Probabilistic Model for the Representation of the Reservoir Water Level of Concrete Dams During Normal Operation Periods

Renato Pereira[1] · António Lopes Batista[2] ·
Luís Canhoto Neves[3]

Abstract The dam's reservoir water level varies over the year according to the water inflow and outflow, as a function of environmental events but also of dam exploitation management policy and human decisions. In the Portuguese dam safety regulation (RSB 2007), the normal water level (NWL) is considered as the optimum exploitation level. However, as proved by the continuous monitoring over the lifetime of a set of dams, the NWL is occasionally exceeded for non-negligible time periods. The reservoir water level, to which the water pressure on the upstream face is related, is a fundamental parameter for the safety and reliability analysis of concrete dams. When water-induced actions are considering the leading loads, only the maximum reservoir water level, usually associated with a high-return-period flood, is relevant. However, for other combinations, in particular, earthquake scenarios, the consideration of the variability of the water level over time is crucial. In reliability analysis of concrete dams, the reservoir water level has been considered either as a deterministic variable, once the loading scenario analyzed assume water-induced actions as leading loads (Westberg 2010), or as a random variable defined using hydrological site information of a specific study case (Altarejos et al. Structural Safety 36-37:1–13 2012). This work proposes a probabilistic model of the reservoir water level of any new dam based only on its geometrical properties, which provides a low-cost alternative to in-depth hydrological analysis. The proposed model can be useful in two stages of the lifecycle of dams: (i) initial design and feasibility stages, and (ii) routine safety assessment of existing structures, as in

✉ Renato Pereira
  r.pereira@campus.fct.unl.pt

[1] Departamento de Engenharia Civil, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Quinta da Torre, 2829-516, Caparica, Portugal

[2] Departamento de Barragens de Betão, Laboratório Nacional de Engenharia Civil, Av. Brasil 101, 1700-066, Lisboa, Portugal

[3] Department of Civil Engineering, University of Nottingham, University Park, Nottingham NG7 2RD, UK

both stages the costs of a complete hydrological analysis is too high for the level of detail required. For that, the recorded reservoir water level of 27 Portuguese large concrete dams is used. A normalized sinusoidal model, with annual period, is adjusted to the reservoir water level annual history of those dams by beta regression. Generally, a good agreement between observations and the proposed model, for most of the annual adjustments, was achieved. The distribution parameters of the random variables were estimated through the maximum likelihood estimation (MLE) method. The physical, model and statistical uncertainties were quantified and can now be included in a reliability analysis procedure.

1 Introduction

The main water loads acting on concrete dams are the hydrostatic and uplift pressures. Both are functions of the reservoir water level, even though the uplift pressures also depend on the rock mass characteristics and treatment works undertaken, namely waterproofing and drainage systems.

Dams are built for energy production, water supply, irrigation, and flood control, among others. They are artificial barriers located in the course of rivers that create reservoirs upstream from them. Most operation rules are designed to minimize current water supply deficit or maximize profit (You 2008). For that reason, the reservoir water level varies over the year depending on the construction purpose. In dams built for water supply and irrigation, the reservoir water level is kept high enough during spring and summer in order to ensure the water supply in dry epochs. In dams built for energy production, following the national energy policy and the Portuguese dam operation guidelines (NEB 1993), when it is possible (reversible hydroelectric power plants), the water shall be pumped back to upstream during the night, when the energy consumption is lower. On the other hand, although the water inflow into the reservoir is greater during winter, when the energy demand is also higher, the reservoir water level is frequently controlled, as a safety measure, in order to sustain the water income of a possible flood. Moreover, if the discharge system is not able to discharge, in safe conditions, the water inflow during a flood, the reservoir water level will rise.

Briefly, the reservoir water level varies over the year according to the water inflow and outflow, dependent on environmental events but also with dam exploitation management policy and human decisions.

In the context of new dams, in which reservoir water level recordings are naturally not available, the tendencies observed in similar dams can be used to define a prediction model of the reservoir water level, only taking into account the dam maximum height and normal water level (NWL). This model can also be seen as an *a priori* prediction model of this quantity.

The continuous monitoring of the reservoir water level of large concrete dams show that the NWL, which is considered in the regulation (RSB 2007) as the optimum exploitation level, is occasionally exceeded over the year and, sometimes, for a non-negligible time period. It has also shown a distinct behavior in dams associated with run-of-river projects and dams built for storage purposes. In run-of-river projects, little or no water storage is provided. The water inflow is discharged toward downstream and the reservoir water level is almost constantly close to the NWL. In storage dams the water is retained in a reservoir for its further use, preventing floods and high demand periods. In this case, the reservoir water level may vary considerably over the year.

Therefore, meeting the recommendation of JCSS (2001) that larger populations shall be divided into sub-populations, when clear behavior differences are noticed, in order to better study and distinguish variability within a population, data was divided into two data sets, corresponding to dams built for storage purposes (STO) and dams integrated in run-of-river projects (ROR).

Rather than predicting the annual maximum reservoir level for a specific dam, which is done by specific hydrological site analysis, this study proposes a probabilistic model to predict the reservoir water level, under normal operation conditions. Thus, the reservoir water level, in normal operation conditions, of any dam can be probabilistically predicted, prior to its operation, using only geometrical information, namely the dam maximum height and the NWL. This is fundamental for the safety analysis of dams when water induced actions are not the leading loads, such as, in particular, earthquake scenarios.

For that, a beta distribution is used, as it is considered that in normal operation conditions the reservoir water level do not overtop the dam crest (upper bound), and a sinusoidal model, representing the distribution mean over the year, is adjusted to the reservoir water level annual history of several Portuguese large concrete dams by beta regression.

2 Reservoir Water Level Monitoring

Following the dam observation and inspection guidelines (NOIB 1993), the dam behavior monitoring, which has been one of the Portuguese National Laboratory for Civil Engineering (LNEC) activity fields, provides useful information that can be used to identify unexpected events. Among other variables, the reservoir water level is periodically recorded. The measurements may be performed daily, weekly or monthly using a limnimetric scale located at the reservoir or continuously by automated monitoring systems. Accuracies of 0.1 m are usually obtained in both situations (Mata 2013).

The reservoir water level is usually monitored since the dam construction. Data from 27 Portuguese concrete large dams was provided by LNEC, under permission of dam's owners, for its use to model the reservoir water level variation over the year.

3 Reservoir Water Level Model

According to the Portuguese dam safety regulation (RSB 2007), for dam structural analysis under normal operation conditions, mean value of the reservoir water level and variations around it may be considered.

A sinusoidal model, with annual period, is proposed to quantify the reservoir water level (H_R) variation over the year. This function is given by,

$$H_R(t) = H_m + S_A \cdot \cos(t - \phi) \qquad (1)$$

where H_m is the annual mean water level, S_A is the semi-amplitude or the peak deviation from the mean value, t is time and ϕ is the phase. This type of function, extensively used to model, for instance, temperature variations (e.g. Silveira 1961) over the year, predicts the maximum value of the reservoir water level at ϕ (considering $t = 0$ the beginning of the year).

As mentioned, the reservoir water level history of a set of Portuguese large concrete dams was compiled. Naturally, those dams have different dimensions which does not allow to draw generalized conclusions about the annual variation trend of the reservoir water level. Therefore, an alternative normalized model shall be considered. This model, representing

the ratio between the reservoir water level and the maximum dam height of each dam, k_R, is given by,

$$k_R(t) = \frac{H_R(t)}{H_{max}} = \frac{H_m}{H_{max}} + \frac{S_A}{H_{max}} \cdot \cos(t - \phi) = k_m + k_{sa} \cdot \cos(t - \phi) \quad (2)$$

where k_m and k_{sa} are the ratio between the annual mean level (H_m) and the semi-amplitude (S_A), respectively, and the maximum dam height (H_{max}).

4 Beta Regression Analysis

The traditional regression analysis is based on the least-square method (LSM) and, consequently, on the assumption that the dependent variable is normally distributed, being the prediction model of k_R, in terms of t, given by,

$$k_R | t \sim N(\mu = k_m + k_{sa} \cdot \cos(t - \phi), \sigma) \quad (3)$$

where μ and σ are the mean and standard deviation, respectively, of the predictive model.

However, once this model represents a proportion, in order to adjust it to data, LSM cannot be used, since its assumptions are often violated when the dependent variable is a proportion (Paolino 2001). In those situations, data have been transformed, by logit, arcsine square root, or other link function, so that it assumes values in the real line. However, this procedure has drawbacks, once typically the LSM assumptions are still not verified and the model cannot be easily interpreted in terms of the original response (Ferrari and Cribari-Neto 2004). An alternative regression model, in which the dependent variable is beta-distributed, designated beta regression, has been proposed to model rates or proportions (Paolino 2001; Ferrari and Cribari-Neto 2004; Kieschnick and McCullough 2003).

The beta density function is given in terms of two positive shape parameters, α and β, as,

$$f_X(x, \alpha, \beta) = \frac{x^{\alpha-1} \cdot (1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1} \cdot (1-u)^{\beta-1} \, du} \quad (4)$$

Since this is a continuous distribution defined on the interval $]0, 1[$, it is appropriate to model proportions.

In this case, the prediction model of k_R, in terms of t, is given by,

$$k_R | t \sim B(\alpha, \beta) \quad (5)$$

where α and β must be such that the model expected value, $E[k_R]$, matches the Eq. 2. According to the definition of beta distribution, the expected value of k_R, in terms of the shape parameters, is given by,

$$E[k_R] = \frac{\alpha}{\alpha + \beta} \quad (6)$$

The model variance contain information about the adjustment residuals, which is a measure of the model uncertainty (JCSS 2001) associated with the level of approximation in the physical formulation of the problem (Faber 2005), and is given by,

$$V[k_R] = \frac{\alpha \cdot \beta}{(\alpha + \beta)^2 \cdot (\alpha + \beta + 1)} \quad (7)$$

Inverting (6) and (7), the shape parameters of the beta distribution can be obtained from the expected value and variance, i.e.,

$$\alpha = \frac{E\left[k_R\right]^2 \cdot \left(1 - E\left[k_R\right]\right)}{V\left[k_R\right]} - E\left[k_R\right] \tag{8}$$

$$\beta = \frac{E\left[k_R\right] \cdot \left(1 - E\left[k_R\right]\right)^2}{V\left[k_R\right]} - \left(1 - E\left[k_R\right]\right) \tag{9}$$

The beta regression (Ferrari and Cribari-Neto 2004) uses the maximum likelihood estimation (MLE) method to estimate the unknown parameters associated with the model expected value (k_m, k_{sa}, ϕ) and variance ($V\left[k_R\right]$).

The likelihood function is given by,

$$L = \prod_{i=1}^{n} f_X\left(x|\theta\right) \tag{10}$$

Table 1 Main characteristics of dams considered in this study

| Name	H_{max}	H_{nwl}	Year	Project type
Alqueva	96 m	94 m	2003	Storage
Alto Lindoso	110 m	109 m	1991	Storage
Alto Rabagão	94.1 m	94 m	1964	Storage
Belver	30 m	28.65 m	1952	Run-of-river
Bouçoais-Sonim	43 m	36 m	2004	Run-of-river
Cabril	132 m	129 m	1954	Storage
Caldeirão	39 m	33.9	1993	Storage
Carrapatelo	57 m	48.5 m	1972	Run-of-river
Castelo do Bode	115 m	111.7 m	1951	Storage
Catapereiro	37.5 m	34.9 m	1999	Storage
Crestuma-Lever	65 m	52.5 m	1985	Run-of-river
Ferradosa	33.4 m	31 m	2005	Storage
Fratel	48 m	35 m	1973	Run-of-river
Fronhas	62 m	56.1 m	1985	Storage
Olgas	34.5 m	30.25 m	2007	Storage
Pedrógão	43 m	33.7 m	2006	Storage
Penha Garcia	25 m	22 m	1979	Storage
Pocinho	49 m	35.5 m	1982	Run-of-river
Pretarouca	28.5 m	24.6 m	2007	Storage
Raiva	36 m	33 m	1981	Run-of-river
Rebordelo	35.5 m	28 m	2005	Run-of-river
Régua	41 m	33.5 m	1973	Run-of-river
Touvedo	42.5 m	37.5 m	1993	Run-of-river
Valeira	48 m	40 m	1975	Run-of-river
Varosa	76 m	75 m	1976	Storage
Venda Nova	97 m	96 m	1951	Storage
Vilarinho das Furnas	94 m	93.5 m	1972	Storage

where n is the sample size, f_X is the probability density function ((4) for beta distribution), and θ is the distribution parameters (α and β for beta distribution).

The maximization of the likelihood function (or its natural logarithm as it is often convenient), provides the best estimators of the unknown distribution parameters.

The proposed model was adjusted to the reservoir water level recorded during 650 operation years of 27 Portuguese large concrete dams. Each operation year provides one observation of the unknown parameters, k_m, k_{sa}, ϕ, and $V[k_R]$. Only reservoir water levels recorded during normal operation periods were considered.

The relevant characteristics of the dams considered in this study, namely the dam maximum height (H_{max}), the height corresponding to the NWL (H_{nwl}), the year of completion and the project type, are synthesized in Table 1.

As expected, in dams associated with run-of-river projects, the reservoir water level is approximately constant over the year leading to an imperceptible sinusoidal oscillation (see Fig. 1b), i.e with reservoir water level permanently close to NWL and, consequently, semi-amplitude is close to 0. On the other hand, reservoir water level variations over the year were noticed in storage dams (see Fig. 1a). The error associated with the sinusoidal oscillation assumption is small in most situations. The obtained values for phase, ϕ, vary widely, showing that the annual maximum reservoir water level can be recorded at any time of the year, depending, as mentioned, on several aspects concerning the dam exploitation management policy.
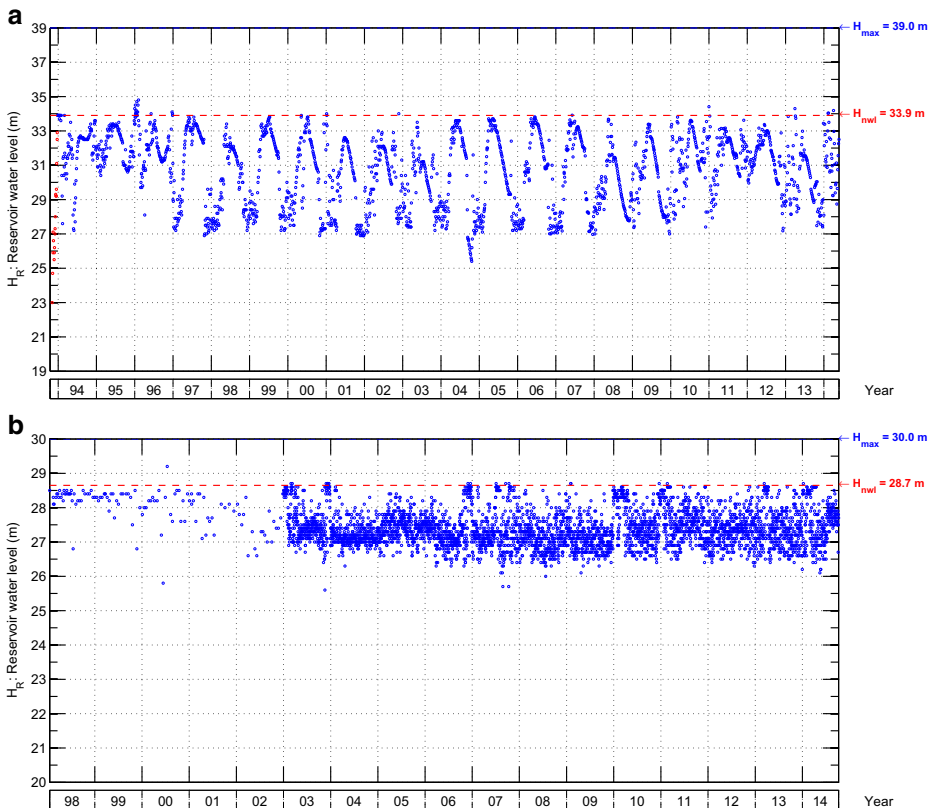
Fig. 1 Reservoir water level history

5 Uncertainty Quantification

The values of each model parameter were divided into two data sets, representing storage dams (373) and dams associated with run-of-river projects (277).

The uncertainty quantification of the model expected value and variance are presented in the sections below.

5.1 Model Expected Value

Since the NWL is the optimum exploitation level (RSB 2007), it is considered that, in average, the reservoir water level cannot be above the NWL, i.e.,

$$E\left[k_R\left(t\right)\right] \leq k_{nwl} \Leftrightarrow \frac{k_m + k_{sa} \cdot cos\left(t - \phi\right)}{k_{nwl}} \leq 1 \tag{11}$$

where k_{nwl} is the ratio between H_{nwl} and H_{max}.

Therefore, being a proportion of k_{nwl}, the expected value of the proposed model is considered a beta-distributed process, i.e.,

$$\frac{E\left[k_R\left(t\right)\right]}{k_{nwl}} \sim B\left(\alpha_E\left(t\right), \beta_E\left(t\right)\right) \tag{12}$$

where $\alpha_E\left(t\right)$ and $\beta_E\left(t\right)$ are the beta shape parameters of the expected value of the model proposed.

The parameter ϕ identifies when, over the year, the maximum reservoir water level occur, which is dependent on the dam construction purpose. However, since this parameter is not
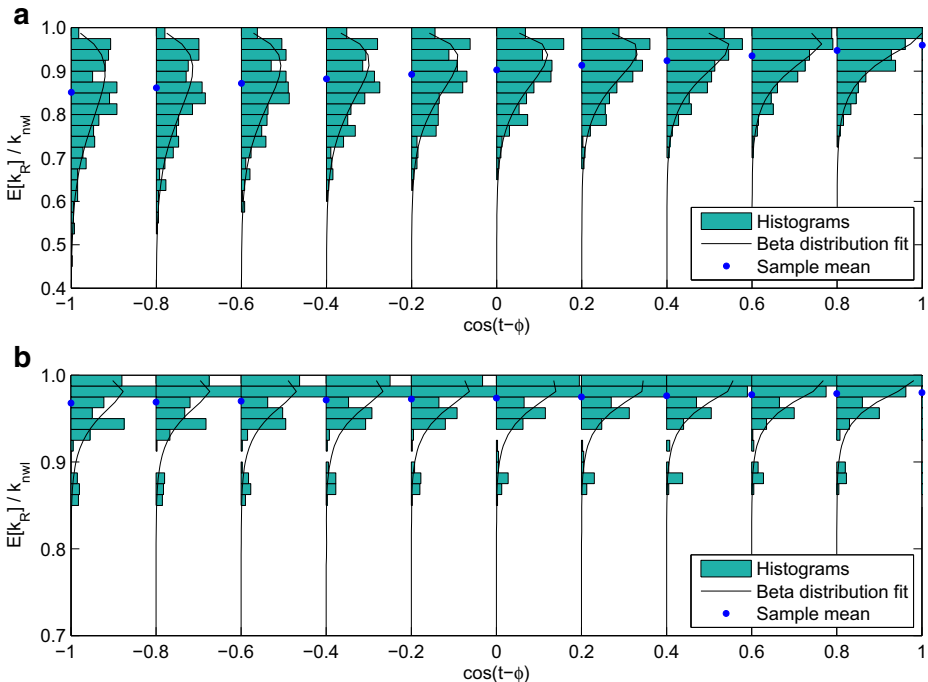
Fig. 2 Histogram and beta fit of $E\left[k_R\right]/k_{nwl}$ for different values of $cos\left(t - \phi\right)$

relevant to quantify the annual variation of the reservoir water level, its statistical treatment is omitted.

Using the pairs of values $[k_m, k_{sa}]$ obtained from the beta regression analysis, values of $E[k_R]/k_{nwl}$ for different values of $cos(t - \phi)$ were computed. The histograms were plotted against $cos(t - \phi)$ (see Fig. 2a, for storage dams, and Fig. 2b, for run-of-river dams).

The best estimator and the covariance of the beta distribution parameters, α_E and β_E, for different values of $cos(t - \phi)$, were obtained by the MLE method. Since the true value of these parameters are unknown, due to statistical uncertainty, but they must be necessarily positive, a log-normal distribution $(LN(\mu, \sigma))$, whose mean value and standard deviation match the best estimator and the corresponding standard error, respectively, is used. A 4th-order-polynomial function $(a_0 + \sum_{i=1}^{4} a_i \cdot cos(t - \phi)^i)$ was adjusted to the parameters of the log-normal distribution. The coefficients of the polynomial functions are synthesized in Table 2.

5.2 Model Variance

The model variance, $V[k_R]$, can be interpreted as a measure of the daily variations around the model mean which, as proved by the monitoring, can be above the NWL.

Since the beta distribution parameters, α and β, must be necessarily positive, from Eqs. 8 and 9, the following condition must be met:

$$V[k_R] \geq E[k_R] \cdot (1 - E[k_R]) \Leftrightarrow \frac{V[k_R]}{E[k_R] \cdot (1 - E[k_R])} \geq 1 \qquad (13)$$

Therefore, being a proportion of $E[k_R] \cdot (1 - E[k_R])$, the proposed model variance is considered a beta-distributed variable, i.e.,

$$\frac{V[k_R]}{E[k_R] \cdot (1 - E[k_R])} \sim B(\alpha_V, \beta_V) \qquad (14)$$

where α_V and β_V are the beta shape parameters of the proposed model variance.

The best estimator and covariance of the beta distribution parameters, α_V and β_V, were obtained by the MLE method. Again, since the true value of these parameters are unknown, due to statistical uncertainty, but they must be necessarily positive, a log-normal distribution $(LN(\mu, \sigma))$, whose mean value and standard deviation match the best estimator and the corresponding standard error, respectively, is used. The lognormal distribution parameters are shown in Table 3.

Table 2 Coefficients of the polynomial functions that quantify the lognormal distribution parameters, representing the statistical uncertainty of α_E and β_E, in terms of $cos(t - \phi)$

	STO				ROR			
	α_E		β_E		α_E		β_E	
	μ	σ	μ	σ	μ	σ	μ	σ
a_0	−0.094	0.0090	−0.1017	0.0006	0.0226	−0.0012	0.0291	−0.0006
a_1	−0.1805	0.0100	−0.2224	0.0036	−0.0101	0.0002	−0.0072	−0.0011
a_2	−0.1363	−0.0027	−0.2608	0.0052	−0.0757	0.0044	−0.0985	0.0029
a_3	0.3866	0.0062	−0.1703	0.0091	0.0771	−0.0021	−0.1320	0.0045
a_4	2.7580	0.0913	0.6836	0.0747	3.8295	0.0937	0.4395	0.1093

Table 3 Beta distribution properties of $k_{v,r}$

Project type	Parameter estimators	Parameter distribution
STO	$\hat{\alpha_V} = 0.75$	$\alpha_V \sim LN\,(\mu = -0.288, \sigma = 0.038)$
	$\hat{\beta_V} = 55.45$	$\beta_V \sim LN\,(\mu = 4.016, \sigma = 0.059)$
ROR	$\hat{\alpha_V} = 0.68$	$\alpha_V \sim LN\,(\mu = -0.379, \sigma = 0.116)$
	$\hat{\beta_V} = 533.27$	$\beta_V \sim LN\,(\mu = 6.279, \sigma = 0.097)$

The proposed model variance is greater in STO dams than in ROR dams. As mentioned, in ROR dams the reservoir water level remains almost constant over the year and the adjusted model with small semi-amplitude values predicts well the reservoir water level annual variation.

## 6 Examples. Comparison of the Model with Recorded Data

The 75% and 95% prediction intervals for the reservoir water level of Caldeirão (storage) and Belver (run-of-river) dams are shown in Fig. 3a and b, respectively. In these examples, the model phase was considered equal to the average phase of all years.
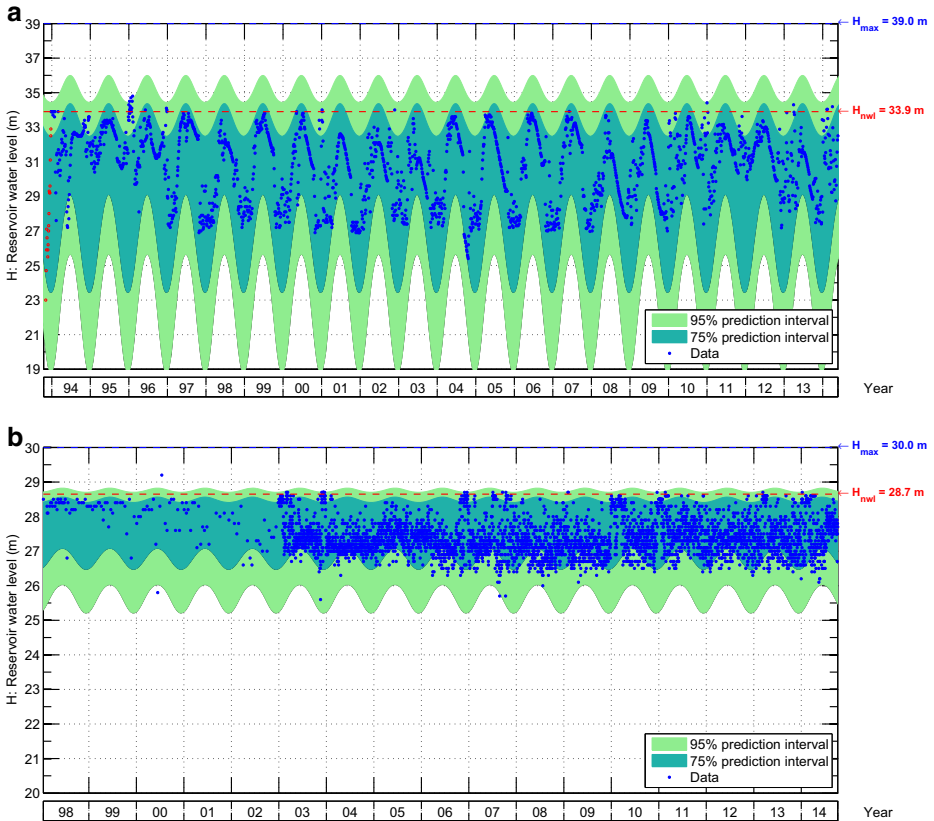
Fig. 3 75% and 95% prediction intervals for the reservoir water level

Alternatively, Fig. 4a and b compare the reservoir water level, recorded during normal operation periods, of Caldeirão (storage) and Belver (run-of-river) dams, respectively, with the prior distribution of the reservoir water level obtained from the proposed model, approximately, by overlap virtual samples generated using the parameters presented in Tables 2 and 3, for equally-spaced time intervals.

As predicted, the variability of the reservoir water level is greater in storage than in run-of-river dams. The prior distribution, obtained by the procedure presented in this paper, using the reservoir water level recorded in several dams, is naturally wider than data recorded in a particular dam. This distribution can be considered in the reliability analysis of new dams, whereas, for existing dams, it should be updated, obtaining more reliable representations of the reservoir water level uncertainty.
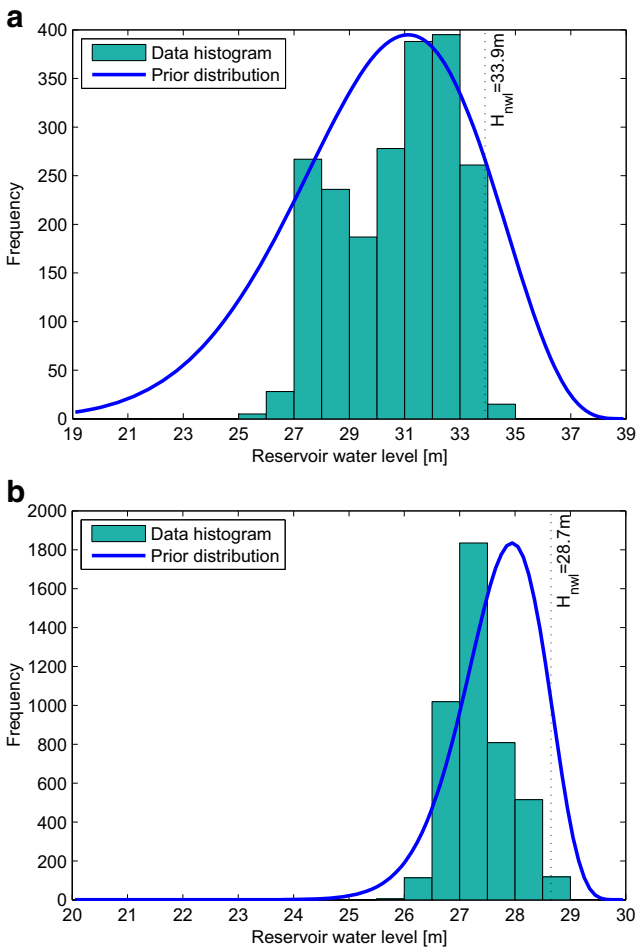
Fig. 4 Recorded data and prior distribution of the reservoir water level

7 Conclusions

Due to both environmental events and the dam exploitation management policy, the reservoir water level varies over the year. In the Portuguese dam safety regulation (RSB 2007), NWL is considered as the maximum reservoir water level in normal operation conditions. However, as proved by the continuous monitoring of this quantity over the lifetime of a set of dams, it is possible that the reservoir water level exceeds the NWL for a non-negligible time period.

The reservoir water level, to which the water pressure on the upstream face is related, is a fundamental parameter for the safety and reliability analysis of concrete dams. When water-induced actions are considering the leading loads, only the maximum reservoir water level, usually associated with a high-return-period flood, is relevant. However, for other combinations, in particular, earthquake scenarios, the consideration of the variability of the water level over time is crucial.

In this work, different than what was done in other works on the reliability analysis of concrete dams in which the reservoir water level was considered either as a deterministic variable (Westberg 2010) or as a random variable defined using specific hydrological site information (Altarejos et al. 2012), it is proposed a probabilistic model of the reservoir water level of any new dam based only on its geometrical properties. This model provides a low-cost alternative to in-depth hydrology analysis which can be useful in two stages of the lifecycle of dams: (i) initial design and feasibility stages and (ii) routine safety assessment of existing structures, as in both stages the costs of a complete hydrology analysis is too high for the level of detail required.

For that, the recorded reservoir water level of 27 Portuguese large concrete dams is used. A normalized sinusoidal model, with annual period, is adjusted to the reservoir water level annual history of those dams by beta regression. Generally, a good agreement between observations and the proposed model, for most of the annual adjustments, was achieved.

This procedure was tested and the 95% and 75% confidence intervals, obtained by Monte Carlo simulation, were overlapped to the reservoir water level recordings of two dams. Also the prior distribution of the reservoir water level obtained from the proposed model was compared to the recorded data.

As predicted, the variability of the reservoir water level is greater in storage than in run-of-river dams. The prior distribution, obtained from the procedure presented in this paper, using the reservoir water level recorded in several dams, is naturally wider than data recorded in a particular dam. This distribution can now be considered in the reliability analysis of new dams, whereas, for existing dams, it should be updated, obtaining more reliable representations of the reservoir water level uncertainty.

References

Altarejos L, Escuder I, Serrano A, Membrillera M (2012) Methodology for estimating the probability of failure by sliding in concrete gravity dams in the context of risk analysis. Structural Safety 36-37:1–13. https://doi.org/10.1016/j.strusafe.2012.01.001

Faber MH (2005) Risk and Safety in Civil, Surveying and Environmental Engineering. In: Lecture Notes, Swiss Federal Institute of Technology. ETHZ, Zùrich

Ferrari S, Cribari-Neto F (2004) Beta regression for modelling rates and proportions. J Appl Stat 31(7):799–815. https://doi.org/10.1080/0266476042000214501

JCSS (2001) Probabilistic model code - part 1: Basis of design. Technical report, Joint Committee on Structural Safety

Kieschnick R, McCullough BD (2003) Regression analysis of variates observed on (0, 1): percentages, proportions and fractions. Stat Modell 3(3):193–213. https://doi.org/10.1191/1471082X03st053oa

Mata JT (2013) Structural safety control of concrete dams aided by automated monitoring systems. PhD thesis, Instituto Superior Técnico

NEB (1993). Dam operation guidelines (in Portuguese). Proposta de portaria

NOIB (1993) Dam observation and inspection guidelines (in Portuguese). Decreto-Lei 847/1993

Paolino P (2001) Maximum likelihood estimation of models with beta-distributed dependent variables. Political Anal 9(4):325–346

RSB (2007) Dam safety regulation (in Portuguese). Decreto-Lei 344/2007

Silveira AF (1961) Temperature variations in dams (in Portuguese). LNEC

Westberg M (2010) Reliability-based assessment of concrete dam stability. PhD thesis, Lund University

You JY (2008) Hedging rule for reservoir operation: How much when and how long to hedge. PhD thesis, University of Illinois