Experimental repetitions and blockage of large stems at ogee crested spillways with piers

Paloma Furlan, Michael Pfister, Jorge Matos, Conceição Amado & Anton J. Schleiss

Research paper

Experimental repetitions and blockage of large stems at ogee crested spillways with piers

PALOMA FURLAN, PhD student, *Laboratory of Hydraulic Constructions (LCH), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. Instituto Superior Técnico (IST), Universidade de Lisboa, Lisbon, Portugal*
Email: paloma.furlan@epfl.ch (author for correspondence)

MICHAEL PFISTER (IAHR Member), Research and Teaching Associate and Professor, *Laboratory of Hydraulic Constructions (LCH), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; Civil Engineering Department, Haute École d'Ingénierie et d'Architecture de Fribourg (HEIA-FR, HES-SO), Fribourg, Switzerland*

JORGE MATOS (IAHR Member), Professor, *CERIS, Instituto Superior Técnico (IST), Universidade de Lisboa, Lisbon, Portugal*

CONCEIÇÃO AMADO, Professor, *Department of Mathematics and CEMAT, Instituto Superior Técnico (IST), Universidade de Lisboa, Lisbon, Portugal*

ANTON J. SCHLEISS ⬤ (IAHR Member), Full Professor, *Laboratory of Hydraulic Constructions (LCH), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*

ABSTRACT
Large wood is often transported by rivers into reservoirs during heavy rainfall events. When a critical section like a spillway is blocked and discharge capacity reduced, an uncontrolled increase of the reservoir water level may occur. This study aims to statistically analyse the importance of repetitions for the accuracy of experimental campaigns when studying blocking probabilities at ogee crested spillways equipped with piers. Systematic and reliable estimations based on physical models are critical for developing preventive measures against large wood blockage. Two statistical methods have been described and applied to calculate confidence intervals. A minimum number of repetitions for a maximum acceptable error is recommended for blocking probabilities. The minimum number of experimental repetitions has been statistically justified in accordance with a reasonable use of resources for experimental campaigns. In addition, a maximum acceptable level of error is proposed as a common metric of accuracy in large wood studies.

Keywords: Large wood; spillway blockage; physical model; Bernoulli trials; confidence intervals

1 Introduction

Reliable spillway operation depends on the knowledge of the characteristics of large wood (LW) at its source, on the process of its generation and transportation to the reservoir, and its potential effect on spillway capacity (Hartford et al., 2016). LW consists of stems longer than 1.00 m and greater than 0.10 m in diameter (Braudrick, Grant, Ishikawa, & Ikeda, 1997; Ruiz-Villanueva, Piégay, Gurnell, Marston, & Stoffel, 2016; Wohl et al., 2016).

Few studies focus on the fact that drifting debris, such as wood, leaves, aquatic vegetation, and other buoyant objects, may clog water-intake or water-release structures. The US National Research Council reported in 1983 that about 30% of spillway malfunctions in the USA were caused by spillway blockage due to floating debris (National Research Council, 1983). In 2015, the National Performance of Dams Program reported 23 accidents at US dams involving debris blockage of spillways (Hartford et al., 2016). Excessive LW accumulations can block flow sections at bridges, weirs or spillways and limit their functionality during extreme flood events, when discharge capacity is mostly needed (Ettema, Arndt, Roberts, & Wahl, 2000; Godtland & Tesaker, 1994; Marche, 2009). By reducing flow capacity, sedimentation processes can also be enhanced given the increasing upstream water levels, thus the potential for upstream flooding may be substantially raised (Lyn, Cooper, & Yi, 2003).

To assess LW interactions with hydraulic structures, simplified physical models have been implemented (Comiti, Lucía, & Rickenmann, 2016). The behaviour and interactions between

LW and structures involves random processes, and so, an experiment can result in different outcomes even though it is performed in the same manner systematically (Montgomery & Runger, 2011). Experiments should be repeated multiple times in order to have statistically sound results, but currently literature gives no guidance regarding the number of repetitions. The purpose of this study was to analyse the influence of test repetitions for the estimation of blocking probabilities when performing experiments in physical models of drifting stems. Hence, for the present study a probabilistic approach was chosen to define the importance of repetitions by means of confidence intervals and accuracy of estimations.

2   LW behaviour in physical models

Model tests are an invaluable tool when designing structures exposed to appreciable amounts of LW. Physical models can reproduce the behaviour of drifting wood except for the mechanical breakage of the stems under hydrodynamic loads (Hartung & Knauss, 1976). Nevertheless, model testing with LW is challenging due to its intrinsic uncertainties, scaling and diversity of governing parameters.

An experimental protocol based on a significant number of test repetitions is of primary importance to obtain statistically sound conclusions on LW dynamics and blocking processes, as a consequence of the complex and stochastic interactions observed (Schmocker & Hager, 2013; Welber, Bertoldi, & Tubino, 2013). The complexity and stochasticity inherent to wood recruitment processes, transportation, and deposition leaves deterministic models as unsuitable tools (Rigon, Comiti, & Lenzi, 2012). Reliable procedures of experimentation to estimate accurate probabilities need to be undertaken, linking statistical existing tools with engineering practice.

Previously, statistical accuracy in experimental campaigns had only been examined in Schmocker and Hager (2011) where an evaluation of number of repetitions was performed. Recent research of Schalko (2017) proposed a maximum standard deviation limit of 0.10 to define the necessary number of experimental repetitions. In spite of their relevance to the topic, these works do not provide general methods or guidelines to be used by others. Like any other statistical tryouts, LW experiments provide estimations of reality (Piton & Recking, 2016), where uncertainty and data scattering are present (Schmocker & Hager, 2013). The scattering of results can be correlated to the accuracy of such estimation and thus, it should be treated with caution.

In the experimental campaigns to analyse LW movement and transport parameters inside streams, Braudrick et al. (1997) and Braudrick and Grant (2000) performed five trials per experiment. Braudrick and Grant (2001) sought to obtain accurate and reliable results minimizing statistical errors by repeating each experiment a maximum of 10 times. Bocchiola, Rulli, and Rosso (2006a,b) used three repetitions per experiment to

Table 1   LW experiments in contact with hydraulic structures, past research

Author	Subject of study	Repetitions
Godtland and Tesaker (1994)	Ogee crested spillway design against LW clogging	?
Johansson and Cederström (1995)	LW blocking probabilities at gated ogee spillway	?
Lyn et al. (2003)	LW accumulations at bridge pier	16 to 50
Bocchiola, Rulli, and Rosso (2008)	LW accumulation patterns	4
Schmocker and Hager (2011)	LW blocking probabilities for bridges	8
Shrestha, Nakagawa, Kawaike, Baba, and Zhang (2011)	LW and debris interaction for slit-check dam	?
Hartlieb (2012)	LW jam at gated ogee crested spillway	20
Schmocker and Hager (2013)	LW accumulation at debris rack	3
Pfister, Capobianco, Tullis, and Schleiss (2013)	LW blocking probabilities at piano key weir	25 to 50
De Cicco, Paris, and Solari (2016)	LW accumulation at bridge piers	10
Gschnitzer, Gems, Mazzorana, and Aufleger (2017)	LW blocking process at bridge piers	8
Schalko (2017)	LW accumulation probability at bridge pier	40

improve accuracy and minimize human errors when estimating entrainment of LW in rivers. Welber et al. (2013) repeated experiments 10 times to study bed morphology and LW dispersal.

Table 1 illustrates the repetitions found in literature for studies of LW interacting with hydraulic structures. Godtland and Tesaker (1994) used physical experiments to define construction recommendations for an overflow spillway. These guidelines, to avoid blocking probabilities higher than 10–20% of the drifting trees, were later used in Galeati (2009), Hartlieb (2015), Johansson and Cederström (1995), Wallerstein, Thorne, and Abt (1996) and Wallerstein and Thorne (1995), among others. However, their methodology of experiments and the respective analysis was not documented.

Blocking probabilities were evaluated in Johansson and Cederström (1995) for an overflow spillway. The main parameters tested were the ratio of stem length to the bay opening and the vertical clearance between a bridge and the spillway. Fifty stems were tested, individually or in pairs. It is not clear if each stem was considered as one experiment or if a stem would represent one repetition of an experiment for the case of individual blockage. The same question arises in the case of pairs, that is, did an experiment consist of providing two stems simultaneously and was it repeated 25 times? For the case of piano key weirs Pfister, Capobianco, et al. (2013) studied individual LW

blocking probabilities. Repetitions ranged between 25 and 50 times.

Bocchiola et al. (2008) used four repetitions to predict wood accumulation patterns that can be used for evaluation of hazard at hydraulic structures, giving a simplified overview of the multifaceted process of LW jams with a statistical approach. To describe LW characteristics and their effect on the jam shape and on the discharge capacity at blocked overflow spillway inlets, Hartlieb (2012) repeated experiments 20 times. In Hartlieb (2015) the LW jamming process and its effects were applied to the design of a debris rack.

Shrestha et al. (2011) studied the interaction between debris flows and LW for slit-check dams with a numerical and physical model. Experiments were repeated many times (unknown quantity) under the same initial conditions. A relation was obtained for blocking probability at a slit-check dam as a function of the number of stems reaching the structure simultaneously.

When estimating blocking probabilities at bridge decks, Schmocker and Hager (2011) performed preliminary experiments seeking a compromise between test effort and accurate probability interpretation. It was noted that from eight to 30 repetitions, no improvement in the statistical accuracy was achieved although the statistical analysis was not documented. A risk assessment tool was defined by using their blocking probability equation to estimate, before a flood, if blockage of single logs or rootstock would occur at a bridge. Gschnitzer, Gems, Aufleger, Mazzorana, and Comiti (2015) and Gschnitzer et al. (2017) used eight repetitions for a statistical assessment on bridge clogging process with LW, associated to flood risk management.

Gschnitzer et al. (2017) provided a probabilistic description of LW related clogging mechanisms at bridges that contributes to a better comprehension of LW interactions. With their work, LW blocking probabilities can be estimated for bridges including geometrical, hydraulic and LW characteristics in one formula. De Cicco et al. (2016) defined 10 repetitions per experiment to investigate LW accumulation at historical bridge piers. For congested transport regime and five different pier shapes, the capacity of LW blocking for each pier shape was tested and compared. Lyn et al. (2003) aimed to have higher accuracy with the experiments using a maximum of 50 repetitions. However, it was concluded that 50 repetitions were not enough to obtain so-called "stable results", in order to understand the physical processes involved in single-pier debris accumulation at bridge crossings. Schalko (2017) performed 300 repetitions of a single experiment to determine a statistical reliable number of repetitions. Under the tested conditions, it was observed that after 40 repetitions the standard deviation calculated was less than 0.10 (maximum limit defined by the authors). Other combinations of parameters tested required 60 repetitions to reach the desired standard deviation.

A sensitivity analysis of repetitions for head increase experiments due to LW blockage of a debris rack was made in Schmocker and Hager (2013). It was verified if repetitions of

tests resulted in the same relative flow depth upstream. For that purpose, it was found that repeating each experiment three times was appropriate.

The above described experimental methods seem to give only an order of magnitude of LW related processes (Piton & Recking, 2016). The intrinsic variability of LW requires enough repetitions to clearly identify causal relationships, as any other estimation of a random parameter. A statistically justified number of repetitions per experiment in the analysis of LW interactions with hydraulic structures will guide towards reliable and accurate blockage assessment. When dealing with probability estimations, in this case LW blocking, the tools given in statistical inference should be used. LW blockage processes will depend of different LW characteristics, the hydraulics of the stream and the involved structure. This estimated probability should be sufficiently accurate and that accuracy can be evaluated by standard error calculations or confidence intervals. In LW-related literature, methods to quantify the accuracy of results were not found. The innovation of this paper is therefore to link statistical tools, to calculate accuracy of estimations, with LW research.

3 Experimental set-up and methodology

3.1 *Physical model*

Experiments were conducted at the Laboratory of Hydraulic Constructions (LCH) of the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. The flume is a straight channel of 1.50 m width, 0.70 m high and 10 m long and has a glass side wall to allow visual observation. Water was supplied through a tank upstream of the channel. A baffle was placed 2.40 m downstream of the channel inlet to assure homogeneous flow (Fig. 1). Water velocity was measured in a cross-section 4 m upstream the model using a flow probe (in total 36 points were measured, every 0.20 m along the channel width, four vertical points separated by 0.16 m) to verify that the velocity field was uniform where the LW was supplied. This was performed for three different water discharges.

The model was situated perpendicular to the flow and included an ogee crested spillway, four round nose piers, and two half piers for completion at the walls of the channel (creating five bays of width $b = 0.26$ m, Fig. 2a). A metallic beam held the piers above the spillway and was attached outside of the flume. The spillway and piers were fabricated from PVC and they were considered hydraulically smooth.

The ogee weir was designed with standard profiles (US Army Corps of Engineers, 1990) for a design head (H_d) of 0.15 m and weir height of 0.42 m (Fig. 2b). This type of weir was chosen due to its effective discharge capacity and frequent application in practice.

Using a round pier nose, three different configurations were tested focussing on the intrusion of the nose in the upstream face of the spillway, namely (Fig. 2a):
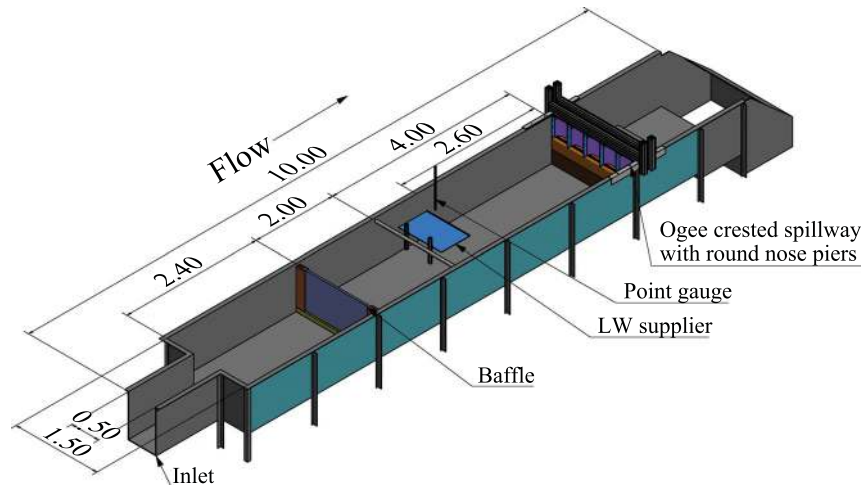
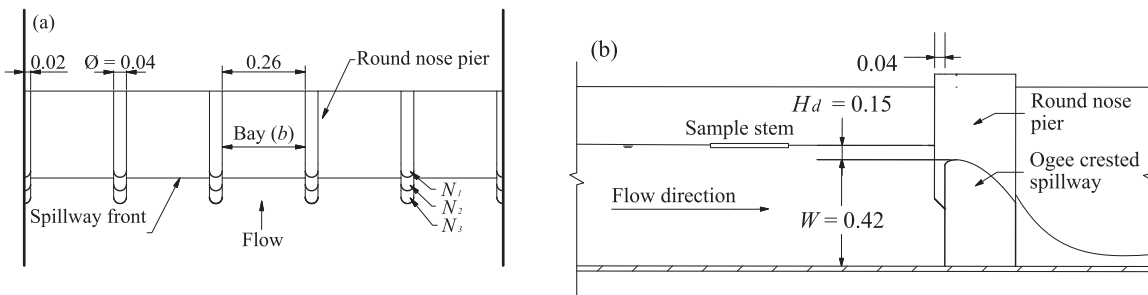Figure 1   3D representation of the channel (dimensions in m)

Figure 2   Schematic representation of the model built (dimensions in m). (a) Top view of the physical model. (b) Longitudinal view of the channel

N_1 : Aligned with the front of the spillway
N_2 : Protrudes 0.04 m from the weir spillway into the reservoir
N_3 : Protrudes 0.08 m from the weir spillway into the reservoir

Furthermore, the number of open bays was varied using vertical gates. For the experiments of five bays, all the gates were fully opened. For one bay scenarios, only the central bay was opened and the four lateral gates were fully closed.

3.2   *Model stems*

Artificial cylindrical plastic stems were chosen to represent LW in the model. The idealized shape excludes geometrical irregularities. Based on the work of Gschnitzer et al. (2015) it can be assumed that, if the geometry of the stems is more complex, the blocking probability may increase. Idealized stems were used to reduce the number of parameters involved in this complex problem. Although the geometry of LW is an important variable, the objective is to quantify the influence of the number of repetitions on the accuracy of blocking probability estimations. Regardless of the idealized or natural shape used to represent LW, the proposed methods for calculating confidence intervals are valid if results can be considered Bernoulli-type experiments.

Stems were separated in classes A to E according to their length L and diameter d (Table 2). For each class, different stem density intervals were defined based on observed data of average dry wood density along Europe and its standard deviation

Table 2   Artificial stems used (model dimensions)

Class	Length L [m]	Diameter d [m]	L/b [%]	Stem density ρ_s [–]
A	0.210	0.010	80	0.59
				0.79
				0.99
B	0.260	0.012	100	0.56
C	0.300	0.016	120	0.43
				0.56
				0.97
D	0.400	0.020	150	0.63
E	0.520	0.025	200	0.40
				0.54
				0.76
				0.99

(Chave et al., 2009). The average dry wood density for Europe taken from a sample of 77 different trees, is $\rho_w = 0.525\,\mathrm{t\,m^{-3}}$ and its standard deviation $\sigma_w = 0.119$ (subscript w stands for wood). Stem's densities were normalized with respect to water density ρ and identified herein with subscript s. Stems were separated into four categories of density: $\rho_{s1} = [0.40 - 0.47]$; $\rho_{s2} = [0.47 - 0.67]$; $\rho_{s3} = [0.67 - 0.88]$; $\rho_{s4} = [0.88 - 0.99]$.

The length range of the stems was chosen in order to cover different relative stem lengths L/b compared to bay opening, but keeping constant $L/d \simeq 20$ as seen in field observations after the 2005 flood in Switzerland (Bezzola & Hegg, 2007). Class A

Table 3 Table of experiments

$N°$	Class	Density ρ_s [–]	H/d [–]	$N°$ of open bays	Nose type	Repetitions	Supply angle [°]
1	E	0.54	0.82	5	N1	60	90
2	E	0.54	1.08	1	N1	40	90
3	C	0.56	0.81	5	N1	60	90
4	E	0.54	0.72	5	N1	60	90
5	E	0.76	0.76	5	N1	20	90
6	E	0.76	1.08	5	N1	70	0
7	B	0.56	1.25	5	N1	60	45
8	D	0.63	0.90	5	N1	70	0
9	E	0.54	0.88	1	N2	58	0
10	B	0.56	1.25	1	N2	70	90
11	B	0.56	0.83	5	N3	60	45
12	E	0.40	0.72	5	N3	60	135
13	C	0.43	0.81	5	N3	70	90
14	A	0.79	0.80	5	N3	60	135

with a stem length of 80% of the bay width is of particular interest since it was recommended by Godtland and Tesaker (1994) as a minimum bay width for spillways in contact with LW.

3.3 *Methodology*

Different flow conditions on the ogee were established by varying the inflow discharge. Ratios of head over the crest to stem diameter (H/d) ranged from 0.72 to 1.25. These limits were defined based on preliminary experiments. A reservoir flow type was analysed, implying close to zero flow velocity. The maximum model discharge used for the repetition analysis was $0.011\,\mathrm{m^3\,s^{-1}}$.

The surface tension effect on the rating curve is negligible if the flow depth exceeds some $0.025\,\mathrm{m}$ (Ettema et al., 2000) and $0.02\,\mathrm{m}$ particularly for standard ogee weirs (Breitschneider (1978), in Pfister, Battisacco, De Cesare, and Schleiss (2013)). Pfister, Battisacco, et al. (2013) state that a head of $0.015\,\mathrm{m}$ generates an error of 5% only in terms of discharge coefficient at a piano key weir. Herein, heads ranging from 0.008 to $0.027\,\mathrm{m}$ were tested. The rating curve of some tests was accordingly potentially influenced by viscosity. However the evaluation of accuracy in stems blockage probability estimation was our focus and not the determination of rating curves for the ogee crested spillway.

Fourteen experiments were defined testing all parameters in random combinations (Table 3). Thus, a wide range of results was assured. The angle of supply was changed testing four different possibilities, taking 0° as parallel to the flow. It was seen that the angle of supply was not influential in the stems movement as they were observed to be capable of self-pivoting and, normally, stems aligned themselves with the flow direction upstream of the ogee.

For every experiment, an initial condition was defined and the flow depth was measured without the presence of any stems. Only one stem was supplied at the time and it was noted if the stem passed or blocked at the spillway inlet. In the case of blockage the stem was removed, to avoid interactions between stems, and the procedure was repeated with the same initial hydraulic conditions.

In order to reduce errors induced by human interaction, every stem was supplied in the flume centre axis, with a mechanical device guaranteeing equal conditions per repetition. The device was placed at approximately $4.00\,\mathrm{m}$ upstream of the ogee. As the centre bay was used in both scenarios (one or five open bays) to be systematic, stems were supplied only in the centre line of the flume.

3.4 *Instrumentation*

The water level in the channel was measured using a point gauge ($\pm 0.5\,\mathrm{mm}$), $2.60\,\mathrm{m}$ upstream of the physical model (Fig. 1). The discharge Q was measured with a magnetic inductive flow metre ($\pm 0.5\%$ at full span). Photographs were taken systematically in order to record each experiment.

4 Statistical analysis

An experiment was defined as one combination of parameters (Table 3) and it was composed of several repetitions or trials under constant initial conditions. The 14 experiments performed represent 818 independent results. To analyse the influence of the number of repetitions on the blocking probability estimation, each experiment was considered as a Bernoulli trial where only two outcomes were possible (blocked or passed). For each experiment, n independent repetitions (trials) were performed, where π denotes the probability of blockage and $1 - \pi$ the probability of passage.

The maximum likelihood estimator, $\hat{\Pi}$, of the blocking probability π is given by (Eq. (1)):

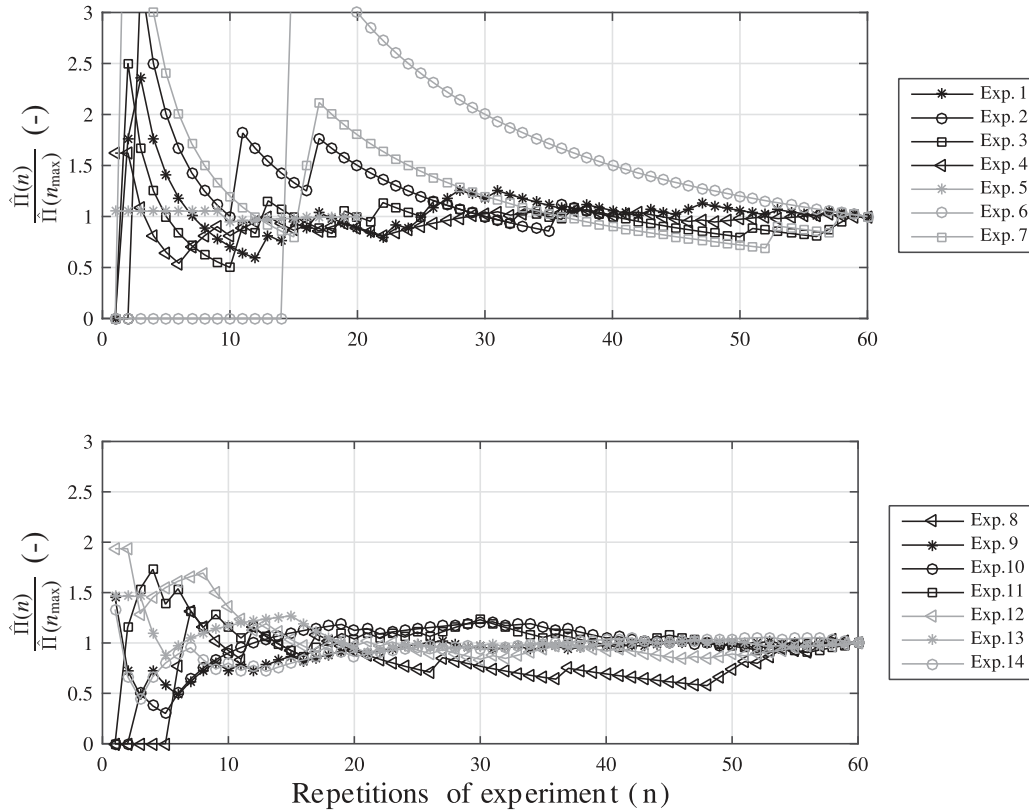$$\hat{\Pi}(n) = \frac{X}{n} \tag{1}$$

Figure 3   Normalized estimated blocking probability, in function of repetitions n

where X is the number of blocked stems, and n is the total number of repetitions. This means that the blocking probability π is estimated as the ratio of the number of stems that blocked with the number of stems that were supplied. Without loss of generality, both the estimator and the estimate will be denoted by the same symbol $\hat{\Pi}$.

The blocking probability of a stem will depend on the parameters defined in Table 3 according to the literature review (Table 1). For each experiment, results were normalized with $\hat{\Pi}$ of the last repetition. This normalization helps to examine the variation of π as a function of the number of repetitions.

Figure 3 shows how normalized blocking probabilities obtained for the 14 experiments performed varies with the repetitions. For example, the normalized estimated blocking probability for experiments 1, 2, 3, 7 and 12 with fewer than six repetitions is more than twice the value than after 60 repetitions. The overall behaviour observed in the figure is that after some 30 repetitions, the dispersion magnitude of the normalized estimated blocking probability starts to range in ± 0.10.

To infer from the experiments, with some confidence, what the real value of π might be, confidence intervals were computed. Confidence intervals allow to understand the variability related to the number of repetitions (sample size) and thus the error of the estimation.

A confidence interval states that, with a given level of "certainty", the true value will likely be in the identified range (Wallis, 2013). The width of the interval will be smaller when more accuracy in the estimation is achieved. A lower and upper

limit defines the interval estimator of π with a pre-defined confidence or coverage, and will be denoted by $[L_L; U_L]$. $[L_L; U_L]$ will have a nominal confidence level of $100(1-\alpha)\%$ (being α the error level) assuring that the interval constructed based on the sampling distribution of $\hat{\Pi}$ will contain the true value of π, that percentage of times (Pires & Amado, 2008). The random variables L_L and U_L depend on X, n and on the method of calculation.

Two common methods were applied to calculate the confidence intervals for π, the Wald method and the Clopper–Pearson method.

4.1   *Wald method*

The Wald method is the most common approach for calculating symmetric binomial confidence intervals, and is based on the approximation of the binomial by the Normal distribution. If X is binomially distributed with parameters n and π, then X has the same distribution as the sum of n independent Bernoulli random variables (Montgomery & Runger, 2011; Ross, 2010). Then, by the central limit theorem, the binomial distribution can be approximated using a standard Normal distribution as n approaches $+\infty$. A rule of thumb to well-approximate a binomial by a normal can be given by the relation: $n\pi(1-\pi) \geq 10$ (Ross, 2010). The $[L_L; U_L]$ Wald interval estimator of π was calculated according to Eq. (2):

$$\hat{\Pi} - z_{\alpha/2}\sqrt{\frac{\hat{\Pi}(1-\hat{\Pi})}{n}} \leq \pi \leq \hat{\Pi} + z_{\alpha/2}\sqrt{\frac{\hat{\Pi}(1-\hat{\Pi})}{n}} \qquad (2)$$

where $z_{\alpha/2}$ denotes the $1 - \alpha/2$ quantile of the standard Normal distribution (Agresti & Coull, 1998; Vollset, 1993). The following term, $\sqrt{\hat{\Pi}(1 - \hat{\Pi})/n}$ is called the standard error (SE) of the point estimator $\hat{\Pi}$.

Apart from being asymptotic, the Wald method may have two limitations. First, when π tends to the extremes $\{0\}$ or $\{1\}$, the product $\pi(1 - \pi)$ tends to 0, leading to an underestimation of the error. Second, the interval can exceed $[0; 1]$ limits (Agresti & Coull, 1998; Wallis, 2013).

4.2 *Clopper–Pearson method*

To overcome the normal theory approximations of Wald intervals, the Clopper–Pearson method is suggested to calculate confidence intervals for π. The Clopper–Pearson confidence interval for π with a coverage probability of at least $1 - \alpha$ can be obtained by solving Eq. (3) for L_L and U_L (Clopper & Pearson, 1934):

$$\sum_{k=x}^{n} \binom{n}{k} L_L^k (1 - L_L)^{n-k} = \alpha' \tag{3a}$$

$$\sum_{k=0}^{x} \binom{n}{k} U_L^k (1 - U_L)^{n-k} = \alpha'' \tag{3b}$$

where $\alpha' + \alpha'' = 1$. For $\alpha' = \alpha'' = \alpha/2$ these correspond to the inversion of the two sided exact binomial test and lead to the central exact interval. For $x = 0$ the solution of Eq. (3) is explicit and given by $L_L = 0$ and $U_L = 1 - (\alpha/2)^{1/n}$. For $x = n$ the solution is also explicit and given by $L_L = (\alpha/2)^{1/n}$ and $U_L = 1$. When $x = 1, 2, \ldots, n - 1$, the lower endpoint is the $\alpha/2$ quantile of a beta distribution with parameters x and $n - x + 1$, and the upper endpoint is the $1 - \alpha/2$ quantile of a beta distribution with parameters $x + 1$ and $n - x$ (Agresti & Coull, 1998). The relation between the beta distribution and the Snedecor's F distribution leads to the following result (Eq. (4)) for the Clopper-Pearson

confidence interval for π:

$$\left[1 + \frac{n - x + 1}{x F_{2x, 2(n-x+1), 1-\alpha/2}} \right]^{-1} < \pi$$
$$< \left[1 + \frac{n - x}{(x + 1) F_{2(x+1), 2(n-x), \alpha/2}} \right]^{-1} \tag{4}$$

for $x = 1, 2, \ldots, n - 1$, and $F_{a,b,c}$ denotes the $1 - c$ quantile from the F distribution with degrees of freedom a and b.

The Clopper–Pearson method provides more reliable confidence intervals with smaller samples than the Wald method (Clopper & Pearson, 1934; Sauro & Lewis, 2005). Clopper–Pearson produces conservative confidence intervals, and are therefore wider (Agresti & Coull, 1998).

5 Discussion

Figure 4 shows the results obtained from an experiment with class E, density ρ_{s1} and pier nose configuration N_3 for $H/d = 0.72$ and five bays open (experiment 12 of Table 3). It can be seen that the first and second stem were blocked but the third one passed. For example, the repetition $n = 30$ considers 13 individually provided stems that blocked, divided by 30 provided stems and results in a (estimated) blocking probability of $\hat{\Pi} = 0.43$. The last repetition performed $n = 60$, shows that 31 of 60 stems blocked at the spillway inlet, resulting in a blocking probability of $\hat{\Pi} = 0.52$. Between repetition $n = 30$ $[\hat{\Pi}(30) = 0.43]$ and $n = 60$ $[\hat{\Pi}(60) = 0.52]$, there is a difference between blocking probabilities of $\Delta \hat{\Pi} = 0.09$. If only four (like in Bocchiola et al., 2008) or 20 (like in Hartlieb, 2012) repetitions of an experiment were considered, the estimation of the blocking probability would be 0.75 and 0.50 respectively. This analysis highlights the importance of knowing the variance of the estimated $\hat{\Pi}$ and how it changes with the number of repetitions.
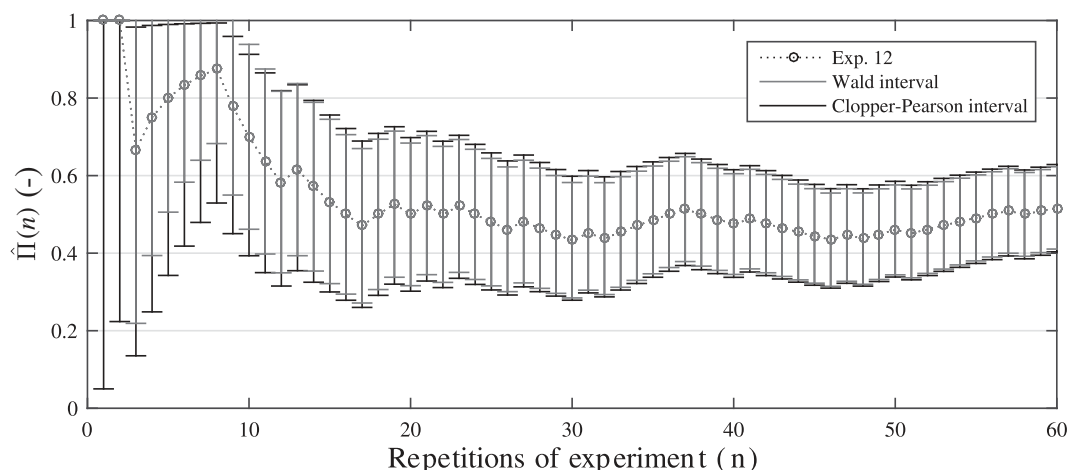

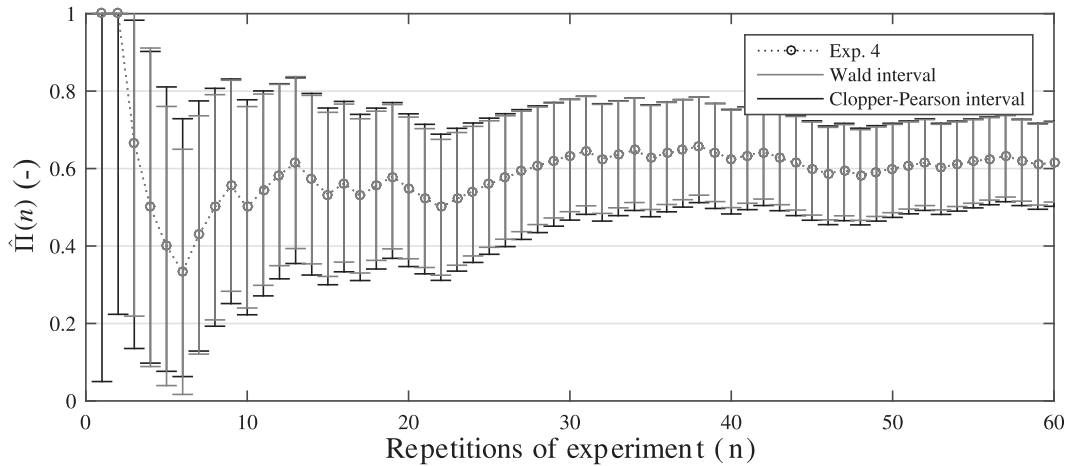Figure 4   Estimated blocking probability for experiment 12 (Table 3)

Figure 5   Estimated blocking probability for experiment 4 (Table 3)

Figure 5 shows the results obtained for experiment 4 (Table 3). In this case, the blocking probability after 30 repetitions is $\hat{\Pi}(30) = 0.63$, meaning that 19 out of 30 stems supplied blocked at the spillway inlet. For the repetition $n = 60$, $\hat{\Pi}(60) = 0.62$. Between repetition $n = 30$ [$\hat{\Pi}(30) = 0.63$] and $n = 60$ [$\hat{\Pi}(60) = 0.62$], there is a difference between blocking probabilities of $\Delta\hat{\Pi} = 0.01$ for individual stems.

For each repetition, 90% confidence intervals are presented in Figs 4 and 5 with the respective error bars. As expected, the confidence interval width reduces when n increases, meaning that the error in the estimation of π decreases, therefore the accuracy of the estimation increases. A confidence level of 95% was also evaluated but the results showed that an excessive number of repetitions ($n \geq 90$) would be needed to achieve errors smaller than 0.10 in the estimation of blocking probabilities.

With the results obtained from the physical experiments, it was shown that the blocking probability depends on the variables of Table 3 as the estimated probability of blockage changed from one experiment to the other. Nevertheless, the main focus of the present paper relies on the influence of repetitions for the accuracy of probability estimations and not on

the parameters involved in the process of blockage. The different magnitude of $\Delta\hat{\Pi}(n)$ of Figs 4 and 5 is influenced by the variables combined for those experiments, but the width of the confidence interval (considering any method) has a similar magnitude, therefore the accuracy of those results can be considered similar.

The results shown in Figs 4 and 5 were obtained in the experimental facility using augmented design. The experiments were repeated 60 times. In order to achieve a consolidated analysis, different sample sizes by subsampling bootstrap (Politis, Romano, & Wolf, 1999) have been computed for the estimated blockage probability. Figure 6 shows experiment 12 (Table 3) as an example of the error bars obtained after 1000 re-samples with different sample sizes. It can be seen that the variability in the width of the confidence interval is similar to that seen in Fig. 4 based on augmented design; as the sample size and the repetitions increased, the confidence in the estimated blocking probability increased accordingly. Nonetheless, due to the time and cost incurred per experiment in physical models, this type of analysis is only feasible numerically. Hence, the augmented design was kept as the remaining methodology.
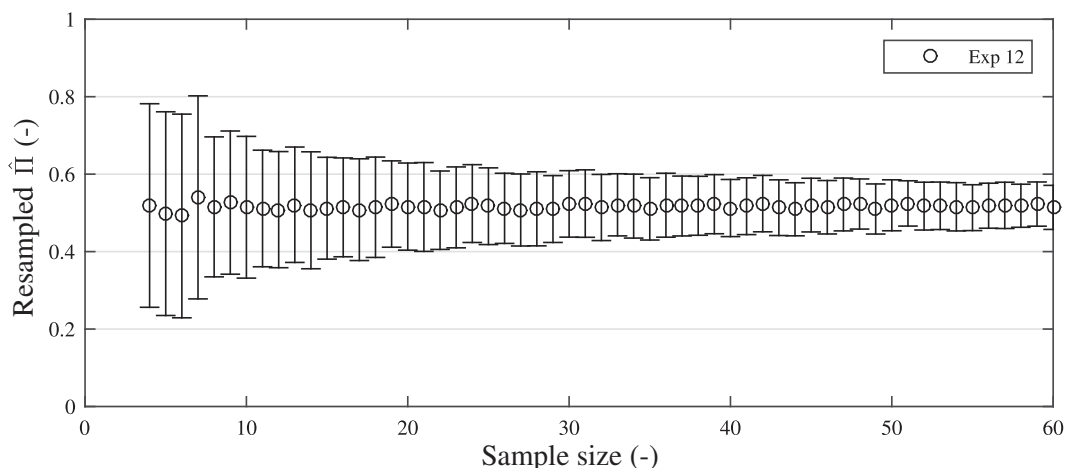

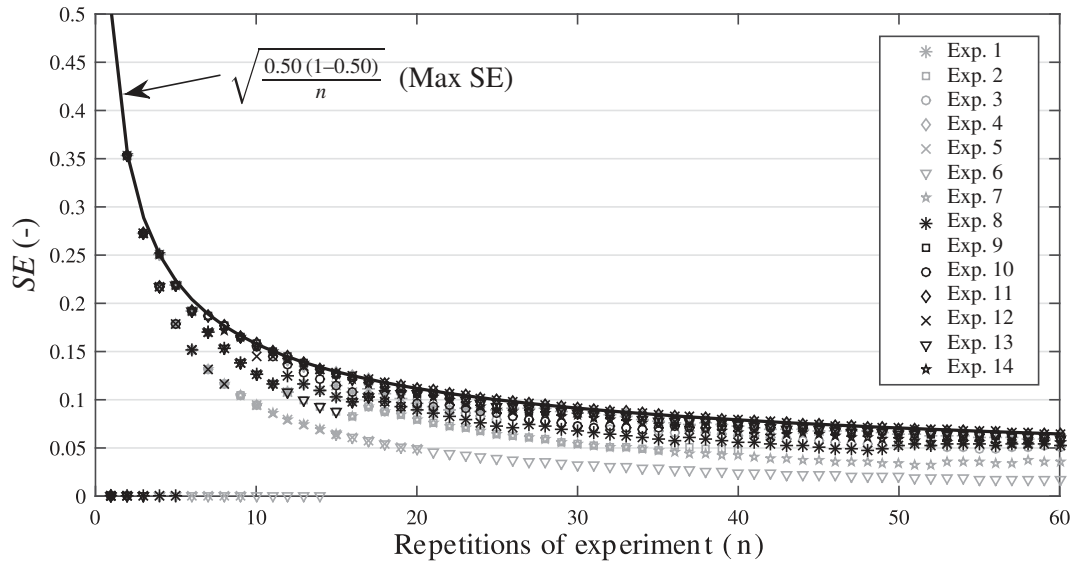Figure 6   Variability of blockage estimation for different sample sizes

Figure 7   Standard error of the point estimator

Table 4   Width of confidence intervals

		$n = 4$			$n = 10$			$n = 30$			$n = 60$	
Experiment	$\hat{\Pi}$	W	CP	$\hat{\Pi}$	W	CP	$\hat{\Pi}$	W	CP	$\hat{\Pi}$	W	CP
1	0.50	0.82	0.80	0.20	0.42	0.47	0.33	0.28	0.31	0.28	0.19	0.20
2	0.25	0.71	0.74	0.10	0.31	0.39	0.10	0.18	0.21	–	–	–
3	0.25	0.71	0.74	0.10	0.31	0.39	0.20	0.24	0.27	0.20	0.17	0.18
4	0.50	0.82	0.80	0.50	0.52	0.56	0.63	0.29	0.31	0.62	0.21	0.22
5	1.00	0.00	0.53	0.90	0.31	0.39	–	–	–	–	–	–
6	0.00	0.00	0.53	0.00	0.00	0.26	0.03	0.11	0.15	0.02	0.05	0.08
7	0.25	0.71	0.74	0.10	0.31	0.39	0.10	0.18	0.21	0.08	0.12	0.13
8	0.00	0.00	0.53	0.20	0.42	0.47	0.17	0.22	0.25	0.22	0.17	0.19
9	0.50	0.82	0.80	0.50	0.52	0.56	0.67	0.28	0.31	0.68[a]	0.19[a]	0.21[a]
10	0.25	0.71	0.74	0.60	0.51	0.55	0.80	0.24	0.27	0.67	0.20	0.21
11	0.75	0.71	0.74	0.50	0.52	0.56	0.53	0.30	0.32	0.43	0.21	0.22
12	0.75	0.71	0.74	0.70	0.48	0.52	0.43	0.30	0.32	0.52	0.21	0.23
13	0.75	0.71	0.74	0.80	0.42	0.47	0.67	0.28	0.31	0.68	0.20	0.21
14	0.50	0.82	0.80	0.60	0.51	0.55	0.70	0.28	0.30	0.75	0.18	0.20

W: Wald method. CP: Clopper–Pearson method. [a]Only 58 repetitions considered

The limitation of the Wald method when dealing with extreme blocking probabilities becomes visible (Figs 4 and 5) as the confidence interval for the first two repetitions could not be calculated. In the case of the Clopper–Pearson method, it can be seen that for $n < 10$ the width of the interval tends to be larger, hence more conservative, than the Wald interval.

The standard errors can be seen in Fig. 7, for all the replications and experiments performed. For a given n, the maximum standard error (SE) was obtained for $\hat{\Pi} = 0.50$. If $n = 60$, the maximum error in the estimation of π is about 0.06. For the case of $n = 30$, the maximum error is about 0.09. If n continues to decrease, it can take values up to 0.50 as it can be seen in Fig. 7 with the continuous line for the maximum standard error.

The width of the intervals according to both methods, for $n = 4$; 10; 30; 60 with 90% confidence is shown in Table 4. For the Wald method, when $[L_L; U_L]$ exceeded $[0; 1]$ it was taken

either 0 or 1. Experiments 5 and 6 are good examples of the limitations of the Wald method. For experiment 5, as the number of repetitions was smaller than 30 and the estimated blockage probability was close to 1, the confidence interval calculated exceeded the $[0; 1]$ interval therefore the external limits had to be taken. Experiments 6 and 7 had 70 and 60 repetitions respectively but $\hat{\Pi}$ was close to 0, leading to forced calculations of confidence intervals. The use and interpretation of the Wald method should be done thoughtfully as the calculation is always feasible but can be misleading if the data cannot be approximated by a Normal distribution.

The maximum standard error after one repetition can be of 0.50 (Fig. 7), being the worst case scenario with a blocking probability of 0.50. For a small number of repetitions $n \leq 4$, the estimation has a wide confidence interval or a significant possible difference with the real blockage probability. For example

experiment 3 has a width of 0.71 (Wald method, Table 4) with a confidence level of 90%. Thus if that experiment is performed twice again and the outcomes, after four repetitions, are $\hat{\Pi}_1(4) = 0.25$ and $\hat{\Pi}_2(4) = 0.75$, both estimations could be inside the first confidence interval although they represent different blockage scenarios. With this perspective, the results obtained in Bocchiola et al. (2008) after four repetitions can have a maximum error of 0.25 in the probability of stems being entrapped into a jam in function of their length, for example.

Increasing the number of repetitions to $8 \leq n \leq 10$ as in Schmocker and Hager (2011), De Cicco et al. (2016) and Gschnitzer et al. (2017) reduces the width of the confidence interval to almost half of the width given for $n \leq 4$ and the maximum standard error is 0.16. With $n \leq 30$ and $n \leq 60$ the width is notably reduced meaning that the difference of the estimation from experiments with the real value is small. Also, the maximum standard error is 0.09 and 0.06 respectively. Although, to repeat an experiment 60 times means a significant use of resources for achieving better accuracy.

The Wald method needs to have at least $n > 30$ and blocking probabilities different than 0 or 1 to be correctly applied. It can provide a false sense of accuracy because of slender intervals (Agresti & Coull, 1998). The Clopper–Pearson method performs better when having small amounts of repetitions and avoids the assumption of a Normal distribution. It allows to calculate confidence intervals even if the probabilities are close to the extremes (for example, experiment 8 Table 4 for $n = 4$), although it gives wider intervals to be more conservative. Wald and Clopper–Pearson methods provide a good overview of the variability of blocking probabilities related to repetitions and can be easily applied with high levels of confidence. Both are adequate for binomial results, having a simple calculation process and being present in most of statistical software. The chosen method will depend on the type of data available and if the assumption of a Normal distribution is possible. If more complex approximations are to be used, they can be found among statistical literature (Correa and Sierra, 2001; Newcombe, 1998; Pires & Amado, 2008; Sauro & Lewis, 2005; Vollset, 1993).

Thus, the achieved accuracy per experiment will be dependant on the number of repetitions. It is recommended that $n \geq 30$ is used, as it will give an estimated blocking probability with standard errors less than 0.09 for 90% confidence levels. This number of repetitions will give a statistically reliable estimation of the unknown value of the blockage probability. Furthermore, it is recommended to express the blockage probability of LW with its confidence interval and its level of confidence, as it remains a point estimator of an unknown quantity. As mentioned in Wohl et al. (2010), some common metrics are needed in complex topics such as LW behaviour.

6   Conclusion

Experiments have been conducted to evaluate the influence of repetitions on the accuracy of LW blockage probability estimations at an ogee crested spillway with piers. This research has pointed out that stem experiments require a systematic approach and common metrics. Physical models are used to study LW processes, hence a statistically justified minimum number of repetitions is valuable. Nevertheless, previous works often ignored the importance of defining this number based on statistical accuracy. As a result, the importance of accuracy in probabilistic estimations has been a minor topic. To provide a blocking probability estimation $\hat{\Pi}$ with any information of how close it is of π is inconclusive. When this point estimate is reported, it is necessary to give an order of the accuracy of that estimation.

A wide range of typical stem dimensions and hydraulic conditions were tested, simultaneously with changing parameters regarding the stems characteristics or the hydraulic structure. Experiments indicate that the variability of the estimated blocking probability has a strong relation with the number of repetitions per experiment. Herein, the application of mathematical tools related to the problem of LW behaviour by determining an interval estimate for a binomial proportion is analysed. Knowing the error in an estimation is the first step in inferential statistics and it allows calculation of how reliable an observation is, without the need for further sampling (Wallis, 2013). Based on the results and the statistical methods presented, it is recommended for experimental campaigns to make $n \geq 30$ repetitions per experiment so that estimations of blockage probabilities with errors smaller than 0.09 occur (with 90% confidence). In accordance with Schalko (2017), the maximum acceptable error of the experimental estimations should be 0.09 to have rigorous assessments of LW blockage risk. This number of repetitions can be decreased if less accuracy of results is tolerable. Blocking experiments with fewer than 10 repetitions are not recommended as they have large scattering of results and large width of confidence intervals. Confidence intervals are valuable methods for estimating accuracy of observations and they should be included when referring to an estimated probability, with their implicit confidence level. As a common metric for calculating confidence intervals, the Wald or Clopper–Pearson methods are proposed with confidence levels of 90%.

7 Notation

b	$=$	bay width (m)
d	$=$	stem diameter (m)
H	$=$	hydraulic head (m)
H_d	$=$	hydraulic design head (m)
L	$=$	stem length (m)
$L_L; U_L$	$=$	lower and upper limit of confidence interval (–)
n	$=$	number of repetitions (–)
N	$=$	nose configuration (–)
Q	$=$	discharge (m^3 s^{-1})
SE	$=$	standard error (–)
X	$=$	number of blocked stems (–)
W	$=$	weir height (m)
α	$=$	error level (–)
θ	$=$	pier nose diameter (m)
π	$=$	blocking probability (–)
$\hat{\Pi}$	$=$	estimated blocking probability (–)
$\hat{\Pi}(n_{max})$	$=$	estimated blocking probability for maximum number of repetitions (–)
ρ	$=$	water density (t m^{-3})
ρ_s	$=$	stem density (–)
ρ_w	$=$	wood density (t m^{-3})
σ_w	$=$	standard deviation wood density (t m^{-3})

ORCID

Anton J. Schleiss http://orcid.org/0000-0003-1559-5740

References

Agresti, A., & Coull, B. A. (1998). Approximate is better than "Exact" for interval estimation of binomial proportions. *The American Statistician*, *52*(2), 119–126.

Bezzola, G. R., & Hegg, C. (2007). *Ereignisanalyse Hochwasser 2005, Teil 1 - Prozessse, Schäden und erste Einordnung [Event analysis flood 2005, Part 1 - Processes, damage and first classifications]* (Technical Report No 0707). Bern: Bundesamt für Umwelt BAFU, Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft WSL.

Bocchiola, D., Rulli, M. C., & Rosso, R. (2006a). Flume experiments on wood entrainment in rivers. *Advances in Water Resources*, *29*(8), 1182–1195.

Bocchiola, D., Rulli, M. C., & Rosso, R. (2006b). Transport of large woody debris in the presence of obstacles. *Geomorphology*, *76*(1–2), 166–178.

Bocchiola, D., Rulli, M. C., & Rosso, R. (2008). A flume experiment on the formation of wood jams in rivers. *Water Resources Research*, *44*(W02408), 1–17.

Braudrick, C. A., & Grant, G. E. (2000). When do logs move in rivers?. *Water Resources Research*, *36*(2), 571–583.

Braudrick, C. A., & Grant, G. E. (2001). Transport and deposition of large woody debris in streams: A flume experiment. *Geomorphology*, *41*(4), 263–283.

Braudrick, C. A., Grant, G. E., Ishikawa, Y., & Ikeda, H. (1997). Dynamics of wood transport in streams: A flume experiment. *Earth Surface Processes and Landforms*, *22*(7), 669–683.

Breitschneider, H. (1978). Bauwerksmodelle: Abflussver-hältnisse, Energieumwandlung, Erosion [Physical models: Discharge conditions, energy dissipation, erosion]. In *Mitteilungsheft4*(DVWW), 195–216.

Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G., & Zanne, A. E. (2009). Towards a world-wide wood economics spectrum. *Ecology Letters*, *12*(4), 351–366.

Clopper, C. J., & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, *26*(4), 404–413.

Comiti, F., Lucía, A., & Rickenmann, D. (2016). Large wood recruitment and transport during large floods: A review. *Geomorphology*, *269*, 23–39.

Correa, M. J. C., & Sierra, L. E. (2001). Intervalos de confianza para el parámetro de la distribución binomial. *Revista colombiana de Estadística*, *24*(1), 59–72. [Confidence intervals for the binomial distribution parameter].

De Cicco, P. N., Paris, E., & Solari, L. (2016, September). Wood accumulation at bridges: Laboratory experiments on the effect of pier shape. In G. Constantinescu, M. Garcia, & D. Hanes (Eds.), *Proceedings of the 8th international conference on fluvial hydraulics river flow 2016* (pp. 2341–2345). St Louis, USA: CRC Press.

Ettema, R., Arndt, R., Roberts, P., & Wahl, T. (2000). *Hydraulic modeling: Concepts and practice* (Technical Report No 97). Reston: Environmental and Water Resources Institute of the American Society of Civil Engineers.

Galeati, G. A. (2009, January). Hydraulic safety of ENEL dams. In *Proceedings of the Colloque CFBR/SHF Dimensionnement et fonctionnement des évacuateurs de crues* (pp. 1–8). Lyon, France: Société Hydrotechnique de France.

Godtland, K., & Tesaker, E. (1994). Clogging of spillways by trash. *Proceedings of the commission internationale des Grands Barrages, 18th Congress of ICOLD* (pp. 543–557), Durban.

Gschnitzer, T., Gems, B., Aufleger, M., Mazzorana, B., & Comiti, F. (2015). On the evaluation and modelling of wood clogging processes in flood related hazards estimation. In G. Lollino, M. Arattano, M. Rinaldi, O. Giustolisi, J.-C. Marechal, & G. E. Grant (Eds.), *Engineering geology for society and territory* (Vol. 3, pp. 139–142). Cham: Springer International Publishing.

Gschnitzer, T., Gems, B., Mazzorana, B., & Aufleger, M. (2017). Towards a robust assessment of bridge clogging processes in flood risk management. *Geomorphology*, *279*, 128–140.

Hartford, D. N. D., Baecher, G. B., Zielinski, P. A., Patev, R. C., Ascila, R., & Rytters, K. (2016). Disturbances. In *Operational safety of dams and reservoirs* (pp. 173–219). London: ICE Publishing.

Hartlieb, A. (2012, June). Large scale hydraulic model tests for floating debris jams at spillways. In P. Rutschmann, M. Grünzner, & S. Hötzl (Eds.), *Proceedings of the 2nd IAHR Europe congress* (Paper C18). München: Taylor and Francis.

Hartlieb, A. (2015). *Schwemmholz in Fließgewässern - Gefahren und Lösungsmöglichkeiten [Driftwood in streams - dangers and solutions]* (Technical report). München: Berichte des Lehrstuhls und der Versuchsanstalt für Wasserbau und Wasserwirtschaft fur Technische Univesität München.

Hartung, F., & Knauss, J. (1976, March). Considerations for spillways exposed to dangerous clogging conditions. In *Proceedings of the commission internationale des Grands Barrages, 12th congress of ICOLD* (pp.741–749). Paris: ICOLD.

Johansson, N., & Cederström, M. (1995, July). Floating debris and spillways. In J. J. Cassidy (Ed.), *Proceedings of the international conference on hydropower - waterpower '95* (pp 2106–2115). San Francisco, California: American Society of Civil Engineers.

Lyn, D. A., Cooper, T., & Yi, Y. -K. (2003). *Debris accumulation at bridge crossings: Laboratory and field studies*. (Technical Report FHWA/IN/JTRP-2003/10). West Lafayette, Indiana: Joint Transportation Research Program.

Marche, C. (2009). Dam-break flow against obstacles and through river bed singularities. In D. De Wrachien & S. Mambretti (Eds.), *Dam-break problems, solutions and case studies* (Vol. 36, pp. 211–239). Southampton, Boston: WIT Press.

Montgomery, D. C., & Runger, G. C.. (2011). *Applied statistics and probability for engineers*. 5th ed. New York: John Wiley & Sons.

National Research Council. (1983). *Safety of existing dams: Evaluation and improvement*. Washington, D.C: National Academies Press.

Newcombe, R. G. (1998). Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Statistics in Medicine*, *17*(8), 857–872.

Pfister, M., Battisacco, E., De Cesare, G., & Schleiss, A. J. (2013). Scale effects related to the rating curve of cylindrically crested Piano Key weirs. In S. Erpicum, F. Laugier, M. Pfister, M. Pirotton, G-M. Cicero, & A. J. Schleiss (Eds.), *Labyrinth and piano key weirs II* (pp. 73–82). Boca Raton: CRC Press.

Pfister, M., Capobianco, D., Tullis, B., & Schleiss, A. J. (2013). Debris-Blocking sensitivity of piano key weirs under reservoir-type approach flow. *Journal of Hydraulic Engineering*, *139*(11), 1134–1141.

Pires, A. M., & Amado, C. (2008). Interval estimators for a binomial proportion: Comparison of twenty methods. *Revstat*, *6*(2), 165–197.

Piton, G., & Recking, A. (2016). Design of sediment traps with open check dams. II: Woody debris. *Journal of Hydraulic Engineering*, *142*(2), 04015046.

Politis, D. N., Romano, J. P., & Wolf, M.. (1999). *Subsampling*. New York: Springer Series in Statistics.

Rigon, E., Comiti, F., & Lenzi, M. A. (2012). Large wood storage in streams of the Eastern Italian Alps and the relevance of hillslope processes. *Water Resources Research*, *48*(1), W01518.

Ross, S. M. (2010). *Introduction to probability models*. 10th ed. Amsterdam: Academic Press.

Ruiz-Villanueva, V., Piégay, H., Gurnell, A. M., Marston, R. A., & Stoffel, M. (2016). Recent advances quantifying the large wood cycle and dynamics in river basins: New methods, remaining challenges. *Reviews of Geophysics*, *54*(3), 611–652.

Sauro, J., & Lewis, J. (2005, September). Estimating completion rates from small samples using binomial confidence intervals: Comparisons and recommendations. In *Proceedings of human factors and ergonomics society 49th annual meeting* (pp. 2100–2104). Orlando, USA: HFES by SAGE Publications.

Schalko, I. (2017, August). Large wood accumulation probability at a single bridge pier. In Ghani (Ed.), *Proceedings of the 37th IAHR world congress* (pp. 1704–1713). Kuala Lumpur, Malaysia: IAHR and Usains Holding SDN BHD.

Schmocker, L., & Hager, W. H. (2011). Probability of drift blockage at bridge decks. *Journal of Hydraulic Engineering*, *137*(4), 470–479.

Schmocker, L., & Hager, W. H. (2013). Scale modeling of wooden debris accumulation at a debris rack. *Journal of Hydraulic Engineering*, *139*(8), 827–836.

Shrestha, B. B., Nakagawa, H., Kawaike, K., Baba, Y., & Zhang, H. (2011). Driftwood deposition from debris flows at slit-check dams and fans. *Natural Hazards*, *61*(2), 577–602.

US Army Corps of Engineers (1990). *Hydraulic design of spillways* (Engineer Manual 1110-2-1603). Washington, USA: USACE.

Vollset, S. E. (1993). Confidence intervals for a binomial proportion. *Statistics in Medicine*, *12*(9), 809–824.

Wallerstein, N. P., & Thorne, C. R. (1995). *Debris control at hydraulic structures in selected areas of Europe*. (Technical report). Nottingham, England: University of Nottingham, Department of Geography.

Wallerstein, N. P., Thorne, C. R., & Abt, S. R. (1996). *Debris control at hydraulic structures contract modification: Management of woody debris in natural channels and at hydraulic structures* (Technical report). Vicksburg, Mississippi: US Army Corps of Engineers.

Wallis, S. (2013). Binomial confidence intervals and contingency tests: Mathematical fundamentals and the evaluation of alternative methods. *Journal of Quantitative Linguistics*, *20*(3), 178–208.

Welber, M., Bertoldi, W., & Tubino, M. (2013). Wood dispersal in braided streams: Results from physical modeling. *Water Resources Research*, *49*(11), 7388–7400.

Wohl, E., Bledsoe, B. P., Fausch, K. D., Kramer, N., Bestgen, K. R., & Gooseff, M. N. (2016). Management of large wood in streams: An overview and proposed framework for hazard evaluation. *Journal of the American Water Resources Association*, *52*(2), 315–335.

Wohl, E., Cenderelli, D. A., Dwire, K. a., Ryan-Burkett, S. E., Young, M. K., & K. D. Fausch (2010). Large in-stream wood studies: A call for common metrics. *Earth Surface Processes and Landforms*, *35*(5), 618–625.