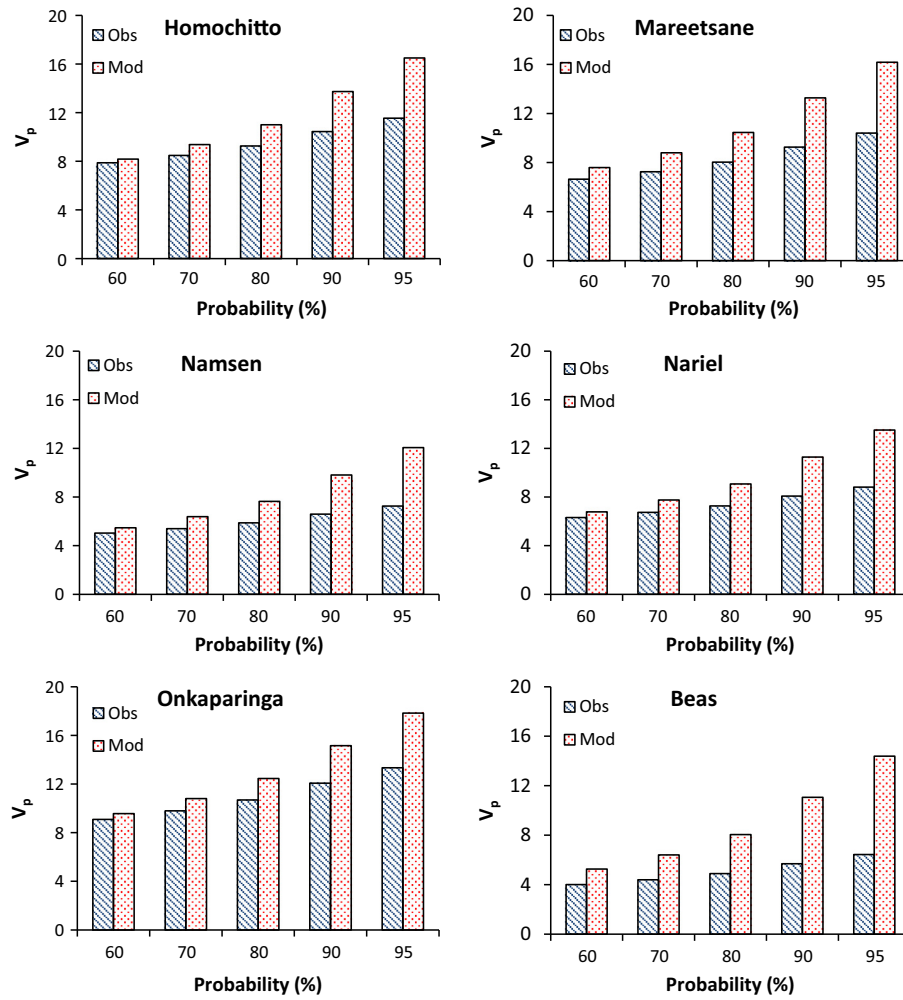
Fig. 3. Observed and Phien generic-model-predicted reservoir capacity quantiles (full regulation, $m = 0.0$).

$$\gamma'_Q = \gamma_Q \left[\frac{1 - \rho^3}{(1 - \rho^2)^{1.5}}\right] \tag{15}$$

where γ'_Q is the serial-dependence adjusted skew coefficient which is used instead of γ_Q in Eq. (14). In general, the Wilson–Hilferty transformation works well for $-3 \leqslant \gamma_Q \leqslant 3$ (McMahon and Miller, 1971).

To use Eq. (13) for data generation at each site, model parameters (μ_Q; σ_Q, ρ, γ_Q) are first estimated using the available historic record at the site. McMahon and Adeloye (2005) provide expressions for unbiased estimates of these and other parameters. Once estimated, the parameters are then used in Eq. (13) to generate 2000 replicates of the historic record, with the length of each replicate being equal to the historic record length. The choice of 2000 was meant to comply with Phien's experiments; to ensure that the replicates are independent the generator will be re-seeded after generating each replicate.

3.2. Sequent peak algorithm (SPA) for reservoir capacity estimation

The SPA was implemented using:

$$K_{t+1} = \max[0.0; K_t + D_t - Q_t]; t = 1, 2, \ldots, n \tag{16}$$

where K_t and K_{t+1} are, respectively, the cumulative sequential deficits at the beginning and end of year t, Q_t is the reservoir inflow dur-

ing t and all other symbols are as previously defined. The SPA initialises with a full reservoir (i.e. $K_o = 0.0$) and the iteration is limited to the single cycle of the data record if the final sequential deficit is also zero, i.e. $K_{n+1} = 0$. If this condition is not met, another cycle of the data record is used but starting with the previous K_{n+1}, i.e. $K_o = K_{n+1}$ for initializing Eq. (16). The second cycle should end with the starting K_{n+1}; otherwise the SPA has failed, a situation that would normally result if too much water (i.e. $\alpha > 1$) is being taken from the reservoir. Once Eq. (16) has converged, the capacity estimate then becomes:

$$K_a = \max(K_{t+1}) \tag{17}$$

where K_a is the reservoir capacity.

3.3. Data records

The analyses used six global rivers as listed in Table 2. The main characteristics of the six rivers are also summarised in Table 2. Although only six, the first thing to note is that the rivers represent a wide range of annual CV (0.185–1.012) which covers most conditions in the world (McMahon et al., 1992). The river catchments vary in size from a minimum of 252 km² to maximum of 12,561 km². The available record lengths at the sites ranged from 15 to 69 years, making all but two (i.e. the Onkaparinga at Clarendon weir and Beas at Pong dam) to be within the 20–50 years range
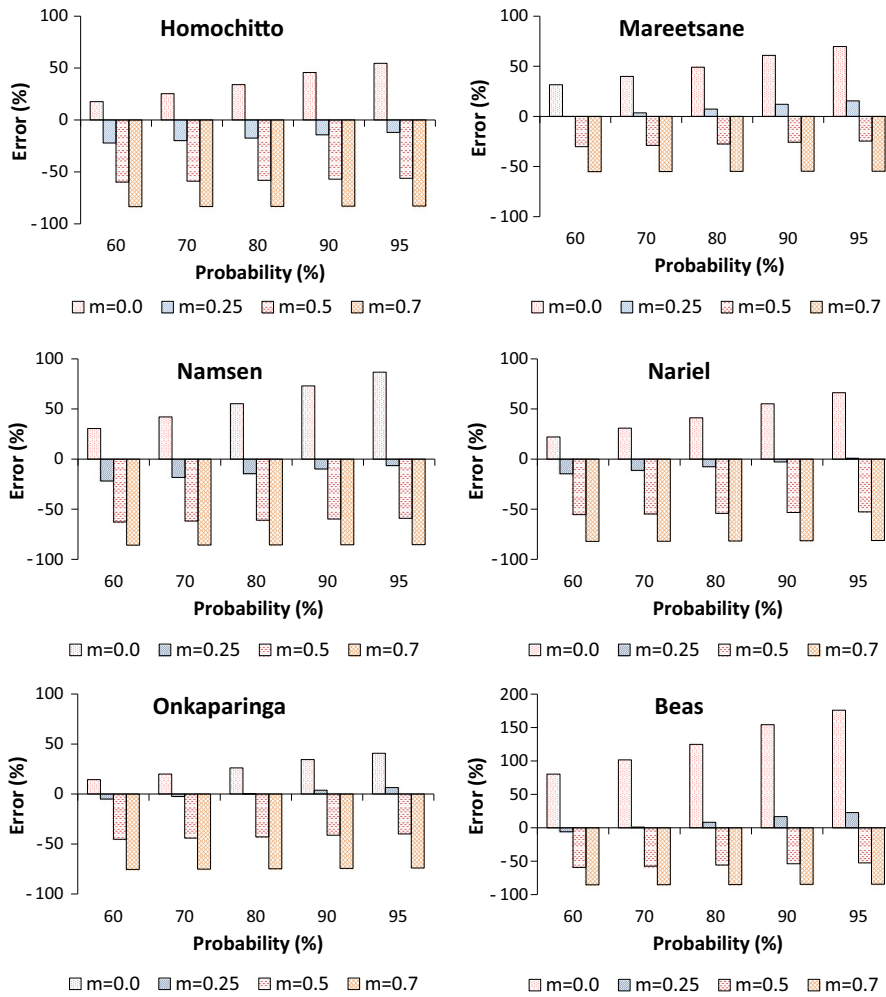
Fig. 4. Errors (%) in reservoir capacity quantile estimates with the Phien generic model.

employed in Phien's analyses. The Onkaparinga data length is 69 years and thus exceeds the upper limit of 50 years while the Beas data length of 15 years is shorter than the lower limit of 20 years considered by Phien. The two records thus offer an opportunity for testing the effect of too short or too long record lengths on the performance of the Phien models.

The estimated lag-1 serial correlation coefficients in Table 2 contain some negative values but these are small that they can be taken as zero. Thus, the selected historic runoff records also meet the lag-1 serial correlation conditions of the Phien models. Much more crucial, however, is whether the runoff for the selected rivers can be described by the gamma distribution or not.

Table 2 contains the 95% CONF for the theoretical gamma skewness and as can be seen from this Table, all the sample skew estimates fall within the 95% CONF limits, implying that there is no statistical evidence to reject the hypothesis that the runoff data follow the gamma distribution function. The results of the more formal Kolmogorov–Smirnov test are shown in Table 3 – the null hypothesis that the runoff records follow the gamma distribution is not rejected at the 5% level.

As remarked earlier, the estimated lag-1 serial correlation coefficients shown in Table 2 are low and hence unlikely to produce huge adjustments to the estimated skew; nonetheless, they have been used to adjust the skew coefficient according to Eq. (15) before using it in the data generation Eq. (13).

4. Results and discussions

4.1. Data generation

Table 4 compares the statistics of the historic and stochastically generated runoff records at the sites. The statistics of the synthetic represent the average over the 2000 replicates generated at each site using Eq. (13). As seen in Table 4, the stochastic model has adequately preserved the mean, standard deviation and skew coefficient of the historic runoff at all the sites. The performance of the models with respect to the lag-1 serial correlation coefficient was not as good but given that the serial dependence was generally low as noted earlier, this should not be of much concern. As noted by Burges and Linsley (1971), the most important runoff statistic that influences reservoir capacity estimate is the CV of annual runoff. Indeed, over-year reservoir capacity will generally vary as the square of the annual runoff CV (see McMahon et al., 2007b,c; McMahon and Adeloye, 2005); thus a doubling of the CV will result in four-fold increase in the estimated reservoir capacity for the same demand level. On the contrary, the effects of both the skew coefficient and serial dependence on capacity estimates are usually marginal; consequently, failure to adequately reproduce especially the serial correlation model is not seen as a major limitation of the analysis.

4.2. Performance of Phien models: mean and standard deviation of reservoir capacity

The mean (μ_V) and standard deviation (σ_V) of the scaled reservoir capacity (V) obtained using the SPA (i.e. the observed) are compared with those predicted by the m-specific models of Phien (i.e. simulated) in Table 5. The observed mean (μ_V) and standard deviation (σ_V) of reservoir capacity were obtained using:

$$\mu_v = \frac{\sum_{i=1}^{N} V_i}{N} \tag{18}$$

$$\sigma_v = \sqrt{\frac{\sum_{i=1}^{N}(V_i - \mu_v)^2}{(N-1)}} \tag{19}$$

where V_i is the scaled capacity estimate from runoff replicate i and N (=2000) is the number of replicates. The corresponding simulated values were obtained using the appropriate equations from Table 1.

For the full regulation situation ($m = 0.0$), Table 5 shows that Phien model is over-predicting the mean of reservoir capacity and significantly over-predicting the standard deviation of reservoir capacity. The performance of the Phien model in respect of the mean of capacity (μ_v) was particularly poor for the Beas data, which is not surprising given that the Beas data record was one of the two that failed to meet the data record length criterion for the Phien models. As noted earlier, the runoff record length for the Beas was a mere 15 years, shorter than the minimum 20 years considered by Phien; thus the fact that the Beas is recording a higher bias in μ_V than Onkaparinga is evidence that violating the record length is less consequential provided the record length is long.

As m becomes higher, i.e. as the demand ratio α reduces, there is a tendency for the performance of the Phien models to be a mixture of under- and over-prediction for the mean of reservoir capacity, although the standard deviation of reservoir capacity is largely still being over-predicted by the Phien models. At the lowest demand ratio level considered by Phien, i.e. $m = 0.5$, the Phien model is under-predicting both the mean and standard deviation of reservoir capacity for all the rivers analysed.

The errors of the Phien estimates relative to those based on the use of the SPA are shown in Fig. 1 for the m-specific models, which further confirm the over-prediction at full regulation ($m = 0$) but under-prediction of the mean capacity as the m increases, i.e. as the demand ratio reduces. For the highest m (i.e. lowest demand ratio) considered by Phien, the highest under-prediction of the mean capacity was 65% while the highest under-design error of the standard deviation of capacity was 60%. This would imply that care should be exercised when applying the models to high values of the drift.

The above discussions relate to the m-specific models but as noted earlier, Phien also developed generic models (see Table 1) that could be applied to any value of the drift (m) in the range [0,0.5]. Table 6 is a comparison of the generic model estimates of both the mean and standard deviation of reservoir capacity with those obtained via SPA simulations. As was the case with the m-specific models, the generic model was also over-estimating the mean and standard deviation of capacity at full regulation ($m = 0$) but as the m increases, the over-estimation gradually became under-estimation. Similar to what was also observed for the m-specific models, the Phien generic model was under-estimating the mean capacity across all the rivers at $m = 0.5$, again signifying the deterioration in performance as the drift increases (or demand reduces).

The error plots for the generic models are shown in Fig. 2, which if juxtaposed with the plots in Fig. 1, will show broadly similar error magnitudes for the generic model when compared with the

m-specific models. Both Table 6 and Fig. 2 include an additional case of $m = 0.7$ which was not included in Phien's analysis but was deliberately included in the current study to test whether or not the generic model could actually be applied beyond the upper $m = 0.5$ limit considered by Phien without a significant deterioration in performance. As shown in Fig. 2, the under-prediction errors in both the mean and standard deviation of capacity for the $m = 0.7$ case were significantly much higher than those recorded for the $m = 0.5$ cases (both m-specific and generic), thus confirming that the generic model is a poor extrapolator. The $m = 0.7$ case also confirms that high drifts (i.e. $m \geqslant 0.5$) will consistently result in the under-design of reservoir capacity with the Phien models.

As noted earlier, there have been very limited studies that investigated the efficacy of the Phien models with real data. The only known study was reported by McMahon et al. (2007b), albeit for the generic model and with regard to only the mean of reservoir capacity (by virtue of being based on single historic records). Nonetheless, there is some resemblance between the outcome of McMahon et al. (2007b) work and the current study. For example, they reported that Phien generic model underestimated the reservoir capacity estimates by about 25% for a demand ratio $\alpha = 0.75$ for rivers that partially met the criteria set by Phien. When McMahon et al. (2007b) included rivers not meeting Phien's criteria in their analyses, they found that the underestimation of reservoir capacity was as high as 80%.

Although the m equivalent of the α investigated by McMahon et al. (2007b) would depend on the CV of the river (see Eq. (8)), thus making a strict comparison with the results obtained in the current study difficult, the errors of 25% and 80% are within the under-design errors recorded in the current study for the mean of reservoir capacity. Thus on this basis, this study has reinforced the caveats by McMahon et al. (2007b) regarding the use of Phien models.

However, it is also important to recognize that the Phien models are not just producing under-design situations but significant over-designs as well, especially at full regulation. Additionally as shown in this study, these biases are not being recorded for the mean alone but also for the standard deviation of reservoir capacity, implying that the entire distribution of reservoir capacity and the resulting capacity quantiles can potentially be subjected to significant uncertainties as will be demonstrated in the next section. While under-design of capacity would mean frequent failures of a reservoir to meet demands, an over-designed reservoir will tie down scarce financial resources that could otherwise be better utilized for other developmental needs. Thus, although cases of full regulation (i.e. $m = 0$) might be rare in practice, the over-design errors in the mean capacity obtained in this study are significant because of the large reservoir capacity often associated with $m = 0$, with significant concomitant financial consequences.

4.3. Performance of Phien models: reservoir capacity distribution and quantile estimates

Observed and model-predicted reservoir capacity quantiles based on the 3-parameter log-normal distribution as described in Section 2.3 are compared in Fig. 3 for the full regulation ($m = 0$) situation of the generic model. The decision to restrict consideration to the generic model is based on the fact it is likely to be one that will be used mostly because it can in theory be applied to any m value. As seen in Fig. 3, the Phien generic model is over-predicting the reservoir capacity quantiles for $m = 0$, which is not surprising, given that the model also over-predicted both the mean and standard deviation of reservoir capacity. However, what Fig. 3 also highlights is that the bias associated with the Phien generic model was particularly poor for the Beas River, a further reinforce-

ment of the need to be wary in attempting to use the model where its underlying assumption in relation to the record length is not met.

Fig. 4 shows the errors associated with the quantile estimates for all the m values investigated with the generic model. Again as was the case with the mean and standard deviation of capacity, the errors are all over-prediction for $m = 0$; however, when compared to the corresponding errors in the mean, the quantile errors are much higher. This should not be surprising given that the quantile error integrates the error in both the mean and standard deviation of reservoir capacity. As the results for $m = 0$ situation clearly show, the over-prediction error (%) in the standard deviation of reservoir capacity was significantly higher than that of the mean; the net effect of this is a much larger error in the quantiles compared to the mean. For $m > 0$, the errors become a mixture of over- and under-predictions again with the tendency towards sustained under-prediction as m becomes larger. Indeed, the largest under-prediction error was obtained with the out-of-range $m = 0.7$ situation, where as seen in Fig. 4 the error was as high as 100% for the Beas River for all the quantiles.

5. Conclusions

For the first time, a complete assessment of the efficacy of all four Phien models in predicting the mean, standard deviation and by extension the quantiles of reservoir capacity has been carried out using observed runoff data records of six global rivers that meet the stated criteria of the models. The results showed that all the models result in over-design of reservoir capacity at full regulation. However, as the drift parameter m increases, the errors become a mixture of both under-design and over-design. For the highest $m = 0.5$ considered by Phien, all the errors are under-design errors. All this will point to the fact the Phien models should be used with caution for reservoir capacity planning, especially where significant under-prediction will result because of the impact of such on the overall performance of a reservoir. Over-prediction situations may be acceptable since they represent built-in safety factors that might be useful for cushioning the uncertainties associated with climate and land-use changes. Since, as revealed in this study such over-predictions are associated with low drift values, i.e. high demands (including full regulation) in which the capacity requirement is already high, the associated financial cost is, however, likely to be very high. A final aspect to reinforce here is that the Phien models should not be applied in situations where the underlying assumptions for the models are not met. In particular, short record lengths below the lower limit of the Phien validity range, such as presented by the Beas River case, will accentuate the estimation errors even when all other criteria of the models are being met.

Acknowledgements

References

Adeloye, A.J., 2012. Hydrological sizing of water supply reservoir. In: Bengtsson, L., Herschy, R.W., Fairbridge, R.W. (Eds.), Encyclopedia of Lakes and Reservoirs. Springer, Dordrecht, pp. 346–355.

Adeloye, A.J., 1990. Streamflow data and surface water resources assessment. J. Water Resour. Technol. – AQUA 39 (4), 225–236.

Adeloye, A.J., 1996. An opportunity loss model for estimating value of streamflow data for reservoir planning. Water Resour. Manage. 10 (1), 45–79.

Adeloye, A.J., Lallemand, F., McMahan, T.A., 2003. Regression models for within-year capacity adjustment in reservoir planning. Hydrol. Sci. J. 48 (4), 539–552.

Adeloye, A.J., Pal, S., O'Neil, M., 2010. Generalised storage–reliability modelling: Independent validation of the Vogel–Stedinger (V–S) model using a Monte Carlo simulation approach. J. Hydrol. 388 (3–4), 234–240.

Bayazit, M., Bulu, A., 1991. Generalised probability distribution of reservoir capacity. J. Hydrol. 126, 195–205.

Bayazit, M., Onoz, B., 2000. Conditional distribution of ideal reservoir storage variables. J. Hydrol. Eng. (ASCE) 5 (1), 52–58.

Burchberger, S.G., Maidement, D.R., 1989. Diffusion approximation for equilibrium distribution of reservoir storage. Water Resour. Res. 25 (7), 1643–1652.

Burges, S.J., Linsley, R.K., 1971. Some factors influencing required reservoir storage. J. Hydraul. Div., ASCE 97 (7), 977–991.

Fiering, M.B., Jackson, B.B., 1971. Synthetic stream flows. Water Resources Monograph 1, American Geophysical Union, Washington, D.C., USA.

Kuria, F.W., Vogel, R.M., 2015. Global storage–reliability–yield relationships for water supply reservoirs. Water Resour. Manage. 29 (5), 1591–1605.

Loucks, D.P., Stedinger, J.R., Haith, D.A., 1981. Water Resources Systems Planning and Analysis. Prentice-Hall Inc., Englewood Cliffs, NJ, USA.

Matalas, N.C., Benson, M.A., 1968. Note on the standard error of the coefficient of skewness. Water Resour. Res. 4 (1), 204–205.

McMahon, T.A., Adeloye, A.J., 2005. Water Resources Yield. Water Resources Publications, Littleton, CO, USA.

McMahon, T.A., Miller, A.J., 1971. Application of the Thomas and Fiering model to skewed hydrologic data. Water Resour. Res. 7 (5), 1338–1340.

McMahon, T.A., Finlayson, B.L., Haines, A.T., Srikanthan, R., 1992. Global Runoff – Continental Comparisons of Annual and Peak Discharges. Catena Verlag, Destedt, Germany.

McMahon, T.A., Pegram, G.G.S., Vogel, R.M., Peel, M.C., 2007b. Revisiting reservoir storage–yield relationships using a global historical streamflow data. Adv. Water Resour. 30, 1858–1872.

McMahon, T.A., Pegram, G.G.S., Vogel, R.M., Peel, M.C., 2007c. Review of Gould-Dincer reservoir storage–yield–reliability estimates. Adv. Water Resour. 30, 1873–1882.

McMahon, T.A., Vogel, R.M., Peel, M.C., Pegram, G.G.S., 2007a. Global streamflows – Part 2. Reservoir storage–yield performance. J. Hydrol. 347 (3–4), 260–271.

McMahon, T.A., Vogel, R.M., Peel, M.C., Pegram, G.G.S., 2007d. Global streamflows – Part 1. Characteristics of annual stream-flows. J. Hydrol. 347 (3–4), 243–259.

NERC, 1975. Flood Studies Report (5 volumes). Natural Environment Research Council, UK.

Phien, H., 1993. Reservoir storage capacity with gamma inflows. J. Hydrol. 146, 383–389.

Silva, A.T., Portela, N.M., 2012. Stochastic assessment of reservoir storage–yield relationships in Portugal. J. Hydrol. Eng., ASCE 18 (5), 567–575.

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. In: Maidment, D.R. (Ed.), Handbook of Applied Hydrology. McGraw-Hill, NY, USA (Chapter 19).

Thomas, H.A., Burden, R.P., 1963. Operations Research in Water Quality Management. Harvard Water Resources Group, Harvard University, MA, USA.

Vogel, R.M., Stedinger, J., 1987. Generalized storage–reliability–yield relationships. J. Hydrol. 89 (3), 303–327.

Wilson, E.P., Hilferty, M.M., 1931. The distribution of chi-square. Proc. Natl. Acad. Sci. 17, 684–688.