Including Epistemic Uncertainty in Engineering Decision Making

Mariano Balbi
PhD candidate, Laboratorio de Materiales y Estructuras, Universidad de Buenos Aires, Buenos Aires, Argentina

David C. B. Lallemant
Assistant Professor, Asian School of the Environment, Nanyang Technological University, Singapore

ABSTRACT: Design problems in engineering require computing the probability of exceedance of decision variables, that can be calculated by arbitrarily complex mathematical models from basic variables and their corresponding distributions. This work tries to explore answers to an important question for engineers: How would my design change if we include uncertainty in the models used? We study the inclusion of the so called epistemic sources of uncertainty in typical design problems in civil engineering, by means of Bayesian predictive posterior estimates. Through relatively simple synthetic examples, we compare these estimates with the traditional approach. We found that the inclusion of epistemic uncertainties in the model relating basic and decision variables can yield significant impact on design estimates that strongly depend on the shape of this relatinoship and the uncertainty characterization of its parameters.

1. THE DESIGN PROBLEM

Most engineering systems are designed on the basis that their probability of under-performance is appropriately low. More sophisticated approaches require the designer to explicitly compute this probability of failure using appropriate probabilistic methodologies (Melchers and Beck (2018)). The failure condition can be generically expressed as the exceedance of a given design, or decision, metric Z over a threshold z as in Eq. (1).

$$p_f = p(Z \geq z) \tag{1}$$

In standard structural reliability theory, for example, the decision variable can be the safety margin $Z = R - S$ (where R is the resistance variable and S the demand on the system) and the threshold simply $z = 0$. In risk assessments of engineering systems, Z might well represent some impact or demand metric, like the water level at a dike from an extreme flood, or the inter-story drift ratio in a building during an earthquake. Depending on the type of decision problem and the stakeholders involved, socio-economic consequence metrics are also typical in engineering decision making, such as the repair cost of a building, or the number of people displaced from their homes.

The decision metric is, in almost every practical scenario, not directly measured. So in order to develop probabilistic models, it is usually computed as a function of a set of basic variables X that are easier to characterize from an analytic and/or empirical standpoint. This model that translates basic variables into the decision variable will be called here, the reliability model. It can range from very simple analytical structural models that translate a load into a stress, to complex environmental and socio-economic models that translate the occurrence of extreme natural events into extended spatial damage.

This model can be generically represented by a mathematical function g that depends on the basic variables X and usually, a set of model parameters

β as per Eq. (2). This parameters are all input necessary to characterize the model, outside of X.

$$Z = g(X, \beta) \quad (2)$$

Combining Equations 1 and 2 we can express the exceedance probability of z as,

$$p(Z \geq z) = 1 - F_X\left(g^{-1}(z, \beta)\right) \quad (3)$$

Where F_X is the cumulative distribution of the basic variable X, and the function g is considered monotonically increasing in X.

The above equations, however, do not take into account the temporary nature of true processes. That is, engineering systems are designed to withstand a given level of performance during a predefined lifetime T. Without losing generality, we can think that the failure condition is a function of time and all variables involved are stochastic processes. Then, the probability of failure can be computed as the probability that $Z(t) \geq z(t)$ occurs during the lifetime T. This is generally known as the 'first-passage probability' (Melchers and Beck (2018)).

In most risk and reliability analysis in civil engineering the occurrence of failures is very rare and can be considered as independent events. The Homogenous Poisson Process (HPP) appears, usually, as a reasonable model to compute the likelihood of failures over time. Under this simplifying assumption the probability of failure over lifespan T can be computed as per Eq. (4).

$$p_{fT}(z) = 1 - \exp\left(-\lambda_0 T p_f(z)\right) \approx \lambda_0 T p_f(z) \quad (4)$$

Where λ_0 is the mean rate of occurrence of extreme, potentially critic levels of X.

The definition of 'event' depends on context, but it can generically be thought of as extreme levels of X that can potentially lead to failure. Many problems naturally adjust to this discrete-events methodologies, like the consequences generated by natural hazards such as earthquakes, hurricanes or thunderstorms. Other problems might intuitively be associated to continuous time variables, such as wind loading, sea-level, or even resistance parameters of structures. In any case, continuous process can be discretized by selecting an appropriate threshold to define an 'extreme event'. For example, we are not interested in the continuous sea-level rise process, but rather in sea-level rises above 2m.

Thus, from a modelling perspective, the problem requires to estimate the distribution F_X of the basic variables for each (discrete) event, and the corresponding design level Z by means of $g(X, \beta)$ (see Figure 1 for an illustrative scheme). The engineering system is then designed so that $p_f T$ is appropriately low, or similarly so that the return period T_r is appropriately high. A very common approach is to calculate the z_p level for a given return period in order to use as a design value or to verify a given design (see Figure 1 for an illustrative scheme). For example, the height of a dike is designed for the water level with an exceedance return period of $500yrs$. Or the resistance of a structural member should be calculated so that the probability of being lower than the demand in 50 years is 1 in 2% (i.e. a return period of 2,500 years approximately). Many more examples are ubiquitous in civil engineering (Melchers and Beck (2018)).

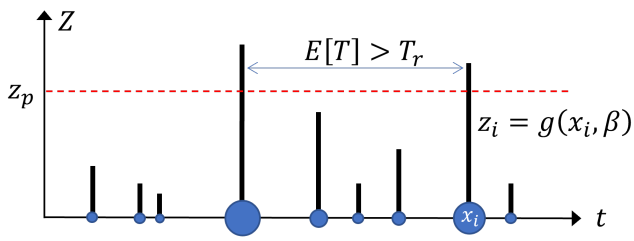

Figure 1: Qualitative scheme of the time dependent design problem in engineering

The objective of the present work is to explore how uncertainty in the construction of the models to compute this probability of exceedance impact in the design of engineering systems. This uncertainty, that stems from limited data and knowledge about the actual evolution of the system, adds a layer of complexity to the computation of Eq. (4). While its inclusion in the distribution model F_X has been conceptually addressed in the literature (Merz and Thieken (2005)), it is still not clear the relative influence of considering uncertainty in the reliability model g. The rest of this work describes how we can include these into the design problem, and how its inclusion is compared to more traditional approaches that do not account for them using simplified synthetic models.

2. UNCERTAINTIES

The engineering design problem is, as described in the previous section, tightly related to predicting an uncertain event in the future. Thus, it is strictly an uncertainty quantification task. At its core, Eq. (4) is a mathematical representation of what is known as 'aleatory uncertainty', here characterized by the HPP and the probability distribution F_X of the basic variable at any event.

Aleatory uncertainty is considered as an inherent component of the physical process and it does not depend on the amount of knowledge and information that the modeller has. However, there are other sources of uncertainty around the estimation of the probability of failure that are related to our incomplete knowledge about the engineering system being analyzed. These are commonly known as 'epistemic uncertainties' (Spiegelhalter and Riesch (2011)).

As described by Spiegelhalter and Riesch (2011), epistemic uncertainty stems mainly from (1) limited information to properly characterize the probabilistic models and (2) limited knowledge to properly describe the true physical processes through the selected models. This more operational description of epistemic uncertainties allows for a more rigorous wat of including them in the decision making process. Limited information appears in practice, as limited-length data, observation errors, or missing variables. It can, typically, be represented through uncertainty in the parameters that describe the models as data is not sufficient to perfectly identify them.

Limited knowledge, on the other hand, is usually represented as uncertain models, such as the distribution family chosen for F_X or the particular physics-based model chosen for g, and simplified hypothesis such as the HPP model. It is much harder to represent this mathematically, although it can, and has been done through model ensembles (i.e. considering and weighting many possible models) or statistical representations of model deficiencies (Kennedy and O'Hagan (2001)). Other, deeper, sources of uncertainty are also described and characterized in (Spiegelhalter and Riesch (2011)) although its inclusion in a reliability framework is not straightforward and beyond the scope of this work.

Besides the epistemological differences between the two, it is not always clear which sources of uncertainty belong to each category, and it can vary depending on the context. In any case, the most important feature that differentiates aleatory and epistemic uncertainty is the fact that the former cannot be practically reduced since it is an inherent property of the system under analysis. The epistemic uncertainty, on the other hand, can be reduced by further collecting information and improving knowledge. This distinction is crucial when allocating resources for model improvement (Der Kiureghian and Ditlevsen (2009)).

2.1. *Epistemic uncertainties in the risk integral*

The inclusion of epistemic uncertainties greatly increases the complexity of the problem from an analytical and computational standpoint. In this context, the risk metrics of Eq. (4) can be understood as conditional to a given set of models and models' parameters.

$$p_{fT}\left(z|\Theta,\mathbb{H}\right) = \lambda_0 T\left[1 - F_X\left(g^{-1}\left(z,\beta\right)|\theta\right)\right] \quad (5)$$

Where Θ include all parameters used to describe the models, and $\mathbb{H}$ represent a set of hypothesis used to build the estimate, such as the HPP.

The set of parameters Θ include the ones that characterize the probability distribution F_X, here named θ, the parameters that characterize the HPP process (i.e. λ_0), and the parameters β that characterize the reliability model. These might also include parameters that assign a probability to an ensemble of different possible models, so it can reliably represent a wide range of epistemic uncertainty sources.

One way of incorporating uncertainty regarding the values of Θ is provided by Bayesian decision theory. In this context, an appropriate estimate of the design value z should take into account the consequences of over or underpredicting its true value. Fawcett and Green (2018) discusses this in the context of return period levels for environmental extreme events, and it suggests the use of the predictive posterior return level as point estimate that incorporates epistemic uncertainty. This is obtained by integrating out the parameters Θ using their posterior distribution $p(\Theta|x)$ as per Eq. (6). In the Bayesian statistical context, is the distribution of the parameters given the available data and modeller's prior knowledge.

$$p_{fT}(z|data,\mathbb{H}) = \int_{\Theta} p_{fT}(z|\Theta)\, p(\Theta|data,\mathbb{H})\, d\Theta \tag{6}$$

3. NUMERICAL SIMULATIONS

3.1. Models and data

We use a semi-synthetic flood hazard case study to study the impact of the inclusion of epistemic uncertainties in the design process. The case study requires to design the height of a riverine levee in order to protect a certain area from flooding with a minimum probability of 10% in 10 years (i.e. a return period of 100 years).

There are no historical measurements of water heights there, but there are flow discharge measurements at a gauge station a few kilometers upstream. Flow discharge is, then, used as basic variable X for our study. Real discharge data was used, from publicly available daily data at Buscot weir, in a small reach of Thames River, obtained from the UK National River Flow Archive. The series spans from 19 years from 1980 to 1998 with some minor gaps that are not expected to affect the extreme statistics analysis to perform.

This dataset was analyzed using Peaks-Over Threshold (POT) theory to identify extreme events and characterize the probabilistic distribution of the discharges. A total of 73 extreme events were identified in 18.8yrs of data, using a threshold of $12m^3/s$ (See Figure 2) and a minimum separation of 7 days between events. The estimated mean rate of occurrences of events was $\lambda_0 = 3.9 events/year$. The selected probability distribution for the peak discharges $F_X(x|\theta)$ above the threshold was the Generalized Pareto Distribution (GPD) as the standard extreme events theory indicates (Bousquet and Bernardara (2021)).

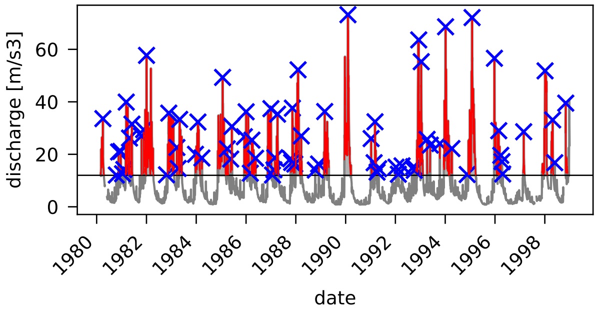

Figure 2: Daily discharge for Buscot gauge station and identification of extreme events. Blue cross indicates event's peak discharge.

The posterior distribution of the GPD shape and scale parameters (ξ and σ respectively) was obtained by Markov Chain-Monte Carlo simulation using an non-informative prior as defined in Castellanos and Cabras (2007). The 'best fit' parameters were selected as the mode of these distributions (also known as maximum a-posteriori estimates), that would be practically equal to standard maximum likelihood parameters in this case. Figure 3 shows the posterior distribution for both parameters.

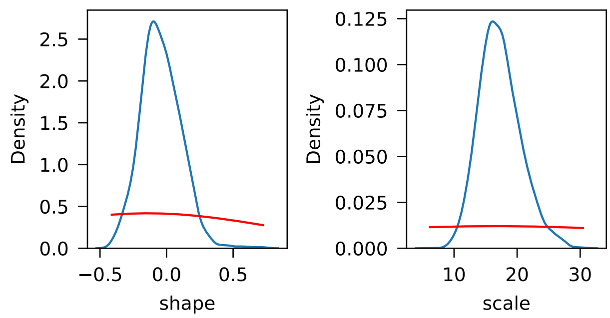

Figure 3: Prior (red) and posterior (blue) distributions of GPD parameters. Best fit parameters are $\xi^* = -0.05$, $\sigma^* = 16.5$

The model g that propagates river flow discharges downstream is called a flow routing model and is typically based on simplified versions of the 1D fluid dynamics equations (Di Baldassarre (2012)). Typical parameters of these models are the roughness of the channel, the cross-section properties, and downstream boundary conditions. For simplicity, and since this is an exploratory study, we used simple mathematical analytical functions to describe g that depend only on discharge X and a single model parameter β.

In the numerical experiments explored here, we examined two very simple formulations for $g(x,\beta)$: (1) A linear model both in x and β, and a (2) non-linear model in x but linear in β.

In this case, we defined three different curves for the posterior distributions for β: (1) a standard Normal distribution, (2) a right-skewed Normal distribution, and (3) a left-skewed Normal distribution. All three models share the same mode $\beta^* = 1$, so they basically represent the same 'best fit' model defined by $g(x,\beta^*)$. In practice, the posterior distribution could also be obtained from available data, or from expert knowledge. By any means, these actual posterior distribution should resemble a Normal, or skewed Normal. We believe, however, that these simple models can still be useful for understanding how they might impact the predictive posterior estimates.

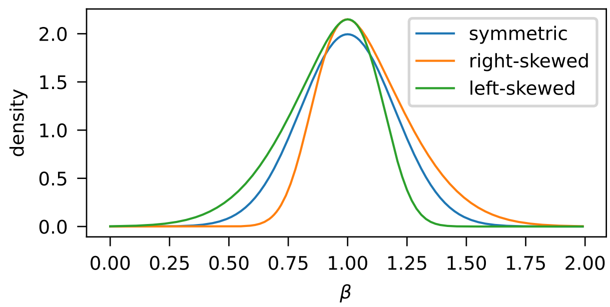

Figure 4: Three different posterior distributions for β

3.2. Simulation methodology

The integral in Eq. (6) does not have an analytic solution for most cases of practical importance. In this work we use a standard Monte-Carlo (MC) approach that simulates draws from the posterior distribution of the parameters $p(\Theta|data,\mathbb{H})$ and then compute the probability from Eq. (5) for each. The predictive posterior estimate is then numerically approximated using the average of the simulated values as in Eq. (7).

$$p_{fT}(z) \approx \frac{1}{N}\sum_{j=1}^{S} p_{fT}\left(z|\Theta_j\right) \tag{7}$$

Computing $p_{fT}\left(z|\Theta_j\right)$ from Eq. (5) is not straightforward, as well, since the forward simulator $g(x,\beta)$ does not typically have an explicit expression (i.e. it generally involves finite differences of finite element methods of resolution). Thus, inverting the model to compute the exceedance probability of z is not possible analytically.

An equivalent simulation-based expression to compute Eq. (1) is given by Eq. (8).

$$p_f = \int_x \mathbf{1}\{g(x,\beta) \geq z\}\, p(x)\, dx = E\left[\mathbf{1}\{g(x,\beta) \geq z\}\right] \tag{8}$$

Where $\mathbf{1}\{\text{cond}\}$ is a function that returns 1 when cond is true and 0 otherwise.

The exceedance distribution of level z can then be numerically approximated by the average of the simulations as per Eq. (9).

$$p_f \approx \frac{1}{N}\sum_j \mathbf{1}\left\{g\left(x_j,\beta\right) \geq z\right\} \tag{9}$$

Where x_i are random draws from $F_X(x|\theta)$ and N the number of simulations used.

The number of simulations N depends on the precision we need on the results, and on the available computational power. For higher return periods, more simulations are needed to achieve a given precision in the estimate by Eq. (9) (Bousquet and Bernardara (2021)). We used here, $N = 20,000$ and $S = 1,000$ since computational time was not an issue.

Finally, it is standard in practice to compute the mean recurrence between events instead of the exceedance probability for a given T. This is computed as per Eq. (10), and results are shown in next section using this metric.

$$T_r(z) = \left(\lambda_0 p_f\right)^{-1} \tag{10}$$

3.3. Results

3.3.1. No epistemic uncertainties

Assuming that best fit parameters are enough to perfectly characterize the probability distribution of discharges and the flow model, we can compute the design height simply by means of Eq. (4). We used the mode of the GPD parameters for the F_X model, and the mode of the β parameter for the routing model as 'best fit parameters'. In this context, it is named as the 'traditional approach'.

Figure 5 shows the exceedance curves for the discharges X, the flow model $Z = g(x, \beta^*)$, and the resulting design curve for water depths. Using this estimate, we should design the levee with a height of 10m.

3.3.2. Epistemic uncertainty in distribution of X

We consider then, the influence of the epistemic uncertainty in the distribution of the discharges, represented by the posterior distribution of GPD parameters. In this case, we compute many exceedance curves for X for each random draw of a set of $\{\xi, \sigma\}$. Figure 6 shows a small subset of the different curves simulated, together with the best fit curves (i.e. the traditional approach), and the predictive posterior estimates.

The comparison shows that the predictive posterior estimate, which includes the epistemic uncertainty, yields more conservative design values. That is, the levee height for 100 years, should be around 11m if we consider the predictive posterior estimate. Furthermore, the difference between the curves grows with return period as is expected, since more rare events are more uncertain to predict from limited-length data.

This is in line with previous studies in the inclusion of epistemic uncertainties when estimating probability models (Merz and Thieken (2005)). It is also a well-known result in Bayesian statistics that the predictive distribution of new observations has heavier tails than the underlying distribution model for the data (Gelman et al. (2013)).

3.3.3. Epistemic uncertainty in reliability model

As a next step, we considered only the epistemic uncertainty in the flow routing model, via the posterior distributions for β plotted in Figure 4. Different cases were computed, combining the linear or the non-linear model, with each of the posterior distributions for β from Figure 4.

Figure 7 shows the case of a linear model and a standard Normal distribution for β. It can be seen that the traditional approach gives the same design curve as the predictive posterior estimate. Theoretically, in this case, it will be sufficient to compute the water depths for the best β parameter only. This is likely related to the fact that, when $g(x, \beta)$ is linear in β, and thus the expected water level for a given discharge is the same as the water level for the expected discharge: $E[Z|X] = Z|E[X]$.

This, however, is not the case when $g(x, \beta)$ is non-linear. Figure 8 shows the results when using the non-linear model and the standard Normal distribution for β. There is a small departure of the predictive posterior curve from the traditional approach towards the non-conservative side in this case. The magnitude of this departure is expected to be very dependent on the non-linearity of the function. The same is expected to happen with the 'direction' of this shift.

A deviation of both design curves can also be seen when using the non-symmetric distributions for β, like the skewed Normal distributions. This will generally yield a non-symmetric distribution for $Z|X$, where the water levels for the best fit parameters will not coincide with the mean water levels, even for a linear model. Figure 9 shows that when uncertainty in β is skewed towards higher Z values, then the predictive posterior design curve will be shifted towards more conservative estimates than the traditional approach. The opposite tends to happen when the distribution for β is skewed in the opposite direction as shown in Figure 10.

The inclusion of epistemic uncertainties in both the probability model for X and the reliability model g, will be a direct combination of both the results described above.

4. CONCLUSIONS

This work intended to compare estimates of a generalizable design problem, using the traditional approach that uses 'best fit' models and a novel approach that includes epistemic sources of uncertainty. This tries to address an important question for engineers: How would my design change if we include uncertainty in the models used?

Results showed that uncertainty in characterizing the probability distribution of the basic variables X will lead to conservative estimates for design as was discussed previously by Merz and Thieken (2005).

On the other hand, results showed that it is less straightforward to determine the differences be-

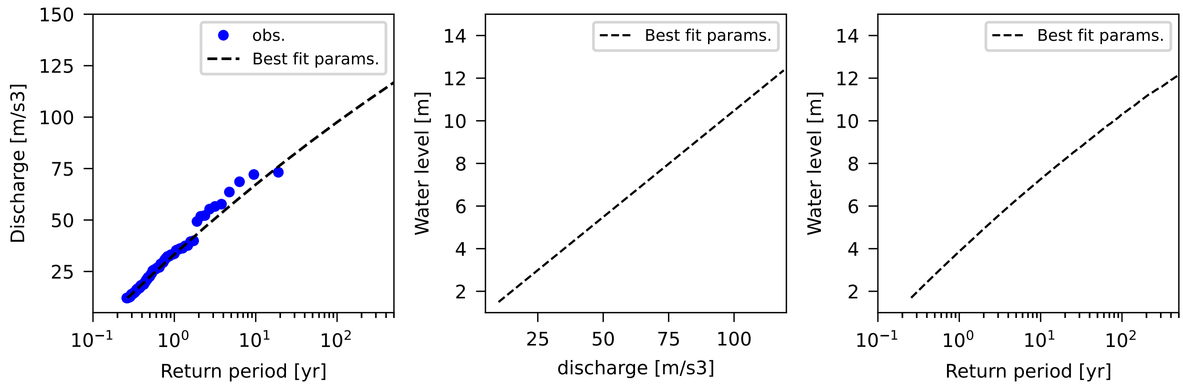

Figure 5: Inclusion of aleatory uncertainty only

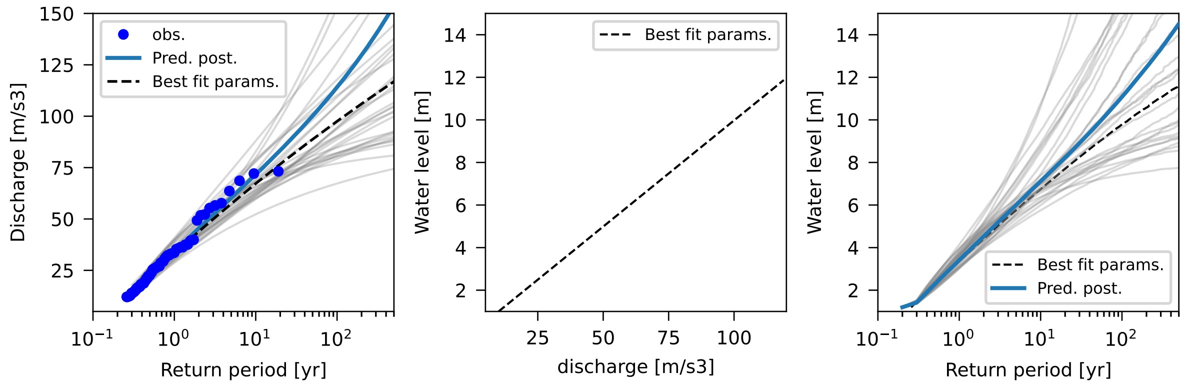

Figure 6: Inclusion of epistemic uncertainty in probability model for X

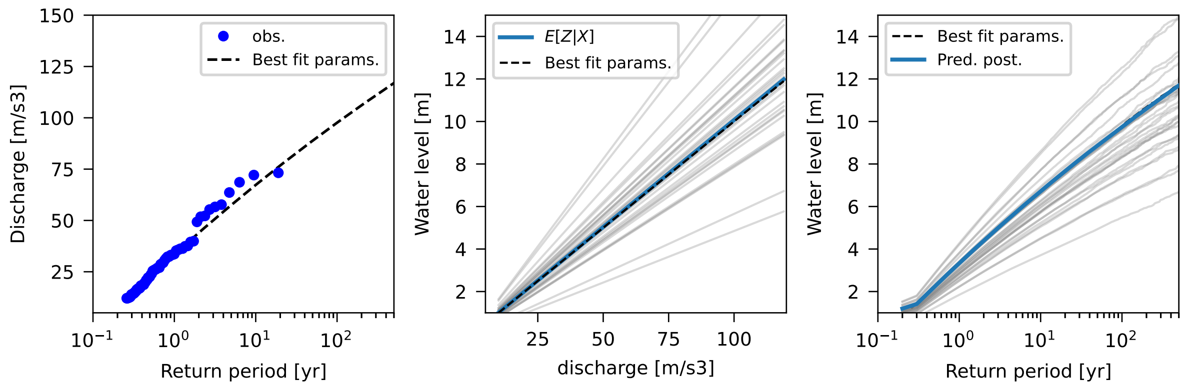

Figure 7: Inclusion of epistemic uncertainty in linear reliability model $g(x, \beta)$

tween the traditional approach and the predictive posterior estimate when including uncertainties in the reliability model g. Generalizations on this matter, seem to require further inspection of the mathematical equations defining the problem. However, the simple numerical examples ran in this work seem to imply that the specific structure of the model (e.g. non-linearity) and the posterior distribution of the parameters (e.g. calibration procedure used) can shift the traditional approach results in vastly different ways, and not necessarily on the conservative side.

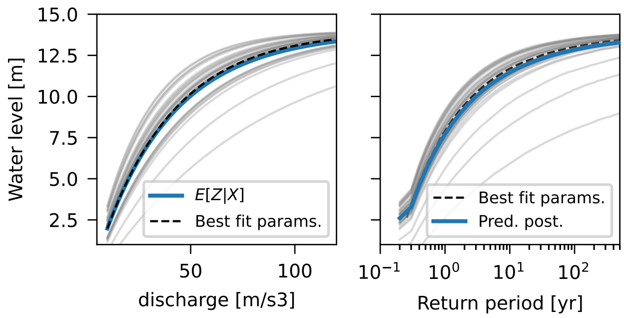

Figure 8: Inclusion of epistemic uncertainty in non-linear reliability model $g(x, \beta)$

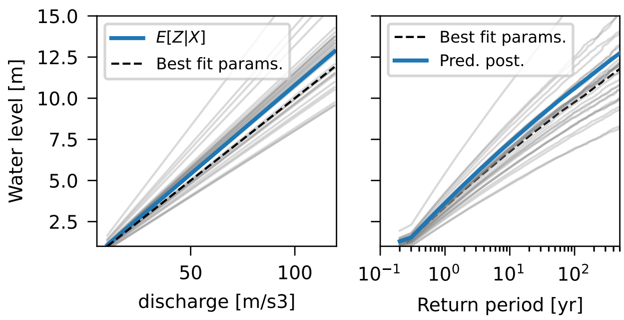

Figure 9: Inclusion of epistemic uncertainty in linear reliability model with right-skewed distribution for β

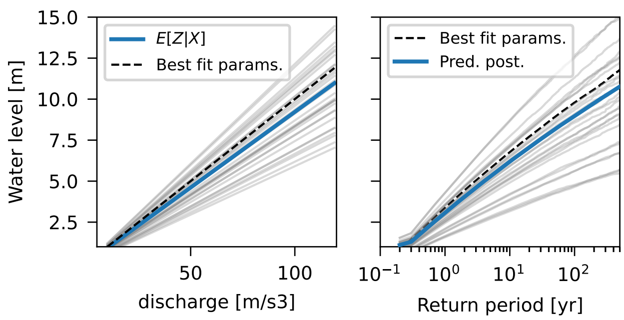

Figure 10: Inclusion of epistemic uncertainty in linear reliability model with left-skewed distribution for β

Rigorous exploration of the influence of the uncertainties in the reliability model are difficult to find in the literature, and generally problem-specific. Further studies could try to generalize some of the findings by exploring different types of non-linearities, posterior distributions for β, or simply by studying general calibration procedures for g.

All code and data used to develop the models and figures is publicly available at GitHub repository Balbi (2023).

5. REFERENCES

Balbi, M. (2023). "Including epistemic uncertainty in engineering decision making: some numerical experiments (2).

Bousquet, N. and Bernardara, P. (2021). *Extreme Value Theory with Applications to Natural Hazards: From Statistical Theory to Industrial Practice*. Springer International Publishing, Cham.

Castellanos, M. E. and Cabras, S. (2007). "A default Bayesian procedure for the generalized Pareto distribution." *Journal of Statistical Planning and Inference*, 137(2), 473–483.

Der Kiureghian, A. and Ditlevsen, O. (2009). "Aleatory or epistemic? Does it matter?." *Structural Safety*, 31(2), 105–112.

Di Baldassarre, G. (2012). *Floods in a Changing Climate: Inundation Modelling*. Number 3 in International Hydrology Series. Cambridge University Press.

Fawcett, L. and Green, A. C. (2018). "Bayesian posterior predictive return levels for environmental extremes." *Stochastic Environmental Research and Risk Assessment*, 32(8), 2233–2252.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis, Third Edition*. CRC Press (November).

Kennedy, M. C. and O'Hagan, A. (2001). "Bayesian calibration of computer models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3), 425–464.

Melchers, R. E. and Beck, A. T. (2018). *Structural Reliability Analysis and Prediction*. Wiley, Hoboken, NJ, third edition edition.

Merz, B. and Thieken, A. H. (2005). "Separating natural and epistemic uncertainty in flood frequency analysis." *Journal of Hydrology*, 309(1-4), 114–132.

Spiegelhalter, D. J. and Riesch, H. (2011). "Don't know, can't know: Embracing deeper uncertainties when analysing risks." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1956), 4730–4750.