Uncertainty quantification of flood mitigation predictions and implications for interventions

Koen D. Berends[1,2], Menno W. Straatsma[3], Jord J. Warmink[1], and Suzanne J.M.H. Hulscher[1]

[1]Department of Marine and Fluvial Systems, Twente Water Centre, University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands
[2]Department of River Dynamics and Inlands Shipping, Deltares, Boussinesqweg 1, 2629 HV Delft, The Netherlands
[3]Department of Physical Geography, Faculty of Geosciences, University of Utrecht, Princetonlaan 8, 3584 CS Utrecht, The Netherlands

Correspondence: K.D. Berends (k.d.berends@utwente.nl)

Abstract. Reduction of water levels during river floods is key in preventing damage and loss of life. Computer models are used to design ways to achieve this and assist in the decision making process. However, the predictions of computer models are inherently uncertain, and it is currently unknown to what extent that uncertainty affects predictions of the effect of flood mitigation strategies. In this study, we quantify the uncertainty of flood mitigation interventions on the Dutch River Waal, based on 39 different sources of uncertainty and twelve intervention designs. The aim of each intervention is to reduce flood water levels. Our objective is to investigate the uncertainty of model predictions of intervention effect and to explore relationships that may aid in decision making. We identified the relative uncertainty, defined as the ratio between the confidence interval and the expected effect, as a useful metric to compare uncertainty between different interventions. Using this metric, we show that intervention effect uncertainty behaves like a traditional backwater curve with an approximately constant relative uncertainty value. In general, we observe that uncertainty scales with effect: high flood level decreases have high uncertainty and conversely, small effects are accompanied by small uncertainties. However, different interventions with the same expected effect do not necessarily have the same uncertainty. For example, our results show that the large-scale but relatively ineffective intervention of floodplain smoothing by removing vegetation, has much higher uncertainty compared to alternative options. Finally, we show how a level of acceptable uncertainty can be defined and how this can affect the design of interventions. In general, we conclude that the uncertainty of model predictions is not large enough to invalidate model-based intervention design, nor small enough to neglect altogether. Instead, uncertainty information is valuable in the selection of alternative interventions.

1 Introduction

The number of people living in areas exposed to river flooding is projected to exceed one billion in 2050 (Jongman et al., 2012). Therefore, it is increasingly important that the river system is designed in such a way that flood risk is minimised. Human intervention in river systems has a long history, to a point that for some rivers human management has become the

dominant factor driving hydrological change (Pinter et al., 2006; Bormann and Pinter, 2017). Today, the decision to change an existing river system (e.g., by leveeing a channel to protect flood-prone areas) is increasingly based on predictions made by computer models. Various software systems can be used to resolve flow conditions accurately, taking into account much of the complexity of rivers — such as terrain geometry, vegetation and hydraulic structures. However, despite increasingly available

5 data, not all model input or model parameters can be reliably measured or represented. For example, vegetation density and vegetation height, which modify vegetation roughness (Baptist et al., 2007; Luhar and Nepf, 2013), are variable both in time and space. This variability is neither captured on the scale of river modelling, nor by the equations that resolve vegetation roughness. Examples such as this introduce uncertainty in the modelling process and in the model output (Oreskes et al., 1994; Walker et al., 2003).

10   Proper understanding and communication of uncertainty is important both for scientists and decision makers (Pappenberger and Beven, 2006; Uusitalo et al., 2015). Maier et al. (2016) distinguished three complementary paradigms for modelling to support decision making under (deep) uncertainty: (1) using the best available knowledge, (2) quantifying uncertainty and sensitivities of key parameters and (3) exploring multiple plausible outcomes. One advantage computational models can bring to decision making for river engineering is an assessment of the impact of the planned intervention on hydraulics, such as water

15 levels, flow velocities and the morphodynamic response of the river bed. While hydraulic effects are not the only impact of interventions (see e.g. Straatsma et al., 2017, for effects on biodiversity) they are considered important: for the 39 interventions of the 2.3 billion Euro 'Room for the River' program in The Netherlands, the hydraulic effect as predicted by models was a precondition for any design "to be taken seriously at all" (Klijn et al., 2013). Mosselman (2018) reported that quantified, large uncertainty in flood water levels is sometimes played down when assessing effects, under the assumption that systematic

20 errors cancel out when subtracting the intervention case from the reference case. However, it was recently demonstrated in an idealised study that uncertainty in flood water levels does not cancel out but could be significant compared to the effect, and sensitive to the specific design of the intervention (Berends et al., 2018). Therefore, uncertainty quantification of effect studies for a real-world case study is needed to assess the implications for the design of interventions.

    Uncertainty quantification of impact analysis for real world solutions suffers from two compounding issues specific to

25 intervention design studies. First, there is a practical need for sufficiently detailed models given the increasingly complicated design of interventions, which moves from traditional flood prevention (building dikes, or levees) to more holistic designs. This is in part motivated by the paradoxical 'levee effect' that states that flood control measures do not decrease, but increase flood risk (White, 1945; Baldassarre et al., 2013; Munoz et al., 2018) and which insight has spurred policy change away from purely flood control and towards alternative options (Pinter et al., 2006; Klijn et al., 2018). In The Netherlands, this

30 has led to designs that focus on increasing conveyance capacity to achieve lower flood levels, while integrating multiple other (ecological, societal) objectives as well (Rijke et al., 2012). Predicting the effects of such more intricate alternative approaches, like the construction of artificial secondary channels (van Denderen et al., 2019), requires detailed models that take into account local geometry, vegetation and other terrain features. Such models are not only computationally expensive, which both complicates Monte Carlo simulation (MSC)-based quantification methods, but also have a high number of model

35 parameters, which complicates the use of (data-based) surrogate model approaches (Razavi et al., 2012). The second issue is

that the analysis relies on at least one unmeasured environmental system, namely, the river system altered by the proposed intervention. Therefore, there is no way to verify that accuracy for the proposed future state of the river. Verification of the accuracy of the current, unaltered state is likewise not necessarily available. For example, flood mitigation measures tend to be designed for a low period of return, e.g. a one in one hundred year or even one in 1250 year flood (Klijn et al., 2018), for which recent measurements are (by definition) sparse. Under such conditions, the uncertainty of model output cannot be expressed in terms of accuracy (e.g. standard deviation of model error) without additional untestable assumptions.

In present paper we address the lack of studies into the model uncertainty of model predictions used in impact design. Our objective is twofold. The first objective is to quantify the effect of parameter uncertainty on the predicted effect of flood mitigation measures, by implementing twelve different interventions of varying type and intensity in a section of the Dutch River Waal, using a detailed hydraulic model. To limit the computational burden we use CORAL, which is an implementation of the efficient uncertainty quantification method of Berends et al. (2018), that does not suffer from limitations with regard to the number of model parameters. Our second objective is to explore the relationship between the expected reduction of flood levels and the uncertainty, to aid in the decision making process. To this end we will introduce the 'relative uncertainty' measure to facilitate inter-comparison between different intervention designs.

Our analysis proceeds in three steps. First, we set up the numerical model for the Dutch River Waal and select the uncertainty sources in section 2. Here, we also introduce the relative uncertainty measure. We then present the results of the uncertainty of the modelled water levels and reduction in water levels in section 3.1 and 3.2. In section 3.3 we show how uncertainty may influence decision making. In the discussion (4) we discuss reducing uncertainty in intervention design and the feasibility of probabilistic analysis for intervention design in practice. A general conclusion on the practical value of uncertainty quantification for decision making in flood mitigation strategies, and specific conclusions related to the objectives, are given in section 5.

2 Methods

2.1 Study area

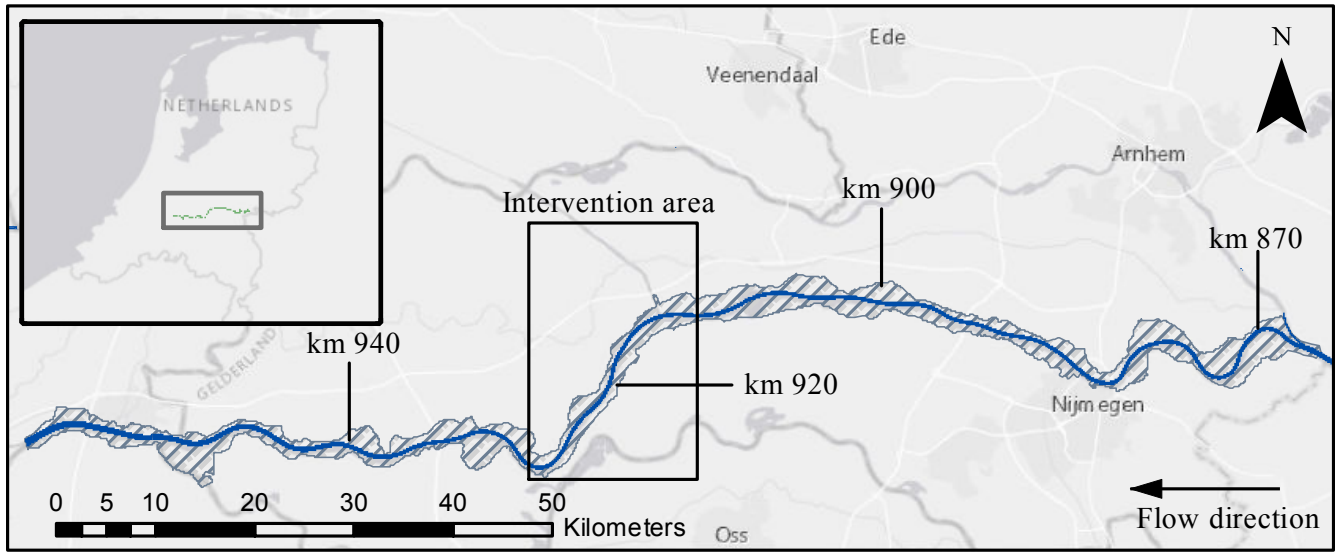
The River Waal is selected as the case study for its extensive history of human intervention and good data availability. The Waal is a distributary of the River Rhine and, by discharge, the largest river in the Rhine-Meuse-Scheldt Delta, located in Western Europe. The present-day river has a main channel about 8 meters deep at bank-full discharge, a bank-full width of about 250 m and relatively narrow floodplains. Before the construction of the dikes from 1000 A.D. onward, it was a meandering river with regular over-bank deposition (Hobo et al., 2014). Recorded river engineering works for flood protection as well as as inland navigation dates back to at least the 18th century (e.g., see Velsen, 1749). In the 19th century, the river was normalised in large governmental projects, obtaining its current narrow channel, groynes (also termed *spur dikes* or *wing dikes*) and relatively straight channel (Sieben, 2009). In present day, the river has a bed slope of approximately $1 \cdot 10^{-4}$ m/m over a stretch of 90 km from the upstream bifurcation to the Merwede bifurcation. At the end of the 20th century, motivated by the near-disaster of 1995 (Chbab, 1995), the large scale 'Room for the River' programme was enacted to reduce flood levels (Rijke et al., 2012;

Esri, HERE, Garmin, © OpenStreetMap contributors, and the GIS user community

Figure 1. The River Waal, situated entirely in The Netherlands. The study area is a stretch of about 15 km between km 913 and km 928. The hatched area shows the total extent of the model domain, which overlaps with the river and its floodplains bounded by the dikes.

Klijn et al., 2018). The largest intervention in the River Waal in this programme was the Nijmegen-Lent dike relocation project, which was predicted to lower the water levels by 27 cm at a cost of approximately 350 million Euro. Most other interventions on the River Waal were projected to lower the water level by less than 10 cm.

Our study domain covers the entire Waal River (Figure 1). Distances along the river are measured using the conventional
5 'Rhine kilometre', which starts from km 0 at the German city of Konstanz. The Waal starts at km 867 at the Pannerden bifurcation and runs through to km 961 where it bifurcates in the Nieuwe Merwede and Boven Merwede rivers. All flood mitigation measures studied will be implemented between km 913 and km 928, depicted by the annotated rectangle in Figure 1.

The intervention area is characterised by a relatively straight river stretch with narrow floodplains and the strongly curved *St.*
10 *Andries* river bend, which provides a known bottleneck during high discharges. Nine towns border directly on the dikes along this 17.5 km stretch. Interventions in this area to increase flood safety, considering the narrow floodplains and the populated surrounding land, present both a technical and a societal problem.

2.2 Hydrodynamic model & interventions

To simulate the hydrodynamic response of the flow to various interventions we use the Delft3D Flexible Mesh modelling
15 system (Kernkamp et al., 2011). Geographical information, such as land cover, bathymetry and hypsometry, as well as information on embankments and weirs, were derived from the Baseline information system (Becker et al., 2014). Our setup uses a two-dimensional unstructured numerical grid, with between 71,000 and 120,000 active grid cells. The grid size varies from

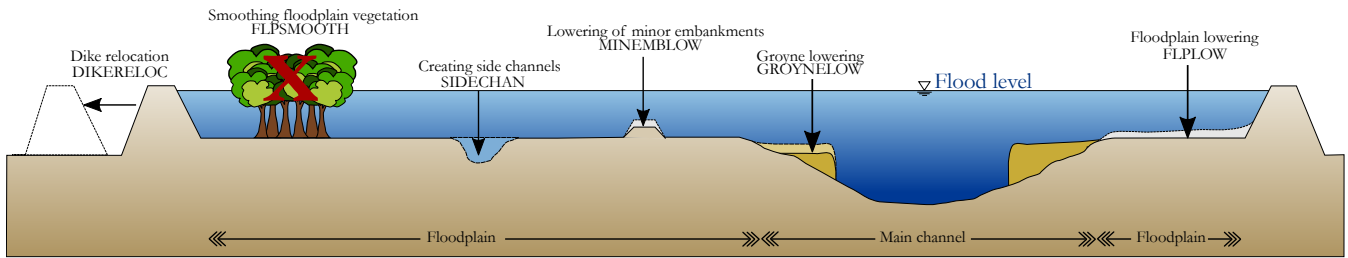



Figure 2. A schematic overview of a typical cross-section of the River Waal and the flood mitigation measures studied in this paper. Figure based on Middelkoop et al. (1999).

40m by 15m in the main channel to a maximum of about 120m by 120m in the floodplains. In all cases, we simulate a steady upstream discharge of 10.165 $\mathrm{m^3s^{-1}}$ and a constant downstream water level of 3.98 m + NAP, for 72 hours. A single model evaluation on our computer resources took about 2.5 hours. These conditions are consistent with a return period of 1250 years (T_{1250}). Initial conditions were derived from a reference run with all considered stochastic variables (see section 2.3) at their
5 median values. The initial time step was set at 30s.

All interventions are modelled as changes to a reference state of the system. This reference state is the Waal River after all interventions from the Room for the River programme, which approximately corresponds to the 2016 situation. We consider twelve additional system states, each one corresponding with a particular human intervention (Figure 2). In all cases, the interventions were procedurally generated by the RiverScape tool (Straatsma and Kleinhans, 2018). We implemented six dif-
10 ferent interventions, detailed below. Each intervention was implemented in a low-intensity and a high-intensity variant. Here, intensity refers to the magnitude of the intervention, for example a small reduction in groyne height (low-intensity) or a large reduction in groyne height (high-intensity). We assume that the intervention is carried out in reality exactly as it was designed. This is known as the *'as designed'* post-intervention state. In reality, there may be a discrepancy between the *as designed* state and the actual ('*as built*') state. In theory, this discrepancy could be considered as an additional source of uncertainty. However,
15 because there is no literature to support any assumptions regarding such discrepancy, we do not take this into account here.

2.2.1   Groyne lowering (GROYNLOW)

Groynes (also known as *wing dikes* or *spur dikes*) are (stone) structures perpendicular to the flow. During normal conditions, groynes restrict the effective channel width to promote navigable depths. However, during high flows groynes obstruct flow. Prior to 2016, many groynes in the Waal were already lowered as part of the Room for the River program. The intervention as
20 implemented in this study further reduces the groynes crest-heights. Groynes were lowered to the crest-height corresponding to the water levels with exceedance frequency of 150 days (low intensity) and 363 days (high intensity) per year.

2.2.2 Minor embankment lowering (MINEMBLOW)

The River Waal floodplains are compartmentalised by minor embankments, which original purpose is to prevent flooding of the floodplains during minor (summer) floods. This intervention lowers the crests of these embankments to reduce their obstruction during high flow. Low and high intensity lowered the crests to a water level with exceedance frequency of 50 days (low intensity) and 150 days (high intensity) per year.

2.2.3 Floodplain lowering (FLPLOW)

Lowering or excavation of the floodplains increases the maximum water volume within the existing bounds of the river corridor, thereby increasing conveyance. In this study, we lower the level of the floodplain without changing existing vegetation or other floodplain configuration to isolate the effect of lowering. Floodplains were lowered to the corresponding water level with exceedance frequency of 20 days per year, with 5% (low intensity) and 99% (high intensity) of the terrestrial floodplain area altered.

2.2.4 Floodplain smoothing (FLPSMOOTH)

Vegetation in the floodplain greatly contributes to resistance during high flows. Replacing existing high-friction vegetation by low-friction vegetation mitigates this problem. Here, the existing vegetation was replaced by production meadow (roughness code 1201 in table A1). Smoothing at low intensity affected the top 5% of the obstructing vegetation. At high intensity all vegetation was converted to meadow.

2.2.5 Dike relocation (DIKERELOC)

In the River Waal, dikes are the primary defence against flooding. However, they also contribute to flood risk by restricting the river corridor. Dike relocation increases the floodplains and allocates more space to the river. While this is perhaps the best example of combatting problematic constriction of the river corridor, it is also the most invasive — considering human settlement near and on the dikes. At low intensity, concave dike sections of less than 700 m are replaced by straight dikes, whereas at high intensity concave sections of 7000 m are straightened, while creating small polders around existing built-up areas.

2.2.6 Side channels (SIDECHAN)

Finally, the construction of secondary (side) channels within the existing corridor both increases the available volume, and decreases vegetation friction. All new channels are assigned "side channel" roughness (Code 105 in Table A1) and a trapezoidal cross sectional shape and slope of 1 to 3. The channels were implemented with widths of 10 m (low intensity) and 100 m (high intensity), and depths of 0.35 m (low intensity) and 3.5 m (high intensity) below the water level with an exceedance frequency of 363 days per year.

Table 1. An summarised overview of the stochastic variables. A full overview is given in Table A1

Type		Number of stochastic variables
Main channel roughness		1
Floodplain roughness		
	Vegetation height (h_v)	14
	Vegetation density (n_v)	17
	Non-vegetation roughness	6
Classification error		1
Total		39

2.3 Uncertainty sources

Modelling real-world rivers on the scale of nearly 100 km necessarily involves various simplifications, discretisations and parameterisations. Sources of uncertainty for this river were identified by Warmink et al. (2011) using expert elicitation. The main sources were (a) boundary conditions and (b) main channel friction and to lesser extent (c) friction by vegetation, (d) geometry and (e) weir and groynes formulations. In this study we follow the design approach taken in the Room for the River program, which assumes the boundary conditions given and stationary for a certain design return rate (which is T_{1250}) and therefore not a source of uncertainty. Warmink et al. (2013b) considered the hydraulic roughness parameters to be the most important parameter based on literature. We adopt this assumption and take into account uncertainty of the main channel friction, vegetation parameters and classification errors in the land-use maps. The total number of stochastic variables is 39 (see Table 1).

2.3.1 Main channel roughness

The hydraulic roughness of the main channel is chiefly determined by the material of the bed ("grain roughness") and the bed forms ("form roughness"). Various models have been proposed that calculate the friction factor based on the characteristics of the bed (van der Mark, 2009). Here, we adopt the approach of Warmink et al. (2013a) who estimated the values for Nikuradse roughness height k_n [m] for a T_{1250} event by extrapolation using a Generalized Extreme Value (GEV, Weibull variant) distribution for five roughness models. Given the GEV percentiles, we fitted a lognormal distribution at the T_{1250} return period for each roughness model. These lognormal distributions were combined into a single distribution using equal weight for each roughness model. Figure 3 shows the resulting individual and combined probability functions from the Weibull extrapolation.

In the following we will use the joint cumulative density function of the five roughness models to sample representative roughness height k_n. The joint cumulative function is a highly asymmetrical distribution with a mean of approximately 0.58 m and 95% confidence limits at 0.31 m and 1.0 m. The sampled values of k_n will be used in the hydrodynamic model as input for the White-Colebrook model:

$$C = 18_{10} \log \frac{12h}{k_n} \tag{1}$$

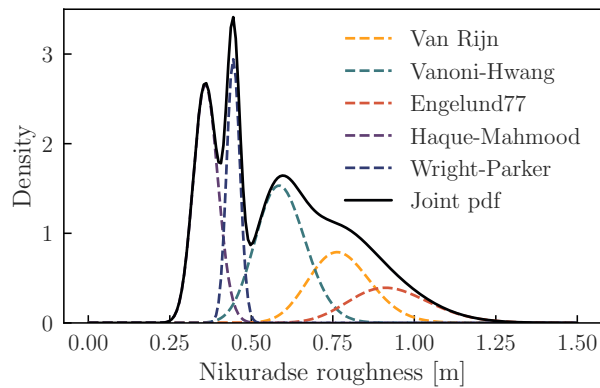



Figure 3. The combined probability function of the main bed roughness is based in Weibull extrapolation of Nikuradse roughness height k_n to the T_{1250} return period using five different bed roughness formulas.

with Chézy coefficient C $[\mathrm{m}^{1/2}\mathrm{s}^{-1}]$ and water depth h [m]. The sampled k_n values are used to resolve friction for all main channel roughness classes, which cover 95.6% of the main channel area. The other 4.4% are non-erodible layers in the outer river bends which are not stochastically determined in this study.

2.3.2 Floodplain roughness

5 The Baseline database divides the floodplain into vegetation classes, riverbanks and water bodies. Each class is given a specific roughness code called a *trachytope*, which is then coupled to an appropriate roughness formula and parameters specific to that formula. Spatially, the vegetation in the floodplains is discretised by polygons, which are designated by either one or a combination of distinct vegetation classes.

For vegetation, the generation of roughness by vegetation is complex, resulting from obstruction by stems and leafs and
10 turbulent flow through the vegetation and over the canopy. The relationship between vegetation and roughness is extensively studied, leading to various models to compute hydraulic friction from vegetation traits (Nepf, 2012; Vargas-Luna et al., 2015; Shields et al., 2017). These traits often include the vegetation canopy height h_v[m], drag coefficient C_D [-], stem density $n_v = m_v D_v$ or vegetation concentration $\lambda = \frac{\pi D_v^2 m_v}{4}$ with number of stems per square meter m_v $[\mathrm{m}^{-2}]$ and stem diameter D_v [m] (Klopstra et al., 1997; Stone and Shen, 2002; Baptist et al., 2007; Huthoff et al., 2007; Yang and Choi, 2010; Li et al.,
15 2015). Warmink et al. (2013b) have shown that the choice for a particular vegetation model has little influence on the outcome uncertainty compared to parameter uncertainty. Therefore, we only used the model proposed by Klopstra et al. (1997), in which vegetation friction is chiefly determined by the parameters h_v, n_v, C_D. These parameter values are determined in the following way. For C_D we use fixed, deterministic values based on van Velzen et al. (2003). The height and density parameters (h_v and n_v) are assigned probability distributions based on the field campaign of Straatsma and Alkema (2009). Based on 206
20 field observations, they approximated the variation in vegetation density and height for 25 homogeneous vegetation classes.

Based on their estimate quantiles we fitted lognormal distributions for each class, assuming the vegetation height and density to be independent parameters. In total, the roughness of 66% of the River Waal floodplain is defined by one of the vegetation trachytopes and 2.5% is defined by combination of two vegetation trachytopes.

Six trachytopes are associated with water bodies or riverbanks, accounting for about 27% of the total areal. We model the friction for these 6 classes with the Manning formula

$$C = h^{1/6} n^{-1} \tag{2}$$

with Manning coefficient n [sm$^{-1/3}$]. The n values for these trachytopes are modelled as triangular probability functions with the minimum, maximum and mean values based on Chow (1959). All parameters of the vegetation parameter distributions are summarised in Table A1. The remaining 4.5% of floodplain area, which is given deterministic roughness values, is mostly (3.7%) covered by buildings which are excluded from flow computations, revetment and pavement.

2.3.3 Classification error matrix

Available vegetation maps are likely to contain 'impurities' (Knotters and Brus, 2012) or classification errors, that have significant impact on model output uncertainty (Warmink et al., 2013a). The probabilities of one class being in reality another is encoded in a confusion matrix (we use Table 1 in Straatsma and Huthoff (2011)). Following this matrix it is found, for example, that "Willow plantation" was, in all cases, found to be "Softwood forest" instead. An overview of the various vegetation classes is given in the Appendix. Following these probabilities we generated an ensemble of 2000 pre-intervention trachytope maps. Each map was then given a fixed index number. We sample from these maps by picking a number from a discrete uniform distribution and finding the corresponding map by index number. In this way, each map is assigned the same probability.

2.4 Quantification of model output uncertainty

2.4.1 Estimation method

In this study we focus on prediction of the maximum flood levels (see Figure 2) in the reference state (denoted by X), the maximum flood levels in the various interventions (denoted by Y) and the difference between these two (denoted by ΔH). Due to the uncertainties in model input discussed in section 2.3, X, Y and ΔH are stochastic. To analytically compute differences between these two (stochastic) states of the river system, we would need to obtain both the model output probability distributions, as well as the covariances between the two distributions (Berends et al., 2018). This problem is analytically intractable, but can be solved numerically by Monte Carlo simulation (MCS; Metropolis, 1987; Stefanou, 2009). In the numerical approach, covariances are not explicitly computed.

However, MCS is not always practically feasible or desirable. Although MCS has been applied for estimation of flood level uncertainty before (Warmink et al., 2013b), the added task of performing a separate MCS for each intervention puts a severe strain on computational resources. Therefore we follow the approach of Berends et al. (2018), schematically depicted in Figure 4, which is computationally more efficient. We refer to this method as 'correlated output regression analysis' (CORAL). The

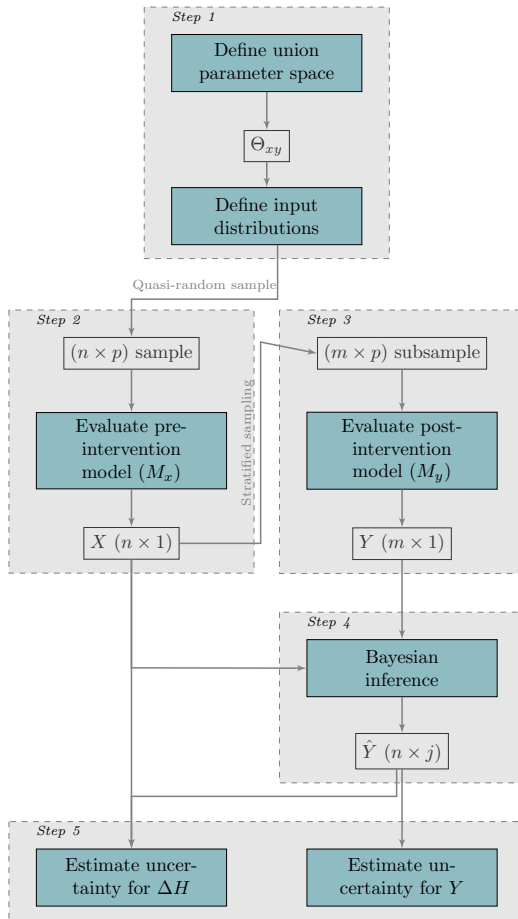



Figure 4. Schematic overview of the efficient uncertainty estimation method CORAL for post-intervention model output and impact analysis. In this figure: n: MCS sample size; m: subsample size; p: number of stochastic parameters; X: pre-intervention model output; j the number of draws from the posterior predictive model; Y: post-intervention model output; $\hat{Y}$: estimated post-intervention model output; ΔH: intervention effect. Figure adapted from (Berends et al., 2018).

key feature of CORAL is that the stochastic intervention output Y is not directly solved by MCS, but instead an estimator $\hat{Y}$ is computed using the reference state and a correlation model. This correlation model is defined within a Bayesian framework and is probabilistic as well. We use the same linear model as Berends et al. (2018):

$$\hat{Y} = \alpha(X - c) + \beta + \epsilon + c \tag{3}$$

5   with $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon)$ and transformation constant c. The constant is introduced to shift data to the origin, which facilitates faster inference, and is defined as $c = \min(X)$. In the Bayesian framework, the regression parameters $\theta = \{\alpha, \beta, \sigma_\epsilon\}$ are considered stochastic variables themselves and estimated via a Markov Chain Monte Carlo (MCMC) algorithm, trained by a limited subsample of simulations from Y. We use a subsample of $m = 20$ for all intervention states. To estimate the distribution of the

reference state (step 2) we use a quasi-random MCS sample ($n = 1000$) with the Sobol' low-discrepancy sequence. The linear model was demonstrated to be applicable to similar intervention studies by Berends et al. (2018), and (visual) inspection of the correlation between the reference state and the intervention models shows this holds for our cases as well.

The main advantage of this approach is a significant decrease of required computational resources by reducing the number of model evaluations. The total number of hydrodynamic model evaluations was 1240, divided over thirteen model states (the reference plus 12 interventions). This is a decrease of computational effort of more than 90% compared to direct MCS with all states, which would have totalled 13.000 model evaluations. It is important to state that the consequence of having a probabilistic correlation model, is that the intervention state estimator $\hat{Y}$ is stochastic for each individual draw from the reference state (X). Since X is stochastic itself, this can be thought of as the 'uncertainty of the uncertainty' or the estimation uncertainty.

To maintain correlation between the X and Y, a union stochastic parameter space is constructed (step 1 in figure 4). Practically, this means that each corresponding member in the reference and intervention model ensembles has identical values for their stochastic variables, with the noticeable exception where those were altered by the intervention. For example, we assume that a tree with a height of 8 m in the reference case still has a height of 8 m after intervention, unless this tree was removed or otherwise affected by the intervention. This is straightforward to implement for parameter values, but requires an additional step for the classification uncertainty. To be able to draw from the confusion matrix (see section 2.3.3) a large number of trachytope maps was generated for the reference case. To maintain correlation, the same number of maps is generated for each intervention in such a way that every map is identical to the corresponding reference map with the same index, except in areas where the intervention has altered the trachytope. For example, an area which was classified as 'willow plantation' in the reference state will still be classified as 'willow plantation' after the intervention, unless that area has been reclassified as part of the intervention.

2.4.2 Uncertainty metrics

As a data reduction step, we summarise the results using the following metrics based on the water level reduction effect ΔH.

Confidence intervals measure the distance between two percentiles (e.g. the 10% and 90%) of a cumulative density function (cdf). We refer to the confidence intervals of the stochastic variables X, Y and ΔH as the model uncertainty confidence interval or MCI. For example, the MCI of ΔH is defined as:

$$\text{MCI}_{\Delta H} = |\hat{F}_{\Delta H}^{-1}(p_u) - \hat{F}_{\Delta H}^{-1}(p_l)| \tag{4}$$

The MCI of $\hat{Y}$ is derived in the same way. Here, $\hat{F}^{-1}$ is the inverse cumulative density function, also known as the *quantile* or *percent point* function and p_l, p_u the lower and upper quantiles. Here, these quantiles are always symmetric around the median. We use a limited subset of quantiles (consistent with the 10%, 20%, 50%, 80% and 90% confidence intervals) to summarise the data. Additionally, we calculate the exceedance probability $P(\Delta H < z) = \hat{F}_{\Delta H}(z)$.

It is important to note that $\text{MCI}_{\hat{Y}}$ and $\text{MCI}_{\Delta H}$ are stochastic — resulting from the estimation method. Therefore, each value belonging to an arbitrary quantile p_z may be expressed in terms of the expected value $E(\hat{F}_{\Delta H}^{-1}(p_z))$ and variance

$\text{Var}(\hat{F}_{\Delta H}^{-1}(p_z))$. This uncertainty may be referred to as the estimation uncertainty. A more detailed account of this can be found in Berends et al. (2018).

To compare the uncertainty of interventions we calculate the *relative uncertainty*, i.e. the uncertainty relative to the expected effect. For this, we use an adapted version of the coefficient of variation, defined as:

$$U_{r;90} = \frac{\text{MCI}_{90,\Delta H}}{\text{E}(\Delta H)} \tag{5}$$

where $\text{MCI}_{90,\Delta H}$ is the 90% confidence interval for ΔH and $\text{E}(\Delta H)$ the average of ΔH. We use $U_{r;90}$ instead of the coefficient of variation (which is defined as the standard deviation on the expected value) to directly express the size of the largest used confidence interval in percentages of the expected effect. Therefore, a $U_{r;90}$ of 100% means that the 90% MCI is as large as the average effect. Since both $\text{MCI}_{90,\Delta H}$ and $\text{E}(\Delta H)$ are stochastic, $U_{r;90}$ is stochastic as well.

3   Results

3.1   Pre-intervention uncertainty

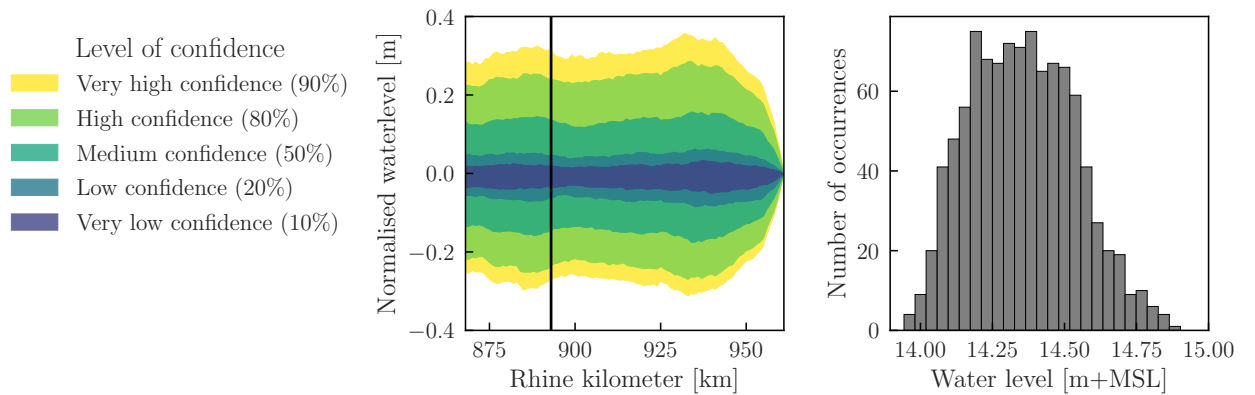



Figure 5. Middle panel: Confidence levels of design water levels, normalised by the average water levels at each location. The water flow is from left (upstream) to right (downstream). The black line indicates km 893. Right panel: the histogram of the (not normalised) water levels at km 893.

Figure 5 shows the uncertainty in the water levels along the River Waal before any intervention has taken place, normalised by the average water levels. The average confidence intervals over the entire stretch range from 55.4 cm for the 90% interval to 5.7 cm for the 10% interval. However, the intervals are not uniform along the river. At the downstream boundary the intervals collapse to zero due to the fixed boundary condition. Upstream from the boundary, there is a noticeably smaller interval, approximately between km 885 and km 925. This is attributed to the relatively wider floodplains in this part of the river, which reduces flow through the main channel (Warmink et al., 2013b). The right hand panel of Figure 5 shows the histogram at

km 893, which enables visual comparison with Figure 6 in Warmink et al. (2013b), who performed their analysis with the WAQUA modelling system for the 1995 River Waal and slightly higher upstream boundary discharge. They reported a 95% confidence interval of 68 cm (13.99 m + MSL to 14.66 m + MSL) at km 893. Our results also show a 95% confidence interval of 68 cm, although at slightly higher water levels (14.05 m + MSL to 14.73 m + NAP). This provides confidence in the model

5   results. However, we highlight two differences with respect to the study of Warmink et al. (2013b). First, we did not include the vegetation roughness model as a source of uncertainty, but we added vegetation parameter uncertainty. This change does not seem to have affected the output uncertainty of flood levels significantly. Second, the increase in flood levels with respect to the study of Warmink et al. (2013b) was not expected, given that between 1995 and 2015 a large scale flood mitigation programme (called 'Room for the River') was carried out aimed at reducing flood levels. However, given the different modelling systems

10   and assumptions, this comparison cannot be used to judge the effectiveness of that programme.

3.2   Uncertainty of along-channel flood level decrease

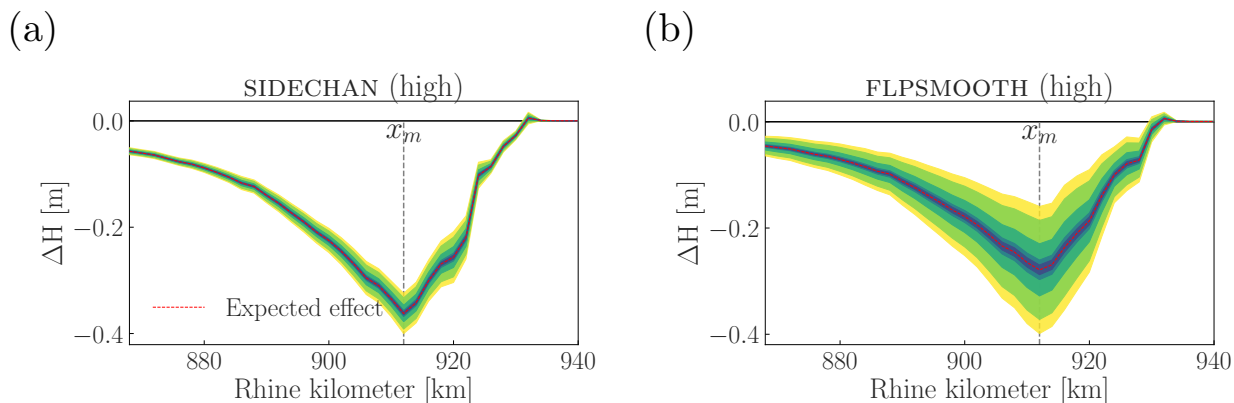



Figure 6. The expected lowering of flood levels for side channel construction (a, SIDECHAN), and floodplain smoothing (b, FLPSMOOTH), both high intensity and flow direction from left to right. The colours in this chart convey identical confidence levels as shown in Figure 5. The location of the maximum expected effect is indicated by the dashed line and annotation x_m.

The primary objective of the flood mitigation measures is to lower flood levels during a given design discharge. In deterministic approaches, this effect is typically communicated with an along-channel diagram showing the difference (i.e., before and after the intervention) in flood levels (see e.g. Figure 3 in Klijn et al., 2018). Figure 6 shows a similar diagram, augmented with

15   information on the uncertainty of the prediction. The expected effect follows the general trend as a deterministic approach. Starting from the downstream end of the intervention (around km 932), water levels gradually decrease due to the local increase of conveyance compared to the reference state. From km 912 upstream, water levels gradually return to the reference

equilibrium. The gradual relaxation to equilibrium levels both over and upstream from the intervention, i.e. backwater curves, is explained through basic sub-critical flow theory. We note that the length of the backwater curves spans tens of kilometres resulting in a residual water level decrease of approximately 5 cm near the upstream boundary (km 868). In theory, this could affect the distribution of discharge at the bifurcation of the River Waal with the Pannerden Canal. This secondary effect is not

5   accounted for in this study.

The novelty of Fig. 6 results is mainly found in the uncertainty, visualised through the confidence intervals. These intervals follow the movement of the average; going down when the average decreases and relaxing toward the reference equilibrium upstream from the location of maximum effect (x_m). This shows that taking into account parameter uncertainty does not fundamentally change established typical effects of flood mitigation measures. However, the movement with the average is not

10 accompanied by parallelity: the range of the intervals is not constant along the river. Instead, the range of the intervals scale with the expected effect: the ranges are small for small effects, and increase with increasing effect. Consequently, uncertainty is largest at the maximum effect. Downstream from the intervention the confidence bands tend to zero, since for a steady computation, there is no difference between the pre- and post-intervention models and no backwater effects. Upstream from the intervention the confidence intervals reduce to zero along with the backwater curve effect.

15   Next, we compute the relative uncertainty ($U_{r;90}$) for each intervention and intensity. We first marginalise the uncertainty in $U_{r;90}$ due to CORAL by calculating the expected value ($E(U_{r;90})$) for each point along the river (Fig. 7). Figure 7, as well as the relatively small standard deviations in Table 2, show that values for $E(U_{r;90})$ are fairly constant over the river length up to the location of the maximum flood level decrease (x_m) for all interventions. A notable exception is the peak at approximately km 885 for low-intensity GROYNLOW. This is attributed to small inaccuracies in the estimation of the expected effect. Since

20 $U_{r;90}$ is a ratio, its value is sensitive to small values of the denominator, which is the expected effect (see Eq. (5)). GROYNLOW has a very minor effect (maximum of 1 cm at low intensity), so small inaccuracies in estimation will greatly affect $U_{r;90}$. For this same reason, we do not compute $U_{r;90}$ upstream from the maximum effect, as the flood level decrease will rapidly return to and cross zero, which will lead to rapidly exploding $U_{r;90}$ values. From the fact that we do not observe such peaks much more often, we gain confidence that the chosen estimation method is suitable for this type of analysis even at very small-scale

25 effects. In general, the constant values for $E(U_{r;90})$ shows that it is a useful parameter to characterise the uncertainty of an intervention, as the increase in uncertainty with the effect, observed in Figure 6 is sufficiently summarised by a single $U_{r;90}$ value.

It is interesting to note that in theory, the adaptation length (i.e. the length scale of the backwater curve) is affected by the equilibrium water level and can therefore result in converging or diverging confidence intervals. Since $E(U_{r;90})$ is relatively

30 constant along the river, the adaptation lengths (i.e. the length scale of backwater effects) are not significantly affected by the intervention. Therefore, we see that all uncertainty is generated over the length of the intervention and subsequently diminishes upstream as water levels return to their equilibrium. In other words, if the adaptation lengths are known, the uncertainty upstream of x_m can be readily estimated from the relative uncertainty at x_m.

A comparison of all interventions (Table 2) shows that small effects ($E(\Delta H)_{xm}$) are accompanied by small uncertainties ($E(\text{MCI}_{90})_{xm}$). This shows that even small changes to rivers can be predicted, even when the absolute uncertainty in water levels (see figure 5) is an order of magnitude larger than the expected effect.

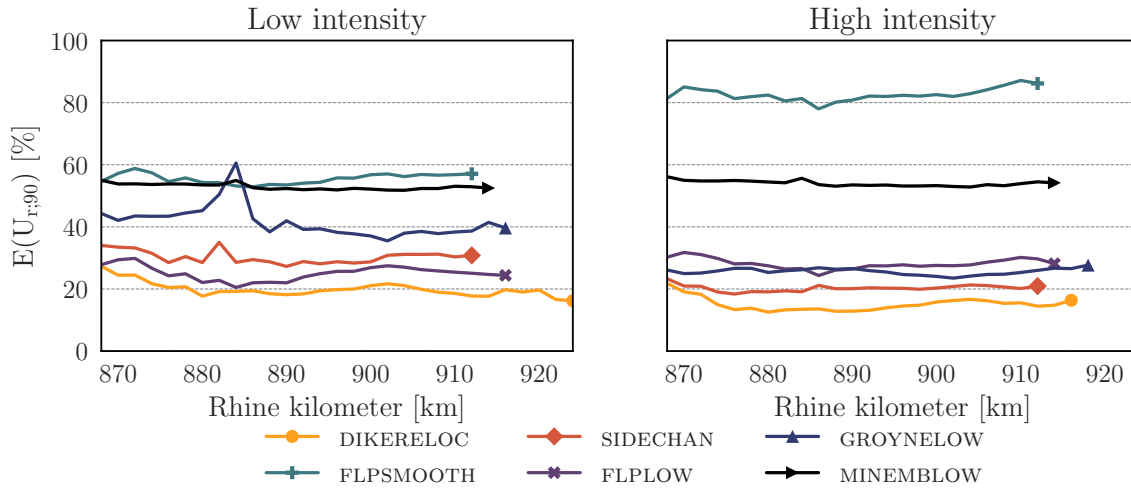



Figure 7. The expected values for the relative uncertainty ($U_{r;90}$) along the river, from the upstream boundary at km 868 to the location of maximum effect x_m indicated by the marker, showing relatively constant values upstream from x_m.

Table 2. Summarised results for each measure. H=high intensity; L=low intensity; x_m= location in Rhine kilometre (km) where maximum effect occurs; $E(\Delta H)_{xm}$= expected effect at x_m; $E(\text{MCI}_{90})_{xm}$=expected 90% uncertainty at x_m; $\widehat{E(U_{r;90})}$= Expected relative uncertainty along the river (average $\pm$ the standard deviation). The rows are sorted from highest expected effect (top) to lowest expected effect (bottom).

Intervention		x_m	$E(\Delta H)_{xm}$	$E(\text{MCI}_{90})_{xm}$	$\widehat{E(U_{r;90})}$
DIKERELOC	H	916	-1.08	0.176	15±2.2%
	L	924	-0.21	0.034	20±2.3%
FLPLOW	H	914	-0.80	0.224	28±1.7%
	L	916	-0.10	0.024	25±2.3%
SIDECHAN	H	912	-0.36	0.076	20±1.0%
	L	912	-0.02	0.007	30±2.1%
FLPSMOOTH	H	912	-0.28	0.239	82±1.9%
	L	912	-0.04	0.023	55±1.6%
GROYNLOW	H	918	-0.04	0.012	26±0.9%
	L	916	-0.01	0.004	42±5.1%
MINEMBLOW	H	914	-0.03	0.014	54±0.9%
	L	914	-0.03	0.013	53±0.9%

3.3 Inter-comparison between interventions

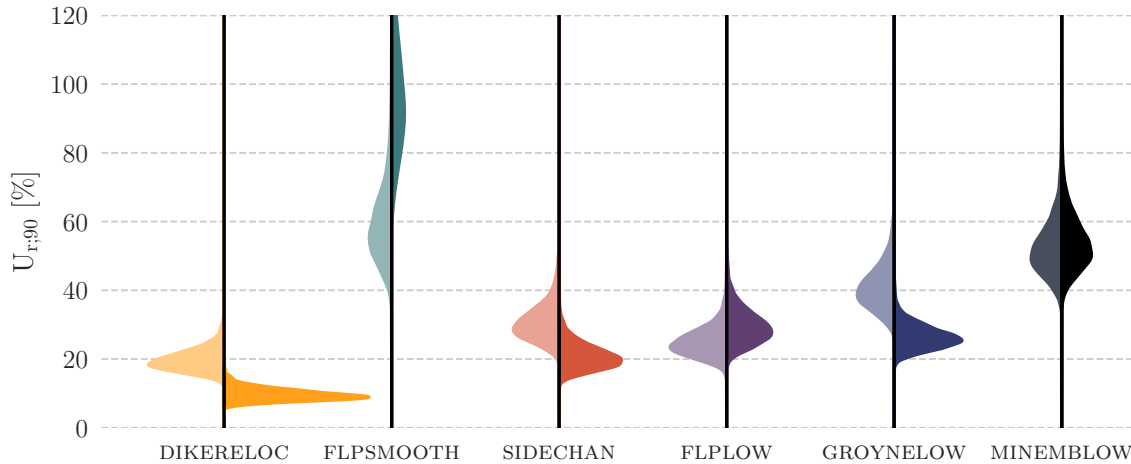



Figure 8. Probability densities of the relative uncertainty ($U_{r;90}$) at the location of maximum effect (x_m), for all interventions, showing that FLPSMOOTH is significantly more uncertain than other interventions. Most interventions show similar distributions for the low intensity (left, light coloured) and high intensity (right, dark coloured) variants.

A straightforward way to compare differences between the effect of two or more interventions is to look at the backwater figures. In Fig. 6 we see two diagrams of interventions that result in flood level decrease within the same order of magnitude, but with very different uncertainty. The uncertainty of FLPSMOOTH is much higher than that of SIDECHAN, as represented by the wider confidence intervals.

To systematically compare interventions we compute the relative uncertainty $U_{r;90}$ at the location of the maximum effect for each intervention (Figure 8). Since $U_{r;90}$ is stochastic, we visualised the kernel density estimation of their probability distributions. In general we observe that (i) FLPSMOOTH and MINEMBLOW are more uncertain than the other four interventions. This can be seen by their higher values for $U_{r;90}$, especially for the high-intensity variants. We also observe that overall, the low-intensity variants have similar distributions compared to their high-intensity variants, both in shape and location. Even in the cases where the high-intensity variant distribution is clearly different from the low-intensity variant (DIKERELOC, GROYNLOW and FLPSMOOTH), the distribution still overlap. This shows that the intensity of an intervention is not a deciding explaining factor for $U_{r;90}$, even though it is evidently important in explaining the absolute uncertainty (MCI$_{90}$) and the expected effect. The latter follows from the earlier observation that the absolute uncertainty scales with the expected effect following a (constant) relative uncertainty. For decision making, it may be useful to estimate how much uncertainty is expected for a given expected effect. When considering different flood mitigation measures, the choice may depend on the level of acceptable uncertainty.

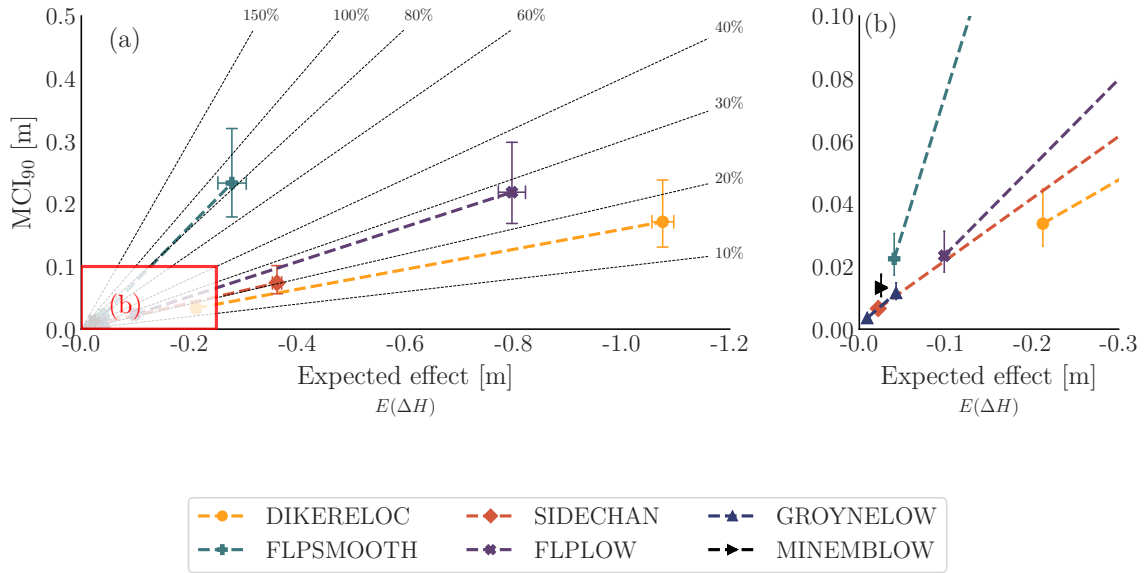



Figure 9. The expected effect ($E(\Delta H)$) at x_m against the 90% confidence interval, linearly interpolated between the low and high intensities. The relationship between the two is described by the relative uncertainty $U_{r;90}$, here shown as dashed grey lines. The error bars depict the estimation uncertainty (90% interval). The right hand panel (b) displays a zoomed-in part of the left-hand graph.

By linearly interpolating between the expected values of the low and high intensities, we obtain a first approximation of the uncertainty for a given expected effect for each intervention (Figure 9). For example, consider that the objective of a planned flood mitigation measure is 25 cm. We see from figure 9 that this can be accomplished by four different interventions (FLPSMOOTH, FLPLOW, SIDECHAN and DIKERELOC), within the bounds marked by the low- and high intensities. The confi-
5   dence intervals differ markedly between those choices. Under the assumption of linear interpolation between the two intensities, the 90% confidence intervals for an expected decrease of 25 cm are estimated at 4 cm (DIKERELOC), 5 cm (SIDECHAN), 7 cm (FLPLOW) and 21 cm (FLPSMOOTH). These intervals may or may not be acceptable.

Instead of looking at the expected effect, the confidence intervals can be marginalised by designing for a given exceedance probability of the effect. The desired effect is related to the expected effect through the exceedance probability. To illustrate
10   how the exceedance probability can be used to guide intervention design, we linearly interpolated between the low-intensity and high-intensity variants as function of the expected effect (Figure 10). In this figure, the estimation uncertainty is not taken into account. Given that we only have two points (low and high intensity), a linear transformation is the only model we can support. Although we acknowledge that other models are possible, we included this as a first approximation, because it illustrates information available for decision makers.

The high-intensity variant of FLPSMOOTH has an expected flood level decrease of 28 cm. However, the likelihood that the decrease is smaller (or higher) than that number is only approximately 50%. This means that the expected effect of 28 cm is as likely as not to be met by the proposed intervention. In contrast, the 90% exceedance probability for the high-intensity variant is 18.7 cm. This means that it is "very likely" (following IPCC terminology, see Solomon et al. (2007)) that the flood level decrease is larger than 18.7 cm.

This approach can be applied to the earlier example of a planned mitigation measure with a 25 cm flood level decrease objective. In figure 10 (b), we see that at an exceedance probability of 66% ("Likely"), all four possible interventions can be implemented. However, to meet this likelihood all interventions have to be "over-designed", i.e. to meet a larger expected flood level decrease than the given objective. For example, floodplain smoothing needs to be designed for an expected effect of 28 cm for a 66% likelihood of reaching the objective of 25 cm. The amount by which the measure needs to be over-designed depends on the uncertainty of that measure and the exceedance probability. Higher exceedance probabilities such as a 95% likelihood ("Extremely likely") can also preclude some flood mitigation measures (Figure 10 (c)). With this level of confidence, FLPSMOOTH can no longer be considered: none of the computed examples reach an 'extremely likely' flood level decrease of 25 cm within the limits of the considered interventions. In this case, there is a physical limit to an even higher intensity FLPSMOOTH, since there is simply no more available land within the study area to turn into meadow. Extrapolation beyond the given limits is therefore not possible, although a higher effect could be obtained by simply enlarging the study area. Within the given limits, the three remaining interventions (DIKERELOC, SIDECHAN and FLPLOW) can still be used. They would need to be designed for an expected effect of -0.27m , -0.28m and -0.29m, respectively. For perspective, it is worthwhile to point out that one of the greatest interventions on the River Waal in the past decades, the Nijmegen-Lent dike relocation project, has (deterministically) predicted effect of -27 cm, at an estimated cost of more than ten million Euro per cm flood level decrease.

4 Discussion

Our results show that some interventions are inherently more uncertain than other interventions. Given the complexity of high-detailed modelling of floodplains, it is not straightforward to explain these differences. The statistical causes of uncertainty are discussed in detail by Berends et al. (2018). They observed that the removal of existing stochastic elements (e.g., a vegetation polygon) in exchange for new ones greatly increased the unexplained variance (ϵ in (eq. 3)) and therefore increases effect uncertainty. We see that here as well: FLPSMOOTH, which removes high-friction vegetation in exchange for meadow, is by far the most uncertain intervention. However, in this study we have argued that the relative uncertainty ($U_{r;90}$) is a more useful parameter than the absolute uncertainty (MCI). In that context, a large absolute uncertainty can be offset by a large expected flood level decrease. This is evidenced by both DIKERELOC and FLPLOW, which both have a large absolute uncertainty at high intensity but a relatively low relative uncertainty. A reduction in relative uncertainty could therefore be achieved by minimising the amount of change to existing floodplains while optimising the expected flood level decrease. This general observation is likely applicable to other rivers and case studies.

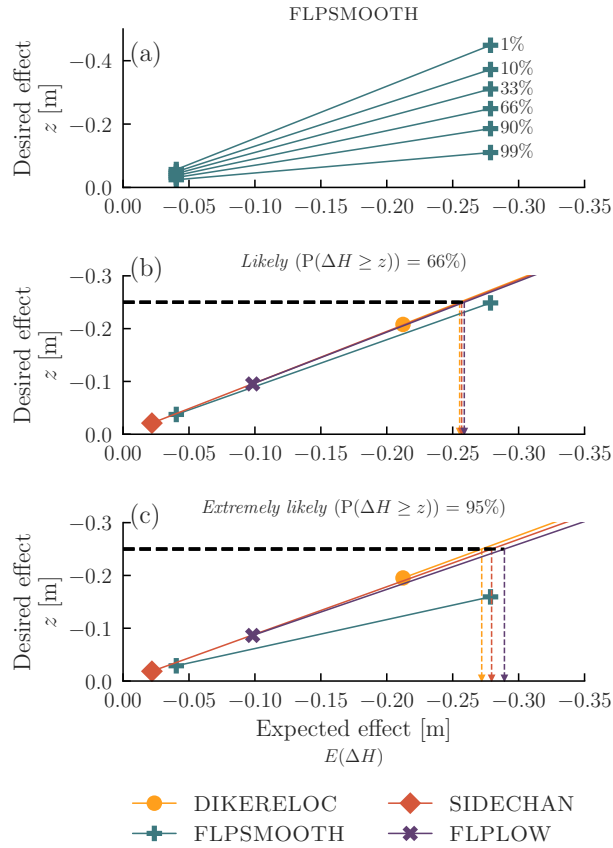



Figure 10. (a) The relationship between the expected effect $E(\Delta H)$ and the exceedance level z for various exceedance probabilities. (b) Inter-comparison of a 66% exceedance probability for four interventions. The dashed black line indicates a z of 25 cm, and the coloured dashed lines indicate which expected level belong to that z, given the exceedance probability (see text). (c) same as (b), but for an exceedance probability of 99%. Note that all lines show the expected value of the exceedance level; the estimation uncertainty is not depicted.

A lack of computational resources is often named as the main obstacle that motivates the use of deterministic approaches over probabilistic approaches. The analysis performed in this study is nonetheless feasible due to the CORAL approach, which reduces the required computer resources by reducing the number of model evaluations. An additional advantage of this method is that the uncertainty of the estimation is known and can be explicitly incorporated in analyses. In this study, we show that the relative uncertainty $U_{r;90}$ can be estimated to a sufficient precision (Fig. 8) to draw conclusions, even with a relatively low number of model runs (20 for each intervention). It is possible to decrease the estimation uncertainty by increasing the number of model runs with the post-intervention models (Berends et al., 2018), if (computational) resources allow.

Our study focused on providing uncertainty estimations for many different interventions, not to provide an optimal intervention for the given case study. Following this approach we made several simplifications, that must be resolved when an optimal intervention design is pursued. The linear interpolation between low intensity and high intensity interventions in figures 9 and 10 was used to obtain a first approximation of the uncertainty for interventions at different intensities. These figures are useful to illustrate how confidence levels and differences in relative uncertainty between interventions affect the expected effect but should not replace new calculations in intervention design – in part because our study provides only two support points, which is insufficient to either support or oppose the assumption of linearity. Rather, they may help to provide a starting point for a new set of calculations. Second, we marginalised the estimation uncertainty in figure 10 by averaging, to show how model uncertainty implicates intervention design. Practical use of exceedance levels should include estimation uncertainty of those levels — which can be reduced by increasing the number model evaluations if necessary. Finally, the interventions studied in this paper are single archetypes, while in reality designs would likely be a combination of multiple archetypes (e.g. side channel combined with lowering of the surrounding floodplain). Our study does not provide support for claims regarding such compound interventions, which would require further study.

Our approach is a form of forward uncertainty analysis: the quantification of model output uncertainty based on uncertainty in model assumptions without taking into account historical evidence of the goal variable (i.e., water level measurements). We justify choosing this approach based on the assumption that no historical information can be used to infer or narrow the distributions of the uncertain parameters, because (a) such information does not exist for our study case and (b) measurements after the intervention will never be available *ex ante*. The results of our analysis are therefore conditional on our assumptions about the validity of the model, the selection uncertainty sources and the probability distributions of these sources. We have taken care to select the stochastic variables, and their distributions, based on previous research specific to our study case. Forward uncertainty analysis is contrasted (e.g., see Beven et al., 2018) with so-called inverse methods, such as GLUE (Beven and Binley, 2014), DREAM (Vrugt et al., 2008) or deterministic calibration, that use historical evidence to potentially narrow the uncertainty bands of input variables. However, the potential use of historical data to constrict uncertainty faces complex challenges which will have to be met. Chief among them are extrapolation to unmeasured conditions and dealing with the large number of uncertain variables (here, 39) which may not be uniquely identifiable (Werner et al., 2005). In any case, like traditional deterministic approaches, the assumptions underlying probabilistic approaches should be readily available, transparent, and open to discussion.

Models can play a role in decision support under (deep) uncertainty and the uncertainty quantification is an important step to model the future (Maier et al., 2016). However, too much focus on quantifiable uncertainty should be avoided. Warmink et al. (2017) documented several examples where too much focus on quantifiable (also "statistical" or "technical") uncertainty resulted in knowledge creation that did not contribute to the policy process ("superfluous knowledge"). A key idea is that

5 other uncertainties exist that cannot be reduced by more data or simply cannot be quantified. Examples of these include 'social uncertainty': fundamental disagreement between experts, lack of trust in scientific results, ambiguity and diverging narratives based on the same empirical basis (Sarewitz, 2004; Brugnach et al., 2011; Warmink et al., 2017) Coping with such uncertainties may require adaptive river management (Pahl-Wostl, 2009) and exploration of multiple alternate futures (See the third paradigm of Maier et al., 2016, and references therein). Given these considerations, there are several probabilistic frameworks

10 of flood risk adaptation strategies to which explicit quantification of model uncertainty is a natural and required extension (e.g. Lendering et al., 2018; Klijn et al., 2018). In this paper, we demonstrated that model predictions of human intervention in rivers are sensitive to uncertainty and can substantially influence design decisions. Therefore, we argue that trusting only on the best available knowledge and, by extension, prediction by the best-available models, is not sufficient for planning interventions in river systems. Finally, our study was limited to the effect of interventions with respect to water level lowering. We did not

15 consider other indicators, such as economic, ecological or societal costs and benefits, as was done by Straatsma et al. (2019) (albeit without addressing uncertainty). Multidisciplinary and multisectoral assessment of interventions including uncertainty is recommended to determine trade-offs in river management. CORAL provides a computationally effective method to do so.

5 Conclusions

In general, our study shows that explicitly quantifying the uncertainty of predicted flood mitigation measures provides valuable

20 information to decision makers and modellers. On the one hand, results show how taking into account uncertainty can lead to different design choices. On the other hand, even small effects on flood levels can be quantified, because small effects are accompanied by small uncertainties. This shows that model uncertainty does not invalidate model-supported decision making in river management, but enriches it.

Our first objective was to quantify the effect of parameter uncertainty on the predicted effectiveness of flood mitigation

25 measures. Based on previous studies, we quantified the parameter uncertainty for 39 variables and estimated the uncertainty of model output. Results show that the absolute uncertainty of the predicted effect of flood level decrease is highly dependent on the type of intervention and location along the river. However, we found that the confidence bounds of flood level decrease along the river can be adequately described by a single "relative uncertainty" metric, defined as the ratio between the 90% confidence interval and the expected effect. This ratio remains relatively constant along the river and between intensities of intervention

30 types, and enables us to make some general observations. First, all uncertainty is 'generated' where the intervention has modified the river system. Upstream from there, the uncertainty gradually diminishes upstream with a constant rate following typical backwater curves. Second, a higher expected flood level decrease lead to a higher uncertainty, and a small flood level decrease was accompanied by a small uncertainty. The ranges of the expected relative uncertainty varied between 15% and

40% for most interventions, which means that the size of the 90% confidence bounds of those interventions is less than half of the expected flood level decrease.

The second objective was to explore the relationship between the expected effect and its uncertainty, to aid in the decision making process. We observe that interventions of different types may reach the same expected flood level decrease, but having different uncertainty. Specifically, a large-scale but relatively ineffective intervention such as floodplain smoothing (by removing high-friction vegetation) has a high relative uncertainty compared to alternative interventions. The intensity of an intervention (e.g. total area of vegetation smoothed) may be increased to reach a higher effect. Our results show that higher intensity also leads to a higher uncertainty, while the relative uncertainty remains approximately constant. This relationship was then used to show how explicit uncertainty quantification and differences in relative uncertainty between various interventions may affect design choices, depending on the level of acceptable uncertainty. For a fixed level of acceptable uncertainty (i.e., by a given exceedance probability), we graphically demonstrated that interventions need to be designed for a larger expected flood level decrease than the given objective.

Code availability. In this article we used the following code and software. For the hydraulic modelling we used Delft3D Flexible Mesh (FM) version 1.1.261.52670. Delft3D FM is available from (https://oss.deltares.nl/web/delft3dfm). For the generation of the various flood mitigation measures we used the Python tool RiverScape (for availability see Straatsma and Kleinhans (2018)). The uncertainty quantification method CORAL is available from GitHub (https://github.com/kdberends/coral.).

Appendix A: Parameter distributions for vegetation types

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study is part of the research programme RiverCare, supported by the domain Applied and Engineering sciences (AES), which is part of the Netherlands Organisation for Scientific research (NWO), and which is partly funded by the Ministry of Economic Affairs, under grant number P12-14 (Perspective programme). This study benefited from cooperation within the Netherlands Centre for River studies (NCR). We thank Jan Kwakkel and Joseph Guillaume for their critical contributions in review, which have improved the quality and positioning of this work significantly.

Table A1. Parameters of the floodplain roughness class distributions. Codes marked with an asterisk * are not used in the Waal model

Class Code	Name	Parameters			
		Empirical distribution			
612 - 637	Alluvial bed				
		Uniform distribution			
n/a	Classification map				
		Triangular distributions			
		min	*mean*	*max*	
102	Deep bed	0.025	0.03	0.033	
104	Natural side channel	0.03	0.035	0.04	
105	Side channel	0.025	0.03	0.033	
106	Pond / Harbor	0.025	0.03	0.033	
111	Sand bank	0.025	0.03	0.033	
121	Field	0.02	0.03	0.04	
	Lognormal distributions	μ_{h_v}	σ_{h_v}	μ_{n_v}	σ_{n_v}
1201	Production meadow	-3.18	0.47	2.40	0.77
1202	Natural grass and hayland	-0.74	0.53	-2.64	0.93
1203	Herbaceous meadow	-1.64	0.32	2.59	0.33
1211*	Thistle herb. Veg.	-1.29	0.33	1.05	0.43
1212	Dry herbaceous vegetation	-0.59	0.39	-3.06	0.65
1213*	Brambles	-0.67	0.21	-0.73	0.36
1214*	Hairy willowherb	-1.89	0.56	-0.25	0.49
1215*	Reed herb. Veg.	0.60	0.22	-1.83	0.27
1221	Wet herb. Veg.	-1.08	0.38	-1.49	0.44
1222*	Sedge	-1.32	0.67	0.04	0.63
1223	Reed-grass	-0.92	0.86	-2.19	0.16
1224	Bulrush	-0.81	0.67	0.04	0.63
1225*	Reed-mace	0.37	0.23	-1.12	0.57
1226	Reed	0.94	0.13	-1.14	0.42
1231	Softwood shrubs	1.81	0.24	-2.20	0.79
1232	Willow plantation	1.05	0.43	-3.23	0.62
1233	Thorny shrubs	1.48	0.64	-1.73	0.41
1241*	Hardwood production forest	*Deterministic*	*Deterministic*	-4.68	0.67
1242	Softwood production forest	*Deterministic*	*Deterministic*	-4.72	0.66
1243*	Pine forest	*Deterministic*	*Deterministic*	-4.18	0.54
1244	Hardwood forest	*Deterministic*	*Deterministic*	-3.45	0.77
1245	Softwood forest	*Deterministic*	*Deterministic*	-3.04	0.99
1246	Orchard low	1.10	0.10	-3.72	0.25
1247	Orchard high	1.78	0.21	-4.61	0.12
1250	Pioneer vegetation	-2.87	0.18	-1.93	0.50

References

Baldassarre, G. D., Kooy, M., Kemerink, J. S., and Brandimarte, L.: Towards understanding the dynamic behaviour of floodplains as human-water systems, Hydrology and Earth System Sciences, 17, 3235–3244, https://doi.org/10.5194/hess-17-3235-2013, 2013.

Baptist, M. J., Babovic, V., Uthurburu, J. R., Keijzer, M., Uittenbogaard, R. E., Mynett, A., and Verwey, A.: On inducing equations for vegetation resistance, Journal of Hydraulic Research, 45, 435–450, https://doi.org/10.1080/00221686.2007.9521778, 2007.

Becker, A., Scholten, M., Kerkhoven, D., and Spruyt, A.: Das behördliche Modellinstrumentarium der Niederlande, in: 37. Dresdner Wasserbaukolloquium 2014 "Simulationsverfahren und Modelle für Wasserbau und Wasserwirtschaft", edited by Stamm, J., pp. 539–548, http://d-nb.info/105023037X, 2014.

Berends, K. D., Warmink, J. J., and Hulscher, S. J. M. H.: Efficient uncertainty quantification for impact analysis of human interventions in rivers, Environmental Modelling & Software, 107, 50–58, https://doi.org/10.1016/j.envsoft.2018.05.021, 2018.

Beven, K. and Binley, A.: GLUE: 20 years on, Hydrological Processes, 28, 5897–5918, https://doi.org/10.1002/hyp.10082, http://dx.doi.org/10.1002/hyp.10082, 2014.

Beven, K. J., Aspinall, W. P., Bates, P. D., Borgomeo, E., Goda, K., Hall, J. W., Page, T., Phillips, J. C., Simpson, M., Smith, P. J., Wagener, T., and Watson, M.: Epistemic uncertainties and natural hazard risk assessment. 2. What should constitute good practice?, Natural Hazards and Earth System Sciences, 18, 2769–2783, https://doi.org/10.5194/nhess-18-2769-2018, 2018.

Bormann, H. and Pinter, N.: Trends in low flows of German rivers since 1950: Comparability of different low-flow indicators and their spatial patterns, River Research and Applications, 33, 1191–1204, https://doi.org/10.1002/rra.3152, 2017.

Brugnach, M., Dewulf, A., Henriksen, H., and van der Keur, P.: More is not always better: Coping with ambiguity in natural resources management, Journal of Environmental Management, 92, 78–84, https://doi.org/10.1016/j.jenvman.2010.08.029, 2011.

Chbab, E. H.: How extreme were the 1995 flood waves on the rivers Rhine and Meuse?, Physics and Chemistry of the Earth, 20, 455–458, https://doi.org/10.1016/s0079-1946(96)00005-5, http://dx.doi.org/10.1016/S0079-1946(96)00005-5, 1995.

Chow, V. T.: Open-Channel Hydraulics, McGraw-Hill, New York, 1959.

Hobo, N., Makaske, B., Wallinga, J., and Middelkoop, H.: Reconstruction of eroded and deposited sediment volumes of the embanked River Waal, the Netherlands, for the period ad 1631–present, Earth Surface Processes and Landforms, 39, 1301–1318, https://doi.org/10.1002/esp.3525, eSP-13-0264.R1, 2014.

Huthoff, F., Augustijn, D. C., and Hulscher, S. J.: Analytical solution of the depth-averaged flow veocity in case of submerged rigid cylindrical vegetation, Water Resources Research, 43, https://doi.org/10.1029/2006WR005625, 2007.

Jongman, B., Ward, P. J., and Aerts, J. C. J. H.: Global exposure to river and coastal flooding: Long term trends and changes, Global Environmental Change, 22, 823 – 835, https://doi.org/http://dx.doi.org/10.1016/j.gloenvcha.2012.07.004, 2012.

Kernkamp, H. W. J., Dam, A. V., Stelling, G. S., and de Goede, E. D.: Efficient scheme for the shallow water equations on unstructured grids with application to the Continental Shelf, Ocean Dynamics, 61, 1175–1188, https://doi.org/10.1007/s10236-011-0423-6, 2011.

Klijn, F., de Bruin, D., de Hoog, M. C., Jansen, S., and Sijmons, D. F.: Design quality of room-for-the-river measures in the Netherlands: role and assessment of the quality team (Q-team), International Journal of River Basin Managament, 11, 287–299, https://doi.org/10.1080/15715124.2013.811418, 2013.

Klijn, F., Asselman, N., and Wagenaar, D.: Room for Rivers: Risk Reduction by Enhancing the Flood Conveyance Capacity of The Netherlands' Large Rivers, Geosciences, 8, 224, https://doi.org/10.3390/geosciences8060224, 2018.

Klopstra, D., Barneveld, H. J., van Noortwijk, J. M., and van Velzen, E. H.: Analytical model for hydraulic roughness of submerged vegetation, in: The 27th Congress of the International Association for Hydraulic Research, San Francisco, 1997; Proceedings of Theme A, Managing Water: Coping with Scarcity and Abundance, pp. 775–780, 1997.

Knotters, M. and Brus, D. J.: Purposive versus random sampling for map validation: a case study on ecotope maps of floodplains in the Netherlands, Ecohydrology, 6, 425–434, https://doi.org/10.1002/eco.1289, 2012.

Lendering, K. T., Sebastian, A., Jonkman, S. N., and Kok, M.: Framework for assessing the performance of flood adaptation innovations using a risk-based approach, Journal of Flood Risk Management, p. e12485, https://doi.org/10.1111/jfr3.12485, 2018.

Li, S., Shi, H., Xiong, Z., Huai, W., and Cheng, N.: New formulation for the effective relative roughness height of open channel flows with submerged vegetation, Advances in Water Resources, 86, 46–57, https://doi.org/10.1016/j.advwatres.2015.09.018, 2015.

Luhar, M. and Nepf, H. M.: From the blade scale to the reach scale: A characterization of aquatic vegetative drag, Advances in Water Resources, 51, 305–316, https://doi.org/10.1016/j.advwatres.2012.02.002, 2013.

Maier, H. R., Guillaume, J. H. A., van Delden, H., Riddell, G. A., Haasnoot, M., and Kwakkel, J. H.: An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together?, Environmental Modelling & Software, 81, 154–164, https://doi.org/10.1016/j.envsoft.2016.03.014, 2016.

Metropolis, N.: The Beginning of the Monte Carlo Method, Los Alamos Science, Special Issue 1987, 125 – 130, 1987.

Middelkoop, H., de Boo, M., Breukel, R., and van Haselen, C.: Twice a river: Rhine and Meuse in the Netherlands, 1999.

Mosselman, E.: Modelling in Applied Hydraulics: More Accurate in Decision-Making Than in Science?, in: Advances in Hydroinformatics, pp. 741–749, Springer Singapore, https://doi.org/10.1007/978-981-10-7218-5_53, 2018.

Munoz, S. E., Giosan, L., Therrell, M. D., Remo, J. W. F., Shen, Z., Sullivan, R. M., Wiman, C., O'Donnell, M., and Donnelly, J. P.: Climatic control of Mississippi River flood hazard amplified by river engineering, Nature, 556, 95–98, https://doi.org/10.1038/nature26145, 2018.

Nepf, H. M.: Hydrodynamics of vegetated channels, Journal of Hydraulic Research, 50, 262–279, https://doi.org/10.1080/00221686.2012.696559, 2012.

Oreskes, N., Shrader-Frechtette, K., and Belitz, K.: Verification, Validation and Confirmation of Numerical Models in the Earth Sciences, Science, 263, 641–646, https://doi.org/10.1126/science.263.5147.641, 1994.

Pahl-Wostl, C.: A conceptual framework for analysing adaptive capacity and multi-level learning processes in resource governance regimes, Global Environmental Change, 19, 354–365, https://doi.org/10.1016/j.gloenvcha.2009.06.001, 2009.

Pappenberger, F. and Beven, K. J.: Ignorance is bliss: Or seven reasons not to use uncertainty analysis, Water Resources Research, 42, https://doi.org/10.1029/2005WR004820, 2006.

Pinter, N., Ickes, B. S., Wlosinski, J. H., and van der Ploeg, R. R.: Trends in flood stages: Contrasting results from the Mississippi and Rhine River systems, Journal of Hydrology, 331, 554–566, https://doi.org/10.1016/j.jhydrol.2006.06.013, 2006.

Razavi, S., Tolson, B. A., and Burn, D. H.: Review of surrogate modeling in water resources, Water Resources Research, 48, W07 401, https://doi.org/10.1029/2011WR011527, http://dx.doi.org/10.1029/2011WR011527, 2012.

Rijke, J., van Herk, S., Zevenbergen, C., and Ashley, R.: Room for the River: delivering integrated river basin management in the Netherlands, International Journal of River Basin Management, 10, 369–382, https://doi.org/10.1080/15715124.2012.739173, 2012.

Sarewitz, D.: How science makes environmental controversies worse, Environmental Science & Policy, 7, 385–403, https://doi.org/10.1016/j.envsci.2004.06.001, 2004.

Shields, F. D., Coulton, K. G., and Nepf, H.: Representation of Vegetation in Two-Dimensional Hydrodynamic Models, Journal of Hydraulic Engineering, 143, 02517 002, https://doi.org/10.1061/(ASCE)HY.1943-7900.0001320, 2017.

Sieben, J.: Sediment management in the Dutch Rhine branches, International Journal of River Basin Management, 7, 43–53, https://doi.org/10.1080/15715124.2009.9635369, 2009.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and (eds.), H. M.: Climate Change 2007 - The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, 2007.

Stefanou, G.: The stochastic finite element method: Past, present and future, Computer Methods in Applied Mechanics and Engineering, 198, 1031–1051, https://doi.org/10.1016/j.cma.2008.11.007, 2009.

Stone, B. M. and Shen, H. T.: Hydrauli Resistance of Flow in Channels with Cylindrical Roughness, Journal of Hydraulic Engineering, 128, 500–506, https://doi.org/10.1061/(ASCE)0733-9429(2002)128:5(500), 2002.

Straatsma, M. and Alkema, D.: Error propagation in hydrodynamics of lowland rivers due to uncertainty in vegetation roughness parameterization, Tech. Rep. ITC-FC2015 report:2009.06.05, ITC, 2009.

Straatsma, M. and Huthoff, F.: Uncertainty in 2D hydrodynamic models from erros in roughness parameterization based on aerial images, Physics and Chemistry of the Earth, 36, 324–334, https://doi.org/10.1016/j.pce.2011.02.009, 2011.

Straatsma, M. W. and Kleinhans, M. G.: Flood hazard reduction from automatically applied landscaping measures in River-Scape, a Python package coupled to a two-dimensional flow model, Environmental Modelling & Software, 101, 102–116, https://doi.org/10.1016/j.envsoft.2017.12.010, 2018.

Straatsma, M. W., Bloecker, A. M., Lenders, H. J. R., Leuven, R. S. E. W., and Kleinhans, M. G.: Biodiversity recovery following delta-wide measures for flood risk reduction, Science Advances, 3, e1602 762, https://doi.org/10.1126/sciadv.1602762, 2017.

Straatsma, M. W., Fliervoet, J. M., Kabout, J. A. H., Baart, F., and Kleinhans, M. G.: Towards multi-objective optimization of large-scale fluvial landscaping measures, Natural Hazards and Earth System Sciences, 19, 1167–1187, https://doi.org/10.5194/nhess-19-1167-2019, 2019.

Uusitalo, L., Lehikoinen, A., Helle, I., and Myrberg, K.: An overview of methods to evaluate uncertainty of deterministic models in decision support, Environmental Modelling & Software, 63, 24–31, https://doi.org/10.1016/j.envsoft.2014.09.017, 2015.

van Denderen, R. P., Schielen, R. M., Westerhof, S. G., Quartel, S., and Hulscher, S. J.: Explaining artificial side channel dynamics using data analysis and model calculations, Geomorphology, 327, 93–110, https://doi.org/10.1016/j.geomorph.2018.10.016, 2019.

van der Mark, C. F.: A semi-analytical model for form drag of river bedforms, Ph.D. thesis, 2009.

van Velzen, E. H., Jesse, P., Cornelissen, P., and Coops, H.: Stromingsweerstand vegetatie in uiterwaarden (versie 1-2003) Deel 1: Handboek, Rijksinstituut voor Integraal Zoetwaterbeheer en Afvalwaterbehandeling (RIZA) Rijkswaterstaat, 2003.

Vargas-Luna, A., Crosato, A., and Uijttewaal, W. S. J.: Effects of vegetation on flow and sediment transport: comparative analyses and validation of prediction models, Earth Surf Proc Land, 40, 157–176, https://doi.org/10.1002/esp.3633, 2015.

Velsen, C.: De Rivierkundige Verhandeling, Arnhem: Jacob Nyhoff, 1749.

Vrugt, J. A., ter Braak, C. J., Gupta, H. V., and Robinson, B. A.: Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, Stochastic Environmental Research and Risk Assessment, 23, 1011–1026, https://doi.org/10.1007/s00477-008-0274-y, 2008.

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., and von Krauss, M. P. K.: Defining Uncertainty: a conceptual basis for uncertainty management in model-based decision support, Integrated Assessment, 4, 5–17, https://doi.org/10.1076/iaij.4.1.5.16466, 2003.

Warmink, J. J., van der Klis, H., Booij, M. J., and Hulscher, S. J. M. H.: Identification and quantification of uncertainties in a hydrodynamic river model using expert opinion elicitation, Water Resources Management, 25, 601–622, https://doi.org/10.1007/s11269-010-9716-7, 2011.

Warmink, J. J., Booij, M. J., van der Klis, H., and Hulscher, S. J. M. H.: Quantification of uncertainty in Design water levels due to uncertain bed form roughness in the Dutch rivier Waal, Hydrological Processes, 27, 1646–1663, https://doi.org/10.1002/hyp.9319, 2013a.

Warmink, J. J., Straatsma, M. W., Huhoff, F., Booij, M. J., and Hulscher, S. J. M. H.: Uncertainty of design water levels due to combined bed form and vegetation roughness in the Dutch River Waal, Journal of Flood Risk Managament, 6, 302–318, https://doi.org/10.1111/jfr3.12014, 2013b.

Warmink, J. J., Brugnach, M., de Kruijf, J. V., Schielen, R. M. J., and Augustijn, D. C. M.: Coping with Uncertainty in River Management: Challenges and Ways Forward, Water Resources Management, 31, 4587–4600, https://doi.org/10.1007/s11269-017-1767-6, 2017.

Werner, M. G. F., Hunter, N. M., and Bates, P. D.: Identifiability of distributed floodplain roughness values in flood extent estimation, Journal of Hydrology, 314, 139–157, https://doi.org/10.1016/j.jhydrol.2005.03.012, 2005.

White, G.: Human adjustment to floods, 1945.

Yang, W. and Choi, S.-U.: A two-layer approach for depth-limited open-channel flows with submerged vegetation, Journal of Hydraulic Research, 48, 466-475, https://doi.org/10.1080/00221686.2010.491649, 2010.