Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: https://www.tandfonline.com/loi/thsj20

Risks and uncertainties in reservoir yield in highly variable intermittent rivers: case of the Castanhão Reservoir in semi-arid Brazil

J.N.B. Campos, F.A. Souza Filho & H.V.C. Lima

To cite this article: J.N.B. Campos, F.A. Souza Filho & H.V.C. Lima (2014) Risks and uncertainties in reservoir yield in highly variable intermittent rivers: case of the Castanhão Reservoir in semi-arid Brazil, Hydrological Sciences Journal, 59:6, 1184-1195, DOI: 10.1080/02626667.2013.836277

To link to this article: https://doi.org/10.1080/02626667.2013.836277

Hydrological Sciences Journal – Journal des Sciences Hydrologiques, 59 (6) 2014
http://dx.doi.org/10.1080/02626667.2013.836277

Risks and uncertainties in reservoir yield in highly variable intermittent rivers: case of the Castanhão Reservoir in semi-arid Brazil

J.N.B. Campos, F.A. Souza Filho and H.V.C. Lima

Department of Hydraulic and Environmental Engineering, University Federal of Ceará, Fortaleza, Ceará, Brazil
jnbcampos@gmail.com

Received 25 August 2011; accepted 9 July 2013; open for discussion until 1 December 2014

Editor D. Koutsoyiannis; **Associate editor** S. Blazkova

Citation Campos, J.N.B., Souza Filho, F.A., and Lima, H.V.C., 2014. Risks and uncertainties in reservoir yield in highly variable intermittent rivers: case of the Castanhão Reservoir in semi-arid Brazil. *Hydrological Sciences Journal*, 59 (6), 1184–1195. http://dx.doi.org/10.1080/02626667.2013.836277

Abstract We analyse the risks and uncertainties in the assessment of reservoir yields in Northeast Brazil using Monte Carlo simulation. The case study considers the Castanhão Reservoir, which is located in the Jaguaribe River in the State of Ceará. The hydrologic regime of the reservoir inflow was obtained from a 37-year historical series. From the data, 600 series of annual inflows for 30 years were generated. The operation of the reservoir was simulated to estimate the uncertainties and risks regarding the water supply by the reservoir. The variability of regulated streamflows with 30-year horizons, measured by their coefficients of variation, is 27%. It was also found that the mean probability of failure for the 600 traces, with a mean release equal to the mean yield and an assumed 10% risk of failure, was 12.5%, which is 25% greater than the value of 10% used to estimate the yield.

Key words reservoir yield; uncertainties; risks; semi-arid

Risques et incertitudes de rendement de réservoir dans des rivières intermittentes très variables: cas du réservoir Castanhão dans le Brésil semi-aride

Résumé Nous avons analysé les risques et les incertitudes dans l'évaluation du rendement des réservoirs dans le Nord du Brésil, en utilisant la simulation de Monte Carlo. L'étude de cas a été consacrée au réservoir Castanhão, situé sur la rivière Jaguaribe dans l'État de Ceará. Le régime hydrologique à l'entrée du réservoir a été obtenu à partir d'une série historique de 37 ans. A partir de ces données, 600 séries d'entrées annuelles sur 30 ans ont été générées. Le fonctionnement du réservoir a été simulé pour estimer les incertitudes et les risques concernant l'approvisionnement en eau par le réservoir. La variabilité des débits régulés avec des horizons de 30 ans, mesurés par leurs coefficients de variation, est de 27%. On a également trouvé que la probabilité moyenne de défaillance pour les 600 séries, avec un lâcher moyen égal au rendement moyen et un risque d'échec supposé de 10%, était de 12,5%, ce qui est de 25% supérieur à la valeur de 10% utilisée pour estimer le rendement.

Mots clefs rendement de réservoir ; incertitudes ; risques ; semi-aride

INTRODUCTION

Water use rights in the State of Ceará are allocated through an instrument called a grant, which is defined in the Brazilian Water Act. Grants are bestowed for 30 years. The maximum discharge that can be granted in a hydrographic basin is a fraction (close to one) of the basin's reference discharge. This reference discharge is adopted in the law as total yield in the basin (Campos 2003), with 90% reliability. However, uncertainties inherent to random phenomena are not taken into account in this process.

In assessments of reservoir yields in semi-arid regions, river discharges, which are the main input variable, have great year-to-year variability. In classic studies, the safe yield, calculated as the regularized streamflow of an infinite reservoir, is estimated by the Rippl (1883) diagram, which is essentially a deterministic method.

Uncertainties inherent to regulated streamflow studies that use synthetic streamflow time series have been the subject of several studies since the beginning of the last century. In a pioneering study, Sudler (1927) generated 100 years of annual discharges and recorded the values onto cards that were randomly selected to compose an annual streamflow time series. This procedure can be considered as the predecessor of the Monte Carlo method. However, it is well accepted that the use of the Monte Carlo method began with the publication of the article, “The Monte Carlo method” in the *Journal of the American Statistical Association*. The method’s authorship is attributed to the mathematicians John Von Newman and Stanislav Ulam (Sobol 1994).

From the 1950s onwards, the use of the Monte Carlo method in hydrology has gained many followers. In reservoir sizing, among the many noteworthy papers, the following deserve mention: Chow (1951), Thomas and Fiering (1962), Fiering (1967), Matalas and Wallis (1971), Salas and Yevjevich (1972), Klemeš (1987) and Koutsoyiannis (2005).

Professor Vujica Yevjevich made important contributions towards establishing concepts concerning risks and uncertainties. For Yevjevich (1972), risk is understood as a permanent feature of the population of any random phenomenon. The level of basic risk is measured by the probability of values that are higher or lower than certain arbitrary values. Risk cannot be changed unless the population is changed. Imprecision, deficiencies and biases, which are always present in hydrological data, as well as in the properties of populations, must be estimated from data, but lead to several mistakes and loss of information. When uncertainties of data are added to the basic risks, the result is the total risk or hydrologic risk. Yevjevich (1972) considers that, for errors stemming from sample variability, uncertainties are measured by the standard deviation, or variance, of $Qi-Q$ differences, where Qi is the value estimated from the ith sample and Q is the population parameter.

During the 1980s, the expansion of computer resources greatly boosted the application of the Monte Carlo method to the assessment of uncertainties in reservoir sizing. Yevjevich and Harmancioglu (1985) analysed the trends in hydrology research and pointed out the potential of using the Monte Carlo technique to estimate uncertainties in hydrological processes. Phien (1993) studied reservoir sizing using the sequent peak method and the premise that inflows follow a gamma probability distribution. It was concluded that reservoir capacities adjust to a lognormal probability function with three parameters.

The analytical determination of the probability function for the yield of a reservoir is not an easy task and has not yet been resolved by reservoir stochastic theory, except when considering very simplified premises (Fletcher and Ponnambalam 1996). This is one strong reason for the high number of applications of the Monte Carlo method in hydrology.

The importance of the variability in regulated streamflows for the hydrologic regime of Northeast Brazil was studied using stochastic hydrology techniques and considering a year subdivided into two seasons: a wet season when all inflows occur and a dry season when all withdrawals occur (Campos *et al.* 1997). Studart and Campos (2001) assessed uncertainties considering the effect of a reservoir’s initial conditions.

Vieira (2001) considers that uncertainties generate or imply risks, and defines risk as the probability of undesirable values, events or phenomena. He also states that measurements, observations and decisions made by water-resource managers have several types of uncertainties that inevitably result in countless types of risks.

Tyralis *et al.* (2013) used the Monte Carlo technique as a tool for computing confidence intervals and derivative quantities for the evaluation of risks and uncertainty. The method, according to the authors, is heuristic and general, and can be applied for any statistical distribution with any number of parameters.

Koutsoyiannis and Montanari (2007) studied long-term persistence (LTP) in hydroclimatic research. Using a temperature series, they demonstrated analytically that the LTP suggests a dramatic increase of uncertainty in statistical estimation and a decrease in significance on statistical tests. The authors suggested that before drawing concrete conclusions, a methodological framework, supported by physical and statistical arguments, should be built.

McMahon *et al.* (2007b) assessed several reservoir sizing techniques available in the literature for 729 rivers worldwide. It was concluded that the simulation of reservoirs using historical series is convenient for the task of reservoir sizing. It was also concluded that, on an annual time scale, to estimate a reservoir’s capacity with 98% certainty, a minimum of 50 years of data is necessary. These data were only available for 27% of the rivers analysed.

Hejazi *et al.* (2008) defined hydrologic uncertainty in a reservoir system as the entropy value of reservoir inflow. The entropy of the observed release data was used as a measure of decision uncertainties. The authors studied 79 reservoirs in California and the Great Plains, USA, using a data-mining approach. The authors found that inflow forecast is the least important indicator in release decision making. Current inflow was found to be the most important indicator in wet seasons, while previous releases were more relevant during dry seasons.

Kerrour *et al.* (2010) used grid-enabled Monte Carlo analysis to evaluate the impacts of uncertain discharges in sea water intrusion in the Korba aquifer in Tunisia. The Monte Carlo method was necessary to know how the model propagates the uncertainties on the spatial distribution of extraction rate to the sea water intrusion model.

Campos (2010) used the Monte Carlo method to develop a model to evaluate the yield, evaporation and spill relationships considering the inflows from a gamma probability distribution function. The inflows were assumed to be serially uncorrelated.

Marton *et al.* (2011) used the Monte Carlo method to evaluate the uncertainties in the calculation of monthly discharge on reservoir storage. They compiled annual series of water stages measured at a limnigraph station and the coordinates of discharge *vs* water stage obtained from hydrometric measurements. They developed a procedure to deal with uncertainties on the elements of the reservoir input discharges series.

Hamed (2012) applied the Monte Carlo method for estimating the reliability of over-year storage under persistent Gaussian inflows using basic probability principles and numerical integration of the standard multivariate normal distribution. The method has the advantage of dealing with arbitrary autocorrelation structures and is recommended for planning/design stages of a reservoir.

The main objective of this paper is to show peculiarities of the hydrologic regimen of the Northeast region of Brazil, in particular the Ceará State. In that State, there is no perennial river; Jaguaribe River, the object of the study, is the major river in the State. Before the construction of Orós Reservoir in 1960, the Jaguaribe was considered the largest dry (intermittent) river in the world. This was ironically a matter of pride for the inhabitants of that semi-arid region. In addition, the coefficient of variation (Cv) of annual discharge is very high (1.41), amongst the largest in the world. According to McMahon *et al.* (2007d), only 10% of a sample of 1221 rivers worldwide has a Cv larger than 0.87. To understand this hydrological regimen and learn how to live in such a region of high uncertainty is a challenge for the people and researchers of Ceará.

The paper initially discusses the concepts regarding uncertainties and risks in hydrological processes. Next, research conducted using an empirical probability function of reservoir yield with an annual certainty of 90% for the 30-year synthetic series generated using a Monte Carlo simulation is presented. Additionally, the reservoir release equal to the mean yield was determined, and the risk of this release not being met during a 30-year period was estimated.

DEFINITIONS

Here, we define the main variables used in the analysis. The term reservoir yield (*M*) refers to the volume set to be released annually from the reservoir whenever there is availability. The regulated discharge is used synonymously with the term reservoir yield, or simply yield. The yield is the amount of water that is intended to be released from the reservoir whenever it is available. When the reservoir volume decreases to a given water level, only a portion or none of the yield is supplied. The amount of water that is actually withdrawn from the reservoir is termed release.

Reservoir performance is assessed through the reliability or its value with respect to one, which is the probability of failure. Reliability (*R*) is the probability that the reservoir meets the demand during the simulation period. The probability of failure (*f*), that is, the probability of a water shortage, is the probability that the reservoir does not meet the demand during the simulation period. The frequency of failure estimated from the synthetic traces is also called the probability of failure. The probability of failure, f_M, and the reservoir reliability, *R*, are estimated by:

$$f_M = \frac{n_M}{N_M} \tag{1a}$$

$$R = 1 - \frac{n_M}{N_M} \tag{1b}$$

where n_M is the number of months in which the reservoir failed to meet the demand, and N_M is the total number of simulated months.

There is an association between the yield and the reliability. A 90% reliability yield (M_{90}) means that the supply is guaranteed in 90% of years. In this paper, the yield is estimated for a yield with 90% reliability. Thus, hereafter, the symbols M and M_{90} have the same meaning.

The term risk indicates the probability of failure of a specific action in terms of an uncertain event that cannot be controlled by the relevant party. For a reservoir's operation, risk refers to the probability that there will not be enough water stock to meet the established demand. Thus, when planning a reservoir with a regulated discharge that is 90% guaranteed, the manager knows beforehand that there is a 10% risk of not meeting the demand.

In statistics, the term uncertainty is associated with the expression of a random variable, which can be defined as a function that combines an experiment's resulting value with a sample space. When the regulated discharge is estimated using a simulation with an effluent streamflow historical time series of N years for a point in the sample space, for example, a value of $M1$ (a random variable) is obtained. For other historical series, other values will be obtained ($M2$, $M3$, …, Mi).

TYPOLOGY OF UNCERTAINTIES

Vicens *et al.* (1975) classified the uncertainties found in hydrological studies into three categories:

- Type I uncertainties are those that result from ignoring the real model that governs the studied hydrological phenomenon;
- Type II uncertainties result from ignoring the parameters of mathematical models; and
- Type III uncertainties are inherent to natural processes and result from the variability of those processes.

In reservoir hydrological studies, Type I uncertainties are difficult to assess, because, in most cases, the true probabilistic models that govern natural processes are not actually known. The evolution of hydrological science, together with the growth of hydrological observation series, should allow the development of increasingly efficient models to describe natural processes in the future.

Type II uncertainties associated with model parameters already have a satisfactory theoretical treatment in many cases. For example, the mean value, which is a variable that is almost always present as a parameter in stochastic models, is conveniently treated by the de Moivre–Laplace theorem (or central limit theorem). Mistakes made in the assessment of model parameters generally arise from limitations related to the size of the hydrologic series. Stochastic theory shows that, for recorded series, mistakes in the assessment of the mean parameter increase rapidly, with an increase in the Cv of the studied random variable.

Type III uncertainties are inherent to stochastic processes. These uncertainties cannot be changed. They should be known and incorporated for water-allocation planning. This is the type of uncertainty assessed in this article.

AREA OF STUDY

The Jaguaribe River hydrographic basin is located almost entirely within the limits of the State of Ceará. The basin covers 75 669 km^2 and occupies 51.9% of the total area of Ceará State. It is located between latitude 4°30′–7º45′S and longitude 37°30′–41°00′W (Fig. 1). The Jaguaribe River has a length of 610 km from its sources to its mouth in the Atlantic Ocean.

The mean yearly precipitation in the basin ranges from 400 mm in the inland areas to 800 mm at the coast. In the upper valley, in the Cariri region, the annual precipitation reaches 1000 mm. The mean precipitation over the basin is close to 700 mm year^{-1}. The Cv of the precipitation ranges from 0.3 to 0.5 for most of the basin, but in a few places it reaches 0.7. Most of the precipitation, >60%, occurs in the period from February to May. Ninety percent of the annual precipitation happens in the first half of the year (Secretaria dos Recursos Hídricos do Ceará 1992).

The association of this highly concentrated rainfall regime with the crystalline soil that covers >80% of the basin results in intermittent rivers. Before the construction of reservoirs, these rivers had zero discharge for around 9 months of the year. The Ceará State has no perennial river. Hence, the safe water supply comes mainly from surface reservoirs. The Castanhão Reservoir is the largest in the state and probably the largest in the world built in an intermittent river.

The Castanhão Reservoir, located in the Jaguaribe River, was selected for this study. This reservoir has a total capacity of 6700 hm^3, of which 250 hm^3 comprise an intangible reserve, 4250 hm^3

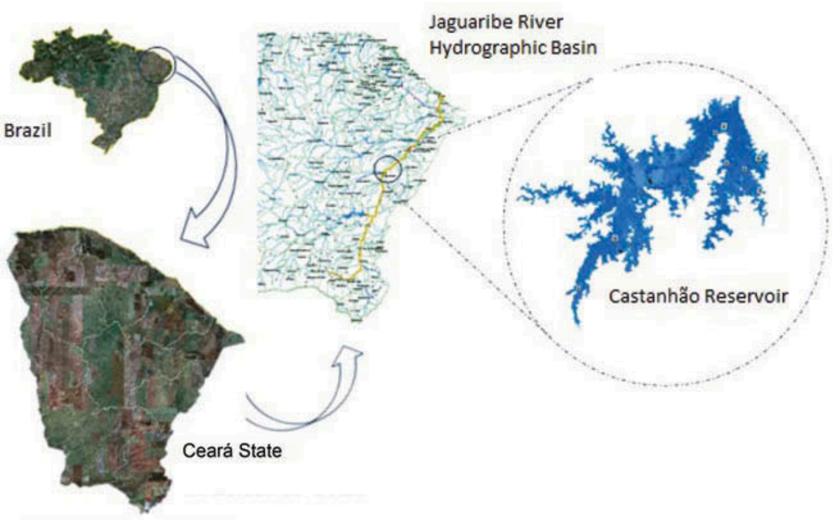

Fig. 1 Map of the Castanhão Reservoir in the Jaguaribe River basin, Ceará State, Brazil.

Table 1 Characteristics of the main reservoirs in the hydrographic basin of the Castanhão Reservoir.

Reservoir name	River	Hydrographic basin (km^2)	Operation volume (hm^3)	
			Minimum	Maximum
Castanhão	Jaguaribe	43 900	249.92	4451.66
Orós	Jaguaribe	24 538	413.12	1953.26
Trussu	Trussu	1775	0.00	263.00
Arneiroz II	Jaguaribe	5560	0.00	239.90
Atalho II	Dos Porcos	1270	7.25	108.25
FariasBrito	Cariús	840	0.00	197.57
Bastiões	Bastiões	2200	0.56	136.74

comprise the conservation volume and 2200 hm^3 are for flood protection. Castanhão controls a hydrographic basin with an area of 43 900 km^2, and there are several dams that modify the hydrologic regime of the Jaguaribe River. The characteristics of the main reservoirs in the Castanhão Reservoir hydrographic basin are presented in Table 1.

DATA AND METHODOLOGY

The Monte Carlo method was used to simulate the reservoir operation using a synthetic streamflow time series. For the case study, the Castanhão Reservoir was chosen as being representative of the hydrologic regime of intermittent rivers in Northeast Brazil.

This first stage of the work consisted of collecting available data on streamflow into the Castanhão Reservoir. A 37-year time series was obtained from

Table 2 Average monthly and annual evaporation and precipitation over the Castanhão Reservoir lake in the Jaguaribe River, Ceará State, Brazil.

	Tank evaporation (mm)	Lake evaporation, E (mm)	Rainfall, P (mm)	$P - E$ (mm)	α
Jan.	190	152	76	76	0.07
Feb.	160	128	89	39	0.03
Mar.	150	120	206	−86	−0.08
Apr.	160	128	160	−32	−0.03
May	180	144	95	49	0.04
Jun.	190	152	45	107	0.09
Jul.	200	160	16	144	0.13
Aug.	210	168	3	165	0.15
Sep.	220	176	2	174	0.15
Oct.	230	184	2	182	0.16
Nov.	210	168	2	166	0.15
Dec.	200	160	17	143	0.13
Total	2300	1840	713	1127	1

Sources Evaporation: GVJ, Política da Águas p.172; Rainfall: PERH, Estudos de Base p. 493.

the *Secretaria dos Recursos Hídricos do Ceará* (the State Water Resources Department) and is presented in Table 2. Using that time series, the annual streamflow into the Castanhão Reservoir was obtained. It was assumed that the streamflow data follow a gamma distribution function with two parameters. Therefore, with the selection of the probability density function and the estimation of its parameters, the reservoir inflow volume is defined. Then, it was possible to generate a synthetic annual streamflow time series.

To determine the monthly streamflow regime of the Jaguaribe River, the fragment method (Svanidze 1980) was used. From monthly fragments of historical data and the annual synthetic time series, it was possible to generate a synthetic monthly streamflow series.

To determine the evaporation and precipitation over the lake, the monthly averages obtained from the Jaguaribe River Water Management Plan (COGERH/ENGESOFT 2000) were used. The monthly values of the mean evaporation minus the mean precipitation, which represent the lake's water losses, are presented in Table 2.

Using the synthetic time series and the average losses of the lake ($E - P$) simulations were run to obtain the reservoir yield and the risk of not meeting the regulated discharge. The detailed steps of the methodology are outlined below.

Annual streamflow regime of the Jaguaribe River into Castanhão

The hydrographic system that forms the discharge of the Jaguaribe River into Castanhão was subdivided into three subsystems: the basin controlled by the Orós Reservoir, the Salgado River system and the complementary system. After an analysis of the available data from those subsystems, it was possible to compose a reliable 37-year time series (1957–1993) of monthly streamflows. The Jaguaribe River runoff regime in Castanhão was defined by: the annual average discharge (1463.50 hm^3 year^{-1}) and the Cv of the annual discharge, Cv (1.41).

To determine the streamflow in the Castanhão, it was assumed that the discharges followed a gamma probability distribution function with two parameters $G(x{:}\alpha,\beta)$. The probability density function (equation (2)) has the form:

$$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \tag{2}$$

where $f(x)$ denotes the probability function; x is the random variable annual affluent volume; Γ denotes the gamma mathematical function; and α and β are the shape and scale parameters, respectively.

The method of moments was used to estimate parameters α and β, as follows:

$$\mu = \alpha\beta \tag{3a}$$

$$\sigma^2 = \alpha\beta^2 \tag{3b}$$

where μ and σ are the mean and the standard deviation of the population, respectively, and α and β are the gamma distribution parameters.

Serial independence of annual inflows

According to Koutsoyiannis (2005), the typical simplifying assumptions in reservoir design, at least for the initial stage, are to: neglect the secondary inflows from precipitation losses due to evaporation and leakage; neglect the seasonality; assume that inflows are independent of time; use for inflows a specific distribution function as normal, lognormal or gamma. In this paper, the following were considered: precipitation over the lake, evaporation losses, seasonality and a two-parameter gamma distribution function. The inflows were considered serially independent, as justified below.

Intermittent rivers in the Brazilian semi-arid region are usually considered to be serially independent for two reasons: (a) a long dry season that lasts from 6 to 9 months, and (b) the predominance of crystalline soils and a high evaporation rate (Campos 1987). To test this hypothesis, the statistical test for serial independence was performed. A null autocorrelation coefficient for a time lag of 1 year was considered to be the null hypothesis (H_0), and a serial correlation was considered as the alternate hypothesis (H_1).

The test statistic, z, is given by:

$$z = \frac{\sqrt{n-3}}{2}\ln\left[\frac{(1+r)(1-\rho_0)}{(1-r)(1+\rho_0)}\right] \tag{4}$$

where z is the standard normal variable; n is the sample size (=37); r is the sample correlation coefficient (0.06); and ρ is the coefficient of correlation of the population, assumed to be zero.

For a significance level of $p < 0.05$, the critical z values are −1.96 and 1.96. Considering that a value of z = 0.36 was obtained, there is no evidence to reject the null hypothesis. Therefore, the hypothesis that the annual streamflows are serially independent is maintained. In summary, it can be concluded that the annual streamflow regime in the Castanhão Reservoir hydrographic basin belongs to a gamma population that is serially independent, with $\alpha = 0.503$ and $\beta = 2910$.

At the global scale, McMahon *et al.* (2007c) evaluated streamflow characteristics of 1221 rivers worldwide. The key features of annual streamflow examined were: mean, variability and skewness, distribution type (normal, gamma or lognormal), flow percentiles and dependence. In that sample, 249 values were significantly different from zero at the 5% significance level. They described the spatial location of rivers with high autocorrelation as: (i) regions with permanent snow cover contributing to the streamflow; (ii) having natural lakes within the basin; (iii) experiencing a significant trend in streamflow during the period of record; or (iv) having a short record length (sampling variability).

The physical conditions of Jaguaribe River basin do not fit these characteristics. The period of 6 to 9 months of zero inflows in the river associated with the intense evaporation prevents there being any carry-over water.

As the streamflow series is relatively short (37 years), two longer series of streamflow in the same river were evaluated: the Oros Reservoir (83 years) and the Arneiroz Reservior (86 years). For the Oros Reservoir, the lag-1 autocorrelation was 0.0177, yielding z values of 0.161 (equation (4)), while, for Arneiroz Reservoir, the autocorrelation was 0.0576 and a value of z = 0.5161 was obtained. Both values are significantly less than the critical value of z for 5% significance level (1.96). Thus, the assumption of serial independence of annual discharges is empirically supported by the streamflow data.

Monthly streamflow regime

Uncertainties in the monthly streamflow time series of 30 years' duration were assessed. The extent of 30 years was chosen for three reasons: (a) it is the duration of water allocation rights grants according to Brazilian law; (b) it is the same order of magnitude as natural streamflow time series; and (c) it coincides with the duration of the time series that are used to compose the climate norm.

The International Mathematics and Statistics Library (IMSL; Rogue Wave Software, Boulder, CO, USA) software routine was used to generate the time series. Initially, 600 time series of annual streamflow data were generated. Then, these series were transformed into monthly streamflow data using the fragment method (Svanidze 1980).

Determining the yield

When determining the behaviour of reservoirs, water budget methodology is most commonly used. A water budget consists of attributing reservoir water withdrawal rules and studying its behaviour for a certain effluent streamflow time series under these rules.

The budget consists of combining reservoir inputs and outputs and is expressed by:

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \text{Input} - \text{Output} \tag{5}$$

where $\mathrm{d}V/\mathrm{d}t$ denotes the rate of change in volume over time, and the Input and Output are all of the water inputs to and outputs from the reservoir, respectively. Water inputs to the system consist of stream discharge and direct rainfall over the lake, whereas outputs are a consequence of evaporation and dam withdrawals. Several procedures are available to solve this equation. In the next section, we present the development and solution of the water budget as it was determined herein.

Water budget equation

The water budget equation is given by:

$$V_{i+1} = V_i + (P_i - E_i) \times \frac{1}{2}(A_{i+1} + A_i) + I_i - R_i - S_i \tag{6}$$

where V_{i+1} and V_i represent the water volumes in the reservoir in months $i + 1$ and i, respectively; P_i is the average rainfall over the reservoir lake in month i; E_i is the average depth of water evaporated from the lake surface in month i: A_{i+1} and A_i represent the area of the reservoir lake in months $i + 1$ and i, respectively; I_i is the reservoir inflow volume in month i; R_i denotes the total of the reservoir withdrawals in month i; and S_i is the volume spilled into the reservoir by the spillway.

The calculation process consists of attributing a value for the withdrawal and assessing the reservoir's behaviour during the simulation period. The 90% reliability yield is computed by trial and error.

Monte Carlo simulation

Stochastic hydrology tools and a reservoir simulation operation were used to estimate the variability of the reservoir yield. The calculations were performed in the following stages: with the annual inflow series obtained from the historical data according to the methodology described previously, the parameters of a gamma distribution probability that was used to represent the discharges of the Jaguaribe Reservoir into the Castanhão River were obtained. The simulation was done in the following steps:

1. Using the IMSL software, 600 30-year time series of annual inflow into the Castanhão Reservoir were generated.
2. By applying the fragment method (Svanidze 1980), the annual streamflow series were transformed into monthly streamflow series. The Castanhão Reservoir was simulated for the 600 synthetic time series, and a sample of 600 reservoir yield values was obtained.
3. The statistical behaviour of the reservoir yield was studied to assess the Type III uncertainty.
4. To estimate the risk of the reservoir not meeting the planned demand, the reservoir was simulated for the 600 traces for a release equal to the mean value of the yield obtained from the 600 synthetic time series.

The computational effort involved was not significant: less than a second for each run of 236 000 months. This is consistent with Koutsoyiannis (2005), who simulated 456 000 months in less than a second on a common PC.

RESULTS

Variability analysis was performed according to two perspectives: the first analysis consisted of establishing a reliability of 90% by calculating the 600 values of M and adjusting a probability density function to that sample. The second analysis consisted of making a withdrawal equal to the mean of M (955.8 hm^3 year^{-1}) and calculating, for each of the 600 values in the series, the probability of failure. The results are presented in the following section.

Table 3 Descriptive statistics of the yields (hm^3/year) of 600 synthetic traces for the Castanhão Reservoir.

Mean (hm^3/year)	955.8	Skewness	0.37
SD (hm^3/year)	262.6	Max. yield (hm^3/year)	1849.8
Cv (-)	0.27	Min. yield (hm^3/year)	392.7

Note Cv: coefficient of variation; SD: standard deviation.

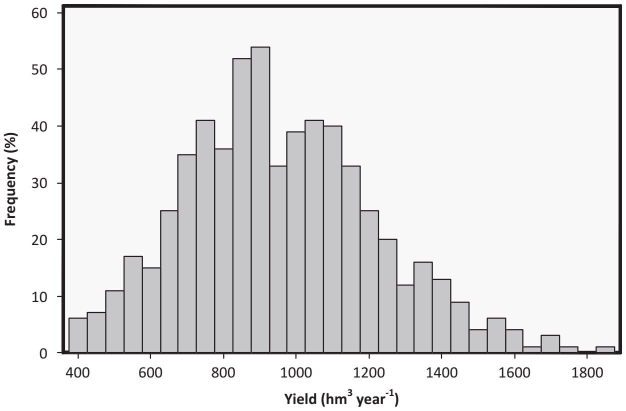

Fig. 2 Histogram of yield for Castanhão Reservoir obtained from a simulation of 600 synthetic 30-year series for withdrawal of 955.8 hm^3 year^{-1}.

Uncertainties in reservoir yield

The statistics for the 600 synthetic series are presented in Table 3; the histogram shown in Fig. 2 is bell shaped, with frequencies grouped in its central part and an approximately symmetrical reduction at its outer extremes. The median is very close to the mean (only 2% different). These characteristics indicate that the yield is derived from a normally distributed population. Nevertheless, the gamma function, as shown next, has an even better fit. The statistics for the samples are given in Table 3.

The Cv value for the annual precipitation over the Jaguaribe basin obtained from the Ceará Water Resources Plan is close to 0.3. Comparing this with the Cv for the Jaguaribe River at the Castanhão Reservoir (1.41), and that for the reservoir yield for the 30-year period (0.27) indicates that: (a) the variability increases greatly from annual precipitation to annual discharge (from 0.3 to 1.4); (b) after flowing into a reservoir for a 30-year period, the variability of the yield decreases to close to that of the precipitation.

In another approach to express the flow variability due to reservoir operation, one long-term synthetic series (5000 years) was simulated. The reservoir natural inflows are divided into two parts: controlled releases (yield) and the spill outflow. The Cv values obtained for the long-term series were as follows:

controlled outflow: Cv = 0.022;
uncontrolled spill: Cv = 3.03; and
controlled plus uncontrolled releases: Cv = 1.08.

That is, the reservoir causes the inflow variability (Cv) to drop from 1.41 to 1.08 in the outflow. A value of Cv = 0.022 was obtained for the controlled outflow; that means very low risk for the release.

The goodness-of-fit statistic, A^2, was introduced by Anderson and Darling (1952) to test the hypothesis that a given sample with the empirical distribution $Fm(x)$ is derived from a given population with the completely specified distribution function $F0(x)$. The criterion for selecting the best fit is the probability distribution function with the lowest value of A^2. A small p-value is an indication that the null hypothesis is false. Usually, the null hypothesis is rejected when $p < 0.05$.

Thus, to select the probability function that best fits the yields, the Anderson–Darling (AD) test was applied to the normal, lognormal and gamma functions by solving the following equation:

$$A^2 = -n - \left(\frac{1}{n}\right)\sum[(2i-1)\log(F(Xi)) + (2n+1-2i)\log(1-F(Xi))] \tag{7}$$

where A^2 is the AD statistic; n is the length of the series (600); i is the sequence number of the variable and $F(Xi)$ is the probability of the ith X variable.

Using the Minitab software (State College, PA, USA), the AD statistic and p-values were obtained, as listed in Table 4. The criterion for selecting the best probability function is to select the one with the lowest AD; the p-value is used to accept or reject the null hypothesis. In Table 4, the lognormal and normal distributions are rejected for the 1% significance level. The gamma function has the lowest AD and $p > 0.25$, that is, the hypothesis that the yield fits a gamma distribution can be accepted even for a 25% significance level. The parameters estimated for the gamma pdf were: shape equal to 12.92 and scale equal to 74.00.

Table 4 Anderson–Darling test statistics for the reservoir yield of the Castanhão Reservoir.

Distribution function	A^2	p-Value
Normal	1.288	<0.005
Gamma	0.339	>0.25
Lognormal	1.438	<0.005

Changes in Cv from precipitation to reservoir yield

The Alto Santo raingauge, which is located near the Castanhão Reservoir and has 33 years of records, was selected to represent the precipitation in the area. The mean annual rainfall at Castanhão is 792.5 mm year^{-1} and the Cv is 0.47. Comparing these values with the mean reservoir inflow of 1463.5 hm^3 year^{-1} and the Cv of 1.41, there is a huge increase in the Cv that comes from the association of rainfall regime with the climate (high evaporation rates) and the crystalline soils of the region.

After passing through the reservoir, the regulated outflows, or yield, have a mean of 955.8 hm^3 and Cv of 0.27 (Table 3). That is, the regulation action of the reservoir, transferring water from the flood years to the drought years, results in a Cv value lower than that of precipitation. To illustrate the changes in Cv from rainfall to reservoir inflow and reservoir yield, these variables were fitted to a gamma probability density function. The histogram and fitted pdf function are shown in Fig. 3.

The explanation for this great difference in variability between the precipitation and the streamflow is provided by the soil–rainfall relationship. In the years of low precipitation, the runoff coefficient decreases, sometimes to zero in very dry years. In high precipitation years, as the rainfall reaches the humid crystalline soils, the runoff coefficient increases to a very high value. In other words, low precipitation results in relatively lower discharge; high precipitation generates relatively higher discharge. In other words, the hydrological conditions of the Brazilian semi-arid region alter (stretch) the shape of the probability density function.

Studying the relationship between Köppen climate and the Cv of streamflow data, McMahon *et al.* (2007a) analysed a sample of 26 rivers with series of 25 years or longer and the Köppen climate BSh. They found that the coefficients of variation of annual discharge ranged from 0.28 to 1.40, with a median of 1.01. The Castanhão Basin, which also has a Köppen BSh climate, has a Cv of 1.41, slightly higher than that of the McMahon *et al.* (2007a) sample.

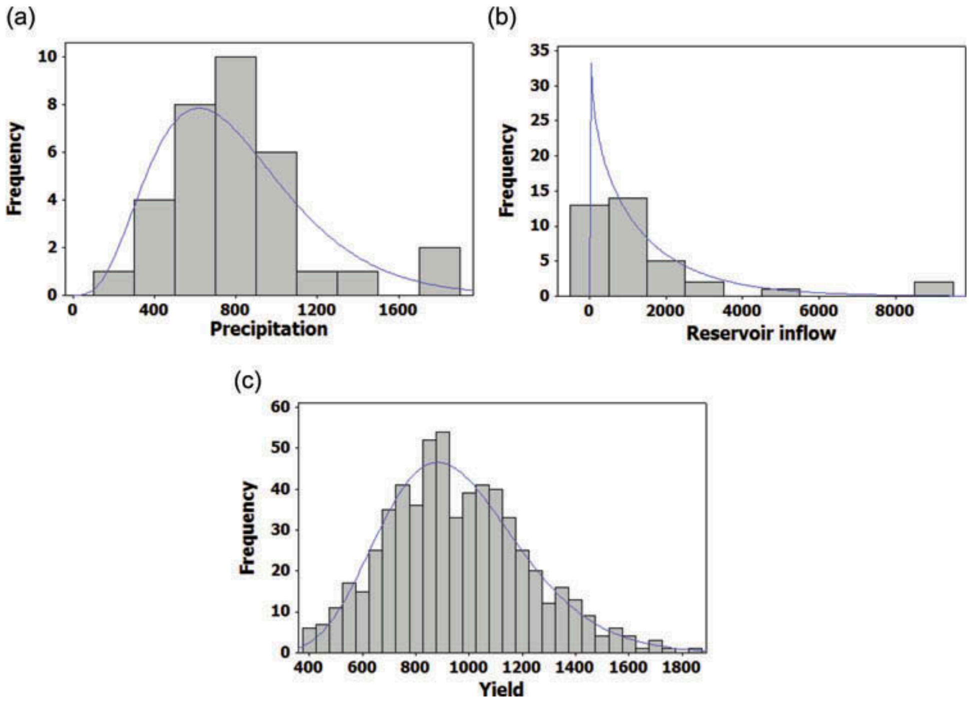

Fig. 3 Histogram and adjusted gamma probability density function for: (a) precipitation, (b) reservoir inflows and (c) reservoir yield.

For the Aw climate, with a sample of 40 rivers, McMahon *et al.* (2007a) found that the Cv varied from 0.14 to 0.72, with a median of 0.29. Vogel *et al.* (1965) studied the streamflow variability in 160 places in the northeastern USA and found a mean Cv of 0.25. Comparing these characteristics with those of the Castanhão, it can be seen that the Cv of Castanhão inflows (1.41) is more than five times that of global streamflows of an Aw climate (0.29) and almost five times that of the northeastern US rivers. Furthermore, the Cv of regulated discharge at Castanhão (0.27) is of the same order of magnitude as these two Cv values.

In conclusion, the variability of the reservoir yield estimated from the 30-year time series is approximately equal to the variability of the annual discharges for the rivers in the northeastern United States and of a Global Aw climate.

Risks for water supply

The probability of failure, that is, the frequency of a shortage in the water supply, was used to measure the risk associated with the reservoir operation. Assuming a release of 955.8 hm^3 $year^{-1}$, the reservoir was simulated for the 600 synthetic time series. The histogram values obtained for the failure frequency simulation are given in Table 5. The histogram shown in Fig. 4 is strongly skewed to the right.

Table 5 Descriptive statistics for the probability of reservoir failure with a withdrawal equal to the mean reservoir yield.

Mean (%)	12.5	Skewness	0.89
SD	9.8	Max. yield	46.4
Cv	0.78	Min. yield	0.0

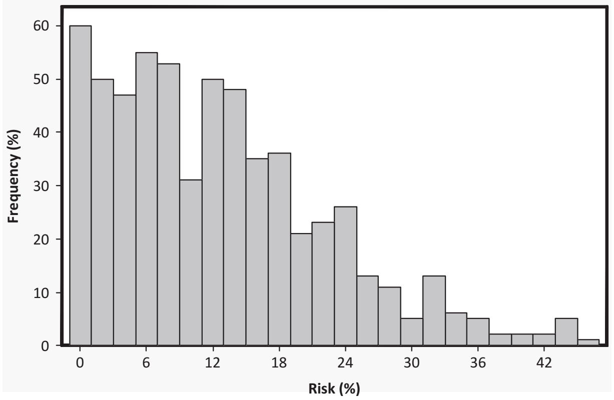

Fig. 4 Frequency of failure (risk) in Castanhão Reservoir obtained from a simulation of 600 synthetic 30-year series for a withdrawal of 955.8 hm^3 $year^{-1}$.

Table 6 Anderson–Darling test statistics (A^2) for the probability of failure for a release from the Castanhão Reservoir of 955.8 hm^3 $year^{-1}$.

Distribution function	A^2	p-Value
Weibull	1.009	0.012
Lognormal	11.581	<0.005
Gamma	2.192	<0.005

It was surprising that the expected value of the frequency of failure was 12.5%, that is, 25% above the 10% frequency of failure used to estimate each of the yields. If a planner assumes that the operational reservoir release is equal to the mean yield, with 90% reliability, a 12.5% risk of failure to meet the demand can be expected. Thus, the expected value of the reliability is 87.5%. The hypothesis test on the mean rejects, at 99% confidence level, the hypothesis that the 12.5% frequency of failure comes from a population with a frequency of failure equal to 10%. Now, if the planner really wants to assume a 10% risk, what risk would be used to estimate the yield?

A zero probability of failure of zero was observed in 37 samples. This result was expected because the zero probability of failure acts as an absorbing barrier. In this situation, a mixed distribution function with a probability mass at zero that is equal to 37/600 and a density function for positive values is the appropriate distribution for that variable. This truncation in the abscissa zero of the pdf probability of failure can explain the bias found in the prior paragraph. However, further studies using Monte Carlo method should be done to evaluate that bias better.

Next, the positive values of f were analysed. Using the Minitab software, the AD statistics and the p-values for the three best fitting pdfs were obtained (Table 6). Using the p-value criterion of a 1% significance level, only the Weibull distribution function is accepted. The pdf that fits the distribution of the probability of failure for the reservoir simulated with a release equal to the mean yield is a mixed function with a proportion of zero failures equal to 0.06 and a Weibull function for the positive values. The parameters of the Weibull function are: shape = 1.374 and scale = 6.038.

CONCLUSIONS

The study found a large Type III uncertainty in the reservoir yield estimation for the rivers with highly variable streamflows located in Northeast Brazil. Using the Cv as a measure of these uncertainties, the values found for the Cv were: 0.47 for precipitation, 1.41 for reservoir inflows and 0.27 for the regulated outflow from 600 traces of 30-year time series. The value of Cv for the Castanhão regulated discharges is of the same order of magnitude as that of unregulated streamflows in the northeastern United States (mean Cv = 0.25) and slightly less than the median of unregulated discharges in 40 rivers in Aw climate worldwide.

These large uncertainties are derived from the large coefficients of variation for the annual streamflows in the area. The association between the soil and rainfall affects the streamflow variability, reducing the Cv of runoff in low-precipitation areas and increasing it in high-precipitation areas. The yields from 600 synthetic traces follow a gamma probability density function with shape parameter equal to 12.92 and scale equal to 74.00.

The risk of not meeting the demand in a 30-year period shows an exponential decrease, with an increase in the frequency of failure, varying from zero (months without any failure) to 46.4%, that is, failing to meet the demand in 167 out of 360 months. Among the 600 traces, 37 traces with zero failure were observed. The concentration of series with no failures comes from the fact that the probability of zero failures acts as an absorption barrier. The sample of probability of failures fits a mixed probability distribution function with 6.2% of the traces having zero failures and a density function for the positive values. The density function fits a Weibull probability distribution function with a shape parameter of 1.374 and a scale parameter of 6.038.

The expected value of the probability of failure, obtained using the mean release with a frequency of failure of 10%, was 12.5%. In other words, when the reservoir is intended to release a discharge equal to the mean yield with a 10% probability of failure, the expected value of the probability of failure is 12.5%. In some way, this result was surprising. A new question arises: what should the frequency of failure be to evaluate the mean yield that gives a 10% risk? Surely, in this case, it should be a number greater than 10%. Further studies are necessary to solve this puzzle.

Funding The first author is thankful for the grant from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Proj. 307,532/2009-7).

REFERENCES

Anderson, T. and Darling, D., 1952. Asymptotic theory of certain "Goodness of Fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23 (2), 193–212. doi:10.1214/aoms/1177729437.

Campos, J.N.B., 1987. *A procedure for reservoir sizing under high evaporation rate*. Dissertation (PhD). Colorado State University, Fort Collins.

Campos, J.N.B., 2003. Política de Águas. *In*: J.N.B. Campos and T. M.C. Studart, eds. *Gestão Das Águas: Princípiose Práticas*. Porto Alegre: ABRH, 27–42.

Campos, J.N.B., 2010. Modeling the yield-evaporation-spill in the reservoir storage process: the regulation triangle diagram. *Water Resources Management*, 24 (13), 3487–3511. doi:10.1007/s11269-010-9616-x.

Campos, J.N.B., Araújo, J.C., and Souza Filho, F.A., 1997. Errors and variability of reservoir yield estimation as a function of the coefficient of variation of annual inflows. *In*: M. English and A. Szollosy-Nagy, eds. *Managing water: coping with scarcity and abundance*, 10–15 August, San Francisco, CA. IAHR, 524–529.

Chow, V.T., 1951. Discussion of long-term storage capacity of reservoirs. *Transactions of American Society of Civil Engineers*, 116 (1), 800–802.

COGERH/ENGESOFT, 2000. *Plano De Gerenciamento Das Águas Da Bacia Do Rio Jaguaribe*. Fortaleza: s.n.

Fiering, M.B., 1967. *Streamflow synthesis*. Cambridge, MA: Havard University Press.

Fletcher, S. and Ponnambalam, K., 1996. Estimation of reservoir yield and storage distribution using moments analysis. *Journal of Hydrology*, 182, 259–275. doi:10.1016/201C-1694 (95)02946-X.

Hamed, K., 2012. A probabilistic approach to calculating the reliability of over-year storage reservoirs with persistent Gaussian inflow. *Journal of Hydrology*, 93–99. doi:10.1016/j.jhydrol.2012.04.051.

Hejazi, M.I., Cai, X., and Ruddell, B.L., 2008. The role of hydrologic information in reservoir operation—learning from historical releases. *Advances in Water Resources*, 31, 1636–1650. doi:10.1016/j.advwatres.2008.07.013.

Kerrour, J., *et al.*, 2010. Grid-enabled Monte Carlo analysis of the impacts of uncertain discharge rates on seawater intrusion in the Korba aquifer (Tunisia). *Hydrological Sciences Journal*, 55 (8), 1325–1336. doi:10.1080/02626667.2010.519706.

Klemeš, V., 1987. One hundred years of applied storage reservoir theory. *Water Resources Management*, 1, 159–175. doi:10.1007/BF00429941.

Koutsoyiannis, D., 2005. Reliability concepts in reservoir design. *In*: J.H. Lerh and J. Keeley, eds. *Water encyclopedia, vol. 4, surface and agricultural water*. New York: Wiley, 259–265.

Koutsoyiannis, D. and Montanari, A., 2007. Statistical analysis of hydroclimatic time series: uncertainty and insights. *Water Resources Research*, 43 (5), WO5429. doi:10.1029/2006WR005592.

Marton, D., Stary, M., and Memsik, P., 2011. The influence of uncertainties in the calculation of mean monthly discharges on reservoir storage. *Journal of Hydrology and Hydromechanics*, 59 (4), 228–237.

Matalas, N. and Wallis, J.R., 1971. Statistical properties of multivariate fractional noise processes. *Water Resources Research*, 7 (6), 1460–1468. doi:10.1029/WR007i006p01460.

McMahon, T., *et al.*, 2007a. Global streamflows—part 3: country and climate zone characteristics. *Journal of Hydrology*, 347 (3–4), 272–291. doi:10.1016/j.jhydrol.2007.09.013.

McMahon, T., *et al.*, 2007b. Revisiting reservoir storage-yield relationships using a global streamflow database. *Advances in Water Resources*, 30 (8), 1858–1872. doi:10.1016/j.advwatres.2007.02.003.

McMahon, T., *et al.*, 2007c. Global streamflows—part 1: characteristics of annual streamflows. *Journal of Hydrology*, 347, 243–259. doi:10.1016/j.jhydrol.2007.09.002.

McMahon, T., *et al.*, 2007d. Global streamflows—part 2: reservoir storage-yield performance. *Journal of Hydrology*, 347, 260–271. doi:10.1016/j.jhydrol.2007.09.021.

Phien, H., 1993. Reservoir storage capacity with gamma inflows. *Journal of Hydrology*, 146, 383–389. doi:10.1016/201C-1694 (93)90285-H.

Rippl, W., 1883. Capacity of storage reservoirs for water supply. *Minutes of Proceedings of the Institution of Civil Engineers*, 71, 270–278.

Salas, J.D. and Yevjevich, V., 1972. *Stochastic structure of water use time series*, Fort Collins, CO: Colorado State University, Hydrology Papers CSU; 52.

Secretaria dos Recursos Hídricos do Ceará, 1992. *Plano Estadual Dos Recursos Hídricos*. Fortaleza: s.n.

Sobol, I.M., 1994. *A primer for the Monte Carlo method*. London: CRC Press.

Studart, T. and Campos, J., 2001. Incertezas nas estimativas da vazão regularizada por um reservatório. *Revista Brasileira De Recursos Hídricos*, 6, 81–94.

Sudler, C.E., 1927. Storage required for the regulation of streamflow. *Transactions of the Institution Society of Civil Engineers*, 91, 622–660.

Svanidze, G., 1980. *Mathematical modelling of hydrological series for hydroeletric and water resources computations*. Fort Collins, CO: s.n.

Thomas, H. and Fiering, M.B., 1962. Mathematical synthesis of streamflow sequences for analysis of basins by simulation. *In*: A. Maas, M.M. Hufschmidt, R. Dorfman, *et al.*, eds. *Design of water resources system*. Cambridge, MA: Havard University Press, 459–453.

Tyralis, H., Koytsoyiannis, D., and Kozanis, S., 2013. An algorithm to construct Monte Carlo confidence intervals for an arbitrary function of probability distribution parameters. *Computational Statistics*, 28 (4), 1501–1527. doi:10.1007/s00180-12-0364-7.

Vicens, J., Rodriguez-Iturbe, I., and Schaake, J., 1975. A Bayesian Framework for the use of regional information in hydrology. *Water Resources Research*, 11 (3), 405–414. doi:10.1029/WR011i003p00405.

Vieira, V.P., 2001. Análise de riscos. *In*: J.N.B. Campos and T.M.C. Studart, eds. *Gestão Das Águas: Princípiose Práticas*. Porto Alegre: ABRH; s.l.: s.n., 147–158.

Vogel, R., Fennssey, N., and Bolognese, R., 1965. Storage-reliability-resiliense_yield relations for the Northeastern United States. *Journal of Water Resources Planning and Management*, 31 (3), 365–374.

Yevjevich, V., 1972. *Probability and statistics in hydrology*. Littleton, CO: Water Resources Publication.

Yevjevich, V. and Harmancioglu, N.B., 1985. Past and future of analysis of water resources time series. *JAWRA Journal of the American Resources Association*, 21, 625–633. doi:101111/j.1752-1688.1985.tb05377.x.