

## Combined point-by-point reply to the comments

*To reviewer comments for discussion paper <https://doi.org/10.5194/nhess-2018-325>  
Berends, K.D., Straatsma, M.W., Warmink, J.J., Hulscher, S.J.M.H., 17 April, 2019*

All referee comments are numbered by referee number and comment number, e.g. [R1.2] for the second comment of the first referee. The author's response is colour coded:

Black - Reviewer comment

Blue - Author reply

Green - Proposed change in manuscript

### Replies to referee 1

We would like to thank the reviewer for taking the time to review our manuscript. We highly appreciate his suggestions and comments, which are helpful in improving the manuscript. Below we have replied to the various comments made by the reviewer.

[R1.1] This article assesses the impact of 12 interventions on flood water levels in the Waal River, including estimation of uncertainty. It goes further than most other studies in 1) estimating uncertainty in uncertainty attributable to the methods used, 2) discussing the potential to approximate the uncertainty, in ways that make the results of the analysis more accessible to a broader audience. The article is well written and clear. The issues I would like to highlight are relatively minor - clarifying some claims rather than fundamentally questioning them. I do suggest some small extensions to the analyses to help in clarifying these claims.

[Reply to R1.1] We are grateful for the kind words, and the acknowledgement of the reviewer that our efforts to quantify uncertainty by approximation makes uncertainty quantification more accessible to a broader audience.

[R1.2] I would rather that the abstract avoid emphasising that "relative uncertainty" is a "newly introduced" metric. As the authors do acknowledge in the main text, it is closely related to the coefficient of variation. If this metric is kept as a contribution, it would require greater discussion of why is preferred over the CV. The greater contribution here, in any case, is that relative uncertainty is identified as a key parameter to obtain approximations of uncertainty in this type of problem.

[Reply to R1.2] We agree with the reviewer that the introduction of relative uncertainty as a metric is not the key contribution of this paper, rather that the relative uncertainty is an effective metric to discuss the uncertainty of engineering interventions, and that this should be reflected in the abstract. In the abstract, we wrote:

“ (...) We show that the uncertainty of an intervention can be adequately described by the newly introduced 'relative uncertainty' metric, defined as the ratio between the confidence interval and the expected effect. (...)”

We changed this sentence to:

[P1 L8-9 in modified manuscript]

“(...) 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. (...)”

[R 1.3] p9 L25 "We would need to obtain both the model output probability distributions, as well as the covariances between the two distributions" This is not entirely true - most of the parameters would be identical in both states of the river, so the difference can be calculated directly as a single distribution - covariances probably do not need to be calculated. The authors do end up doing this with the "union stochastic parameter space" (p11), so it seems worthwhile to clarify that you do not explicitly estimate covariances. It is still correct that this is probably not tractable for 12 interventions.

[Reply to R 1.3] The reviewer is correct in noting the covariances are not explicitly computed (or estimated). To be clear, we are not referring to the covariances between the stochastic parameters (which are in our case identical between the two states of the river), but the covariances of the distributions of the computed water levels. These should always be considered, either explicitly (if done analytically) or implicitly. Otherwise, the uncertainty of the intervention will be needlessly overestimated. Furthermore, the problem is \*analytically\* intractable because of the numerical models involved, regardless of any covariance.

We wrote [p9, L23-L27]:

“To 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 not analytically tractable, but can be solved numerically by Monte Carlo simulation (MCS; Metropolis, 1984; Stefanou, 2009).”

We changed these sentences to:

[p9, L18-21 in modified manuscript]

“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 these two distributions. This problem is analytically intractable, but can be solved numerically by Monte Carlo simulation (MCS; Metropolis, 1984; Stefanou, 2009), while accounting for covariance by sampling from a union stochastic parameter space

(Berends et al., 2018). In the numerical approach, covariances are not explicitly computed.”

[R1.4] p15 To support the claim that values are relatively constant, it would be useful to show the range of  $E(Ur90)$  over the length of the river, either in Figure 7 or Table 2.  $E(\text{range}(Ur90))$  would also be useful, given it is available, i.e. the range of  $Ur90$  over the length of the river, averaged over all realisations. There is enough variation there assuming relative uncertainty is constant seems like it would still be a substantial simplification.

[Reply to R1.4] The first suggestion of the reviewer is to add the range of  $E(Ur;90)$  to either Fig. 7 or Table 2. In response, we will add the standard deviation of  $E(Ur;90)$  to Table 2 to explicitly show the variation along the channel. In the text, we wrote

[p14, L14-L16]:

“Next, we compute the relative uncertainty ( $Ur;90$ ) for each intervention and intensity. Since  $Ur;90$  is stochastic, we first reduce this by calculating the expected value ( $E(Ur;90)$ ) for each point along the river (Fig. 7). Results show that values for  $E(Ur;90)$  are fairly constant over the river length up to the location of the maximum flood level decrease ( $x_m$ ) for all interventions.”

We changed these sentences to refer to the new information in Table 2:

[p14, L8-11]

“Next, we compute the relative uncertainty ( $Ur;90$ ) for each intervention and intensity. We first marginalise the uncertainty in  $Ur;90$  (which is due to CORAL estimation) by calculating the expected value ( $E(Ur;90)$ ) 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  $E(Ur;90)$  are fairly constant over the river length up to the location of the maximum flood level decrease ( $x_m$ ) for all interventions.”

The second suggestion of the reviewer, is to show  $E(\text{range}(Ur;90))$ , as well. This would show the variation in the estimation uncertainty along the channel. In figure 8, we visualized the estimation uncertainty at  $x_m$ . We feel that including variation along the channel as well, would conflate the estimation uncertainty with the variation along the channel and as such not provide meaningful additional information, while making the figure more difficult to understand.

[R1.5] p16 Figure 8 "FLPSMOOTH is significantly more uncertain" The use of the word "significantly"

suggests this is a statistical statement. It might be worth putting a p-value on this, given you have uncertainty on the uncertainty. Visually, it seems high intensity FLPSMOOTH could be significantly more uncertain ( $p=0.05$ ), but that low intensity is

comparable to groyneflow or minemblow, also confirmed by Figure 7. L6-14 could also comment on statistical significance of differences in uncertainty.

[Reply to R1.5] Here, it was not our intention to provide a statistical significance of differences between the distributions, given that that the answer from any statistical test used to obtain a p-value would be conditional on both the intervention subsample size and the number of MCMC steps used in the approximation. To prevent confusion, we removed the word ‘significant’.

We wrote:

[caption Figure 8, p 16]:

“(…) showing that /flpsmooth/ is significantly more uncertain than other interventions.”

We changed this to:

[caption Figure 8, p 16 in modified manuscript]

“(…) showing that /flpsmooth/ is more uncertain than other interventions.”

[R1.6] p17 Figure 9 This figure is illustrative of the interpolation method, but currently only hints at how good the interpolation method would be. You have two sources of information on this: 1) the uncertainty in MCI90 and E(dH) 2) the variation in these values along the river The error bars for (1) in the figure do suggest that there is quite a bit of uncertainty involved in the interpolation. I would suggest showing a ribbon around the line encompassing all potential lines that could be interpolated using 1 and 2. This doesn’t detract from the utility of using linear interpolation, but would provide a better sense of the confidence one should have in using this simplification. If the uncertainty due to using CORAL is too large, perhaps more runs might also help make your point more effectively.

[Reply to R1.6] We agree with the reviewer that there is quite a bit of uncertainty involved in the interpolation and that multiple lines could be drawn between the high and low intensities. In fact, showing a ribbon around the line would still not resolve all uncertainty, as this still is conditional on the assumption of a linear model (i.e. a ribbon would show parameter uncertainty of the interpolation model, but not the model uncertainty of the interpolation model structure).

The main purpose of this figure, as with Figure 10 (see reviewer comment R1.7), is to show how uncertainty implicates intervention design. We believe that introducing more complexity, by showing every possible line that could be drawn (under the assumption of the linear model), would detract from this message. On the other hand, it should be clearly stated that the linear line is, at best, a first approximation. We made two changes in the manuscript to reflect this.

In the results section we wrote

[p16, L16 – p17, L1]:

“By linearly interpolating between the two low and high intensities, we can estimate the uncertainty for a given expected effect for each intervention (Figure 9).”

We changed this sentence to:

[p16, L11 – p12, L1 in modified manuscript]

“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).”

In the discussion we wrote

[P18, L25-28]:

“In studying the effect of uncertainty on intervention design (figures 9 and 10) we linearly interpolated between the high intensity and low intensity interventions. The assumption of linearity is reasonable given the approximately constant  $U_{r;90}$  between the two intensities. However, in practice this should be only seen as a first approximation. Multiple design iteration would be necessary to arrive at a design that meets the requirements.”

We changed and expanded upon this section:

[p20, L1 – L13 in modified manuscript]:

“Our study focused on providing uncertainty estimations for many different interventions, not to provide an optimal intervention 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.”

We further note that the variation alongside the river is no source of information for Figure 9, as this is strictly based on model prediction as  $x_m$  (the location of maximum effect). To

clarify this, as well as clarify the use of the error bars in this figure, we changed the caption of Figure 9 from:

“The expected effect ( $E(\Delta H)$ ) against the 90% confidence interval, (...)”

To

[caption of figure 9, p17 in modified manuscript]:

“The expected effect ( $E(\Delta H)$ ) at  $x_m$  against the 90% confidence interval, (...) The error bars depict the estimation uncertainty (90% interval). (...)”

[R 1.7] p17- exceedance probabilities If I understand correctly, this now linearly interpolates exceedance probabilities rather than relative uncertainty. It's not clear why this is a permissible inference. Are there characteristics of the distributions that allow this jump? What makes (all!) exceedance probabilities linear if relative uncertainty is constant with regard to expected effect? I assume at least some kind of symmetry comes into play? Even if the assumption of linearity doesn't end up holding completely, I would be in support of keeping this analysis in the paper, with some representation of the uncertainty in uncertainty involved (as for Figure 9). I think it's very important to be able to say something about the extremes of the distribution, not just about the size of the uncertainty.

[Reply to R 1.7] Since the exceedance levels are directly related to the support points on the (cumulative) density function, the implicit assumption in the linear interpolation is that the transformation of the distribution function for low to high is linear with respect to the expected value. This does not necessarily mean that either distribution should be symmetrical, or even that the distributions share the same distribution. Given that we only have two points (low and high intensity), a linear transformation is the only model we can support. To better clarify the purpose of this figure, as well as the uncertainty involved in the estimation, we made the following modifications:

We wrote

[P17, L8-L10]:

“For example, figure 10 (a) shows various exceedance probabilities for /flpsmooth/, linearly interpolated between the low-intensity and high-intensity variants as function of the expected effect. “

We changed this to:

[P17, L2-L7 in modified manuscript]

“To illustrate 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.”

We added the following to the caption of Figure 10:

[caption of figure 10, p19 in modified manuscript]

“Note that all lines show the expected value of the exceedance level; the estimation uncertainty is not depicted.”

[R1.8] p19 Discussion It would be useful to mention that the study only considered individual interventions, whereas it would be possible to combine measures in practice. I assume the combined effect is unlikely to be a linear combination of individual effects, so I don't think this study can be used to support any claims regarding such combinations. It would, however, be useful to highlight it as future work and speculate about possible issues that might crop up.

[Reply to R1.8] We agree with the reviewer that a combination of effects would not be linear and added the following section to the discussion (this follows our modification of P15, L25-L28 in response to R1.6):

[p20, L10-L13 in modified manuscript]

“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.”

Technical comments

[R1.9] Table 1 The link with subsections in 2.3 could perhaps be emphasised, e.g. by using consistent ordering and grouping terms

[Reply to R1.9] In the manuscript, the grouping in table one was [Classification error, Main channel Roughness, Vegetation Height, Vegetation density, non-vegetation roughness], while the subsections in 2.3 were [main channel roughness, floodplain roughness, classification error matrix].

We now changed table 1 such that the order and grouping are identical to the subsections.

[R1.10] p7 L14 "general extreme values distribution" Should be "Generalized Extreme Value distribution"?

[Reply to R1.10] That was indeed intended. We changed 'general extreme values distribution' to 'Generalized Extreme Value (GEV, Weibull variant) distribution'

[R1.11] p7 L20 "a 95% confidence limits at 0.31 m and 1.0 m." drop the "a``

[Reply to R1.11] We dropped the "a", such that the sentence now reads:

“(..) highly asymmetrical distribution with a mean of approximately 0.58 m and 95% confidence limits at 0.31 m and 1.0 m.”

[R1.12] p14 Table 2 Should "coefficient of variation" be average relative uncertainty E(Ur90)?

Otherwise, this raises the question of whether relative uncertainty was indeed needed?

[Reply to R1.12] We corrected the caption to “Expected relative uncertainty along the river (average  $\pm$  the standard deviation)”

[R1.13] Figure 7 y axis label should read Ur90? (9 is missing). Figure 9 also seems to have some text rendering issues

[Reply to R1.13] Unfortunately, the labels on the y-axis in both figures render correct in our PDF readers (showing Ur90, indeed); so we are not sure how to fix this issue. We will take this up if the manuscript is approved for final typesetting.

[R1.14] p18 L17 "greatly the unexplained variance" Word missing (increased?)

[Reply R1.14]

We wrote

[P18, L16-L17]:

“They observed that the removal of existing stochastic elements (e.g., a vegetation polygon) in exchange for new ones greatly the unexplained variance (...) “

We changed this to:

[P18, L17-L18 in modified manuscript]:

“They observed that the removal of existing stochastic elements (e.g., a vegetation polygon) in exchange for new ones greatly increased the unexplained variance (...) “

Reply to referee 2

We thank the reviewer for his time spend reviewing our manuscript and his suggestions for improvement. Below we have replied to the various comments made by the reviewer.



[R2.1] Overall the paper is well structured, well written and quite clear. However, I have a number of reservations regarding the manuscript. Several of these overlap with the points raised by Joseph in his comments, so I won't discuss those here in any further detail.

[Reply to R2.1] We thank the reviewer for his kind remarks. We acknowledge that some comments overlap with the first review and will refer to our reply to the first review where applicable.

[R2.2] My main reservation is regarding the contribution that the paper is trying to make. In the abstract the authors in the second sentence mention the role of model-based analyses in supporting decision making, while ending the abstract with claims about model-based decision making. In the middle of the abstract, the authors mention a new metric 'relative uncertainty'. Next, they argue that using this new metric they can provide insight into the uncertainty about the effects of a variety of flood risk reduction intervention.

I have two reservations here. First, regarding the claims related to decision making. In my view, these claims are not well developed in the paper. Outside of the introduction and conclusion, the term itself only appears once. Moreover, if I look at the presented results and their visualization, I doubt these would be used by a decision maker or even someone directly advising a decision maker. Rather, the presented analyses are useful for people working on the design of flood risks reduction strategies, while some of the results might indirectly be used in decision making. As such, I would suggest removing the term from the title as well as tone done any other claims made in the abstract and introduction. Alternatively, the authors would need to expand their discussion in the main text on how providing insights into the range of uncertainty about expected effects of measures can assist decision making on flood risk. This, however, requires discussing notions such as robustness (see e.g., McPhail et al., 2018) and flexibility, as well as discussions on well characterized uncertainty (i.e., you have a meaningful pdf) and Knightian or deep uncertainty (i.e., for whatever reason you don't have a meaningful and uncontested pdf).

[Reply to R2.2] We acknowledge that our use of the term "decision making" was intended to be narrower, and specific to decisions regarding the design of interventions. As the reviewer might well know, the established procedures for the design and approval of river interventions in The Netherlands (in Dutch, "Rivierkundig beoordelingskader") is highly model-based. It was inspired by this context that we intended to frame our work. We think it is to this the reviewer refers when he says "Rather, the presented analyses are useful for people working on the design of flood risks reduction strategies, while some of the results might indirectly be used in decision making."

So rather than being used directly in decision making, these models, and the extension offered by our work, would assist in decision making.

To avoid a too liberal use of the term 'decision making' in the context of the literature suggested by the reviewer, we made the following modifications.

The title was:

“Uncertainty quantification of flood mitigation predictions and implications for decision making.”

We changed it to:

[P1 title in new manuscript]

“Uncertainty quantification of flood mitigation predictions and implications for interventions.”

In the abstract we wrote:

“Finally, we show that for a defined standard of acceptable uncertainty, interventions need to be over-designed to meet this standard, and by how much. In general, we conclude that the uncertainty of model predictions is not large enough to invalidate model-based decision making, nor small enough to neglect altogether. Instead, uncertainty information can be used to improve intervention design and enrich the decision making process.”

We changed this to:

[P1 L14-17 in modified manuscript]

“Finally, we show that for a defined standard of acceptable uncertainty, interventions need to be over-designed to meet this standard. 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.”

[R2.3] My second reservation is with the new metric itself. As also indicated by Joseph, this metric is closely related to the coefficient of variation. Moreover, the more theoretical discussion of this new metric is confined to one short paragraph around equation 5. Too play devils advocate: what is the merit of publishing a paper whose only contribution is a single equation closely related too an already established metric? The case study mainly serves to establish the value of this metric, while the models are taken from earlier work. If you insist on having the metric as a key contribution, than a comparative perspective would be more appropriate. So, what other metrics already exist that could serve a similar function? Classic robustness metrics as well as the coefficient of variation would be logical candidates. How is this metric different from these, and what are the relative merits of the new metric relative to the others?

[Reply to R2.3] We agree that the key contribution is not the relative uncertainty. In response to remark [R1.2] we have removed the focus on this metric.

[R2.4] Basically, both reservations have to do with how the authors currently position their work. I am having profound reservations regarding the first two paragraphs on page 3.

First, the authors cite Pinter on fuzzy math, suggesting that not providing explicit uncertainty quantification leaves decision makers free to interpret the uncertainty in any way they like. My argument would be that the converse also often happens. As for example elaborated nicely in Pilkey and Pilkey-Jarvis (2007), a lot of quantitative analysis for supporting decision making becomes useless arithmetic that is used strategically. In a different line of literature, researchers working on post normal science often claim that many models have not even one significant digit. That is, models give only a false sense of precisions. A third line of research, as exemplified by Sarewitz (2004) shows how more research and increasing efforts to quantify uncertainty can often make environmental controversies worse. In short, the relationship between models, model results, uncertainty about models results and decision making processes is quite a bit more complicated than the authors seem to think. It would benefit the manuscript if the claimed contribution is better positioned relative also to these strands of literature.

[Reply to R2.4] As motivated in [Reply to R2.2], we did not aim to focus on the decision-making process. The authors are aware of the post-normal science discussion, and concede that quantitative information can be used strategically, especially if there is too much focus on technical uncertainties to create superfluous knowledge which is not relevant to the policy process (see e.g. Warmink et al., 2017). We would not argue for the results of such uncertainty quantification (or model results in general) to take the centre stage in decision making, but merely to provide valuable information to assist in that process. We appreciate the suggestions of the reviewer to embed our work in the suggested literature, and have made the following modifications considering these suggestions and the intended scope of our work (cf. [Reply to R2.2]).

*Warmink et al. 2017: (P4592) in Water Resour Manage (2017) 31:4587–4600. DOI 10.1007/s11269-017-1767-6.*

The following paragraph is removed:

[p3 L3-10]

“To this date, there is little evidence in literature of explicit uncertainty quantification for model predictions of flood mitigation measures, or of other human intervention in river systems. Lack of explicit uncertainty quantification leaves room for what Pinter (2005) calls ‘‘fuzzy math’’ --- free interpretation of model uncertainty --- by decision makers. In his example, building permits for floodplains were easily granted on the premise that individual effects are negligible compared to model uncertainty. Since the cumulative effect of all those permits is not negligible (Pinter et al., 2008), this interpretation likely overestimates the effect of uncertainty. An opposite example is given by Mosselman (2018), who reports that large, quantified uncertainties in absolute flood levels are sometimes ignored for intervention effects under the assumption that uncertainties ‘cancel out’.

This assumption is incorrect Berends et al. (2018), but further study is needed to determine how large the uncertainty actually is.”

We added the following paragraph:

[P2 L9-21 in modified manuscript]

“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 effect of the planned intervention on hydraulics, such as water 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 ‘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 errors cancel out when subtracting the intervention case from the reference case. However, it was recently demonstrated in an idealized 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.”

Furthermore, the rest of the introduction was revised to accommodate this new paragraph.

In the discussion, we wrote:

[p20, L17-21]

“In literature, there are several examples of probabilistic and risk-based approaches in designing flood adaptation strategies (Lendering et al., 2018; Klijn et al., 2018). Model uncertainty is generally not explicitly quantified as part of such approaches, but could be a natural addition to such approaches. This is especially so when model uncertainty influences intervention design (Hall and Solomatine, 2008; Beven et al., 2018). The quantification and visualisation of uncertainties here may be a step towards this goal.”

We changed this to:

[p20, L27 – p21, L4 in modified manuscript]

“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 other uncertainties exist that cannot be

reduced by more data or simply cannot be quantified. Examples of ‘social uncertainty’ include fundamental disagreement between experts, lack of trust in scientific results, ambiguity and diverging narratives based on the same empirical basis (Sarewitz 2004, Bruchnach, 2011, Warmink 2017). Coping with such uncertainties may require adaptive river management (Pahl-Wostl, 2009) and exploration of multiple alternate futures (The third paradigm of Maier et al. (2016)). Given these considerations, there are several probabilistic frameworks of flood risk adaptation strategies to which explicit quantification of model uncertainty is a natural and required extension. 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 models, is not sufficient for planning interventions in river systems.”

#### Minor remarks

[R2.5] Might it not be more convenient to show the interventions (2.2.1-2.2.6) in a table?

[Reply to R2.5] Our choice for subsections instead of a different format (e.g. a ‘box’) was mainly motivated by the limitation of the NHESS template rules. If approved for final typesetting, we will confer with the journal whether a different format is possible.

[R2.6] What is the runtime of the detailed model for a single run?

[Reply to R2.6] The runtime of a single model run was approximately 2.5 hours. We added this information to section 2.2.

[R2.7] Page 18, line 17, “for new ones greatly the unexplained” I guess some words are missing here

[Reply to R2.7] The missing word is “increased”. For our answer we refer to our reply to [R1.14]

[R2.8] Page 17, line 4, partly overlapping with Joseph’s comment, but what justifies the linear interpolation?

[Reply to R2.8] We refer to our answer to [R1.7]



# Uncertainty quantification of flood mitigation predictions and implications for ~~decision making~~.interventions

Koen D. Berends<sup>1,2</sup>, Menno W. Straatsma<sup>3</sup>, Jord J. Warmink<sup>1</sup>, and Suzanne J.M.H. Hulscher<sup>1</sup>

<sup>1</sup>Department of Marine and Fluvial Systems, Twente Water Centre, University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands

<sup>2</sup>Department of River Dynamics and Inlands Shipping, Deltares, Boussinesqweg 1, 2629 HV Delft, The Netherlands

<sup>3</sup>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 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 establish relationships to aid in decision making. We ~~show that the uncertainty of an intervention can be adequately described by the newly introduced ‘relative uncertainty’ metric~~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 a 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 that for a defined standard of acceptable uncertainty, interventions need to be over-designed to meet this standard, ~~and by how much~~. In general, we conclude that the uncertainty of model predictions is not large enough to invalidate model-based ~~decision making~~intervention design, nor small enough to neglect altogether. Instead, uncertainty information ~~can be used to improve intervention design and enrich the decision making process~~is valuable in the selection of alternative interventions.

## 20 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 data, not all model input or model parameters can be reliably measured or represented. For example, vegetation density and 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).

Proper understanding and communication of ~~this~~ uncertainty is important both for scientists and decision makers (Pappenberger and Beven, 2006; Uusitalo et al., 2015).

~~Predicting the effect of flood risk mitigation measures with models is similarly affected by uncertainty~~ 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 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 ‘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 errors cancel out when subtracting the intervention case from the reference case. However, there are two compounding issues particular to intervention design studies. First, the detailed models used in effect studies have a high number of parameters that need to be determined. There 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 intervention design studies. First, there is a clear practical need for 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). ~~In some countries, this 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 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,



such as like the construction of artificial secondary channels (van Buuren and Warner, 2014)(van Denderen et al., 2019), requires detailed models that take into account local geometry, vegetation and other terrain features.

5 A second compounding 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 parameters, which complicates the use of (data-based) surrogate model approaches (Razavi et al., 2012). The second issue is that effect studies *ex-ante* rely (by definition) on simulations of the analysis relies on at least one unmeasured environmental system; namely, the river system altered by the proposed intervention. This complicates using measurements to estimate model accuracy, because Therefore, there is no way to verify that accuracy for the proposed future state of the river. A further complication is that the hydraulic conditions, for which Verification of the accuracy of the current, unaltered state is likewise not necessarily available. For example, flood mitigation measures are  
10 designed, tend to have 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). Therefore, it is likely that measurements of the current, unaltered state are likewise unavailable. For these reasons, 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.

15 Model output uncertainty can alternatively be expressed in terms of the In present paper we address the lack of studies into the model uncertainty of model input and parameters. This approach, also known as forward uncertainty analysis (Beven et al., 2018), is stochastic: both uncertain model input and parameters, as well as model output are expressed as probability density functions. For most practical applications, neither the deterministic problem (i.e. solving the flow equations) nor the stochastic problem (i.e. resolving the model output probability density functions) are analytically tractable. Monte Carlo simulation (MCS) is a well-known method to solve the stochastic problem numerically (Stefanou, 2009). The disadvantage of MCS  
20 is that it comes with a significant, often prohibitive, computational cost. An alternative, computationally less expensive, approach was proposed by Berends et al. (2018). This approach reduces the number of post-intervention model runs by using the pre-intervention model as a surrogate in a statistical Bayesian framework, which becomes more efficient with an increasing number of interventions.

To this date, there is little evidence in literature of explicit uncertainty quantification for model predictions of flood mitigation  
25 measures, or of other human intervention in river systems. Lack of explicit uncertainty quantification leaves room for what Pinter (2005) calls “fuzzy math” — free interpretation of model uncertainty — by decision makers. In his example, building permits for floodplains were easily granted on the premise that individual effects are negligible compared to model uncertainty. Since the cumulative effect of all those permits is not negligible (Pinter et al., 2008), this interpretation likely overestimates the effect of uncertainty. An opposite example is given by Mosselman (2018), who reports that large, quantified uncertainties in  
30 absolute flood levels are sometimes ignored for intervention effects under the assumption that uncertainties ‘cancel out’. This assumption is incorrect (Berends et al., 2018), but further study is needed to determine how large the uncertainty actually is.

Given the problem that effect studies use complicated models in unmeasured situations, it is likely that there is significant uncertainty involved with model predictions. Yet, lacking uncertainty quantification leads to a free interpretation of the importance of this uncertainty. To approach this problem we present a quantitative, real-world study on the effect of model uncertainty on  
35 designing flood mitigation measures. Here we restrict the definition of effect to the initial reduction of flood water levels by a

~~proposed intervention, without considering secondary effects such as subsequent morphological or ecological system response. model predictions used in impact design.~~

Our objective is twofold. ~~First, we~~ 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. ~~The,~~ using a detailed hydraulic model. To limit the computational  
5 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 establish a 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.

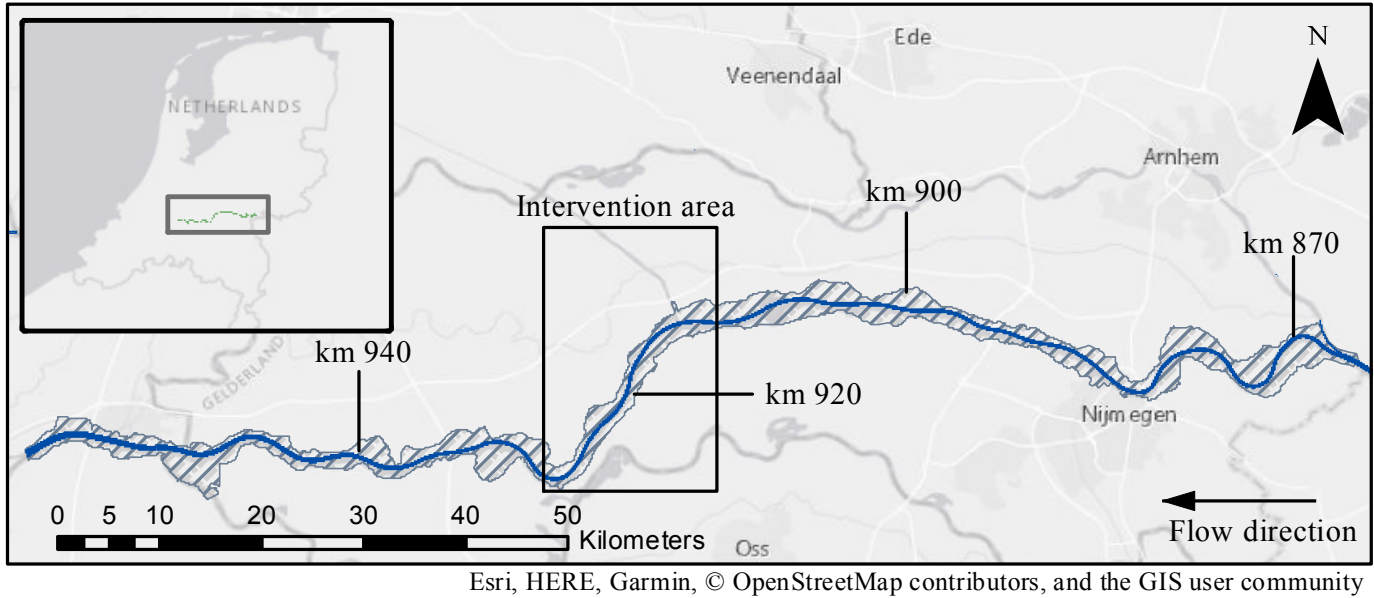
Our analysis proceeds in three steps. First, we set up the numerical model for the Dutch River Waal and select the uncertainty  
10 sources in section 2. Here, we also introduce the relative uncertainty measure. We then present the results of the uncertainty of the modelled water levels and reduction 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  
15 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  
20 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 inland navigation dates back to at least the 18<sup>th</sup> century (e.g., see Velsen, 1749). In the 19<sup>th</sup> century, the river was normalised in large governmental projects, obtaining its current narrow channel, groynes (also termed *spur dams* or *wing dams*) and relatively  
25 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 20<sup>th</sup> 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; 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  
30 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 ‘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



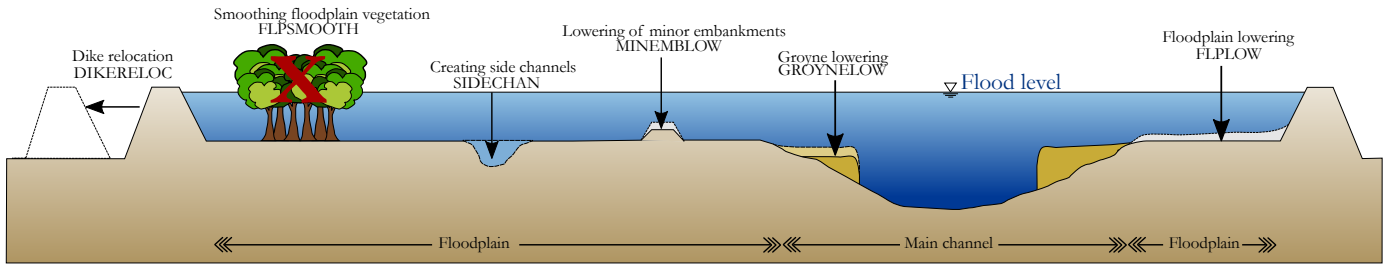
**Figure 1.** The River Waal, situated entirely in The Netherlands. The study area is a stretch of about 15 km between km 913 and km 928. The hatched area shows the total extent of the model domain, which overlaps with the river and its floodplains bounded by the dikes.

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. 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 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 resolution varies from 40m by 15m in the main channel and a gradually increasing resolution in the floodplains to a maximum of about 120m by 120m. In all cases, we simulate a steady upstream discharge of  $10.165 \text{ m}^3\text{s}^{-1}$  and a constant downstream water level of  $3.98 \text{ m} + \text{NAP}$ , 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 median values. The initial time step was set at 30s.



**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).

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 different interventions, detailed below. Each intervention was implemented in a low-intensity and a high-intensity variant. 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.

### 2.2.1 Groyne lowering (GROYNELOW)

Groynes (also known as *wing dams* or *spur dams*) 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 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  
5 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  
10 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.

#### **15 2.2.6 Side channels (SIDECHAN)**

Finally, the construction of secondary (side) channels within the existing corridor both increases the available volume, as  
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  
20 of 363 days per year.

### **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)  
25 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. Following Warmink et al. (2013b), we take into account uncertainty in the main channel and classification errors in the land-use maps. Additionally, we take into account uncertainty in vegetation parameters. This leads to a total of 39 stochastic variables (see Table 1).

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

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

### 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 ~~general extreme values distribution~~ 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.

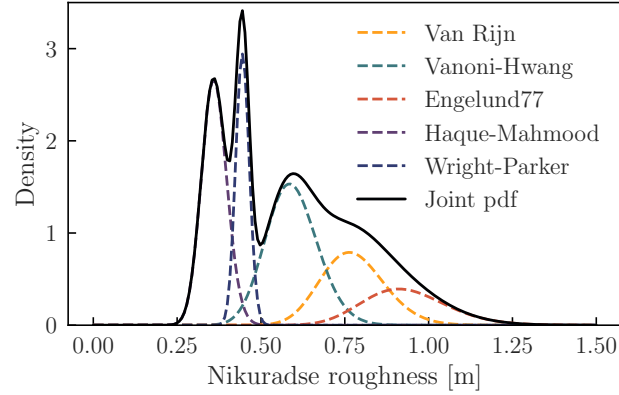
In the following we will use the joint cumulative density function of the five roughness models to sample representative roughness height  $k_n$ . The joint cumulative function is a highly asymmetrical distribution with a mean of approximately 0.58 m and a 95% confidence limits at 0.31 m and 1.0 m. The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

$$C = 18_{10} \log \frac{12h}{k_n} \quad (1)$$

with Chézy coefficient  $C$  [ $\text{m}^{1/2}\text{s}^{-1}$ ] 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

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 designed either one or a combination of distinct vegetation classes.



**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.

For vegetation, the generation of roughness by vegetation is complex, resulting from obstruction by stems and leafs and 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  $m_v$  [ $\text{m}^{-2}$ ] 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., 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 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

$$C = h^{1/6} n^{-1} \quad (2)$$

with Manning coefficient  $n$  [ $\text{sm}^{-1/3}$ ]. 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 trachytopes 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  $\Delta H$ ). Due to the uncertainties in model input discussed in section 2.3,  $X$ ,  $Y$  and  $\Delta H$  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 not-analytically tractableanalytically intractable, but can be solved numerically by Monte Carlo simulation (MCS; Metropolis, 1987; Stefanou, 2009). In the numerical approach, covariances are not explicitly computed.

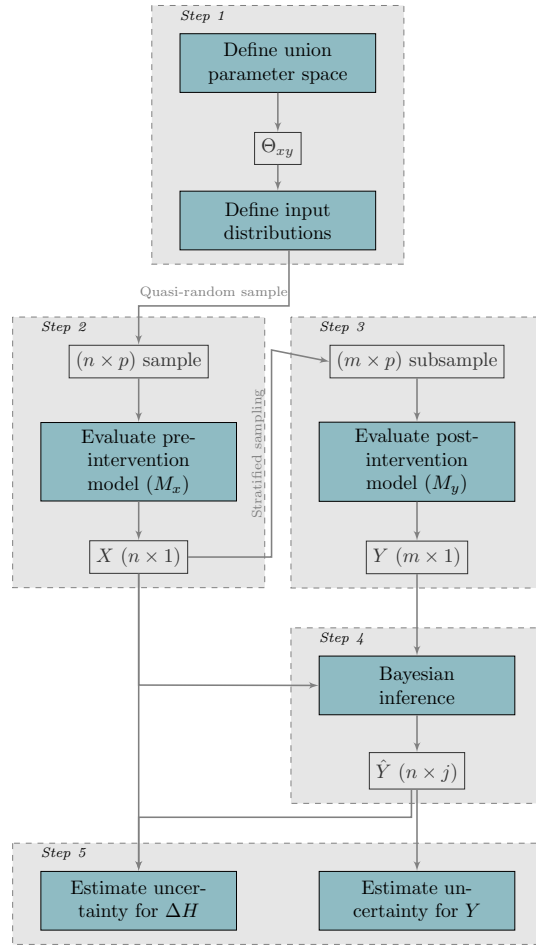
However, MCS is not practically feasible; although always practically feasible or desirable. Although MCS has been applied for estimation of flood level uncertainty before (Warmink et al., 2013b), the added task of performing a separate MCS for each intervention puts a severe strain on computational 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 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):

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

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





**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;  $\Delta H$ : intervention effect. Figure adapted from (Berends et al., 2018).

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.

- 5 The main advantage of this approach is that 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 trachytopes 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 trachytopes. 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  $\Delta H$ .

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  $\Delta H$  as the model uncertainty confidence interval or MCI. For example, the MCI of  $\Delta H$  is defined as:

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

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).

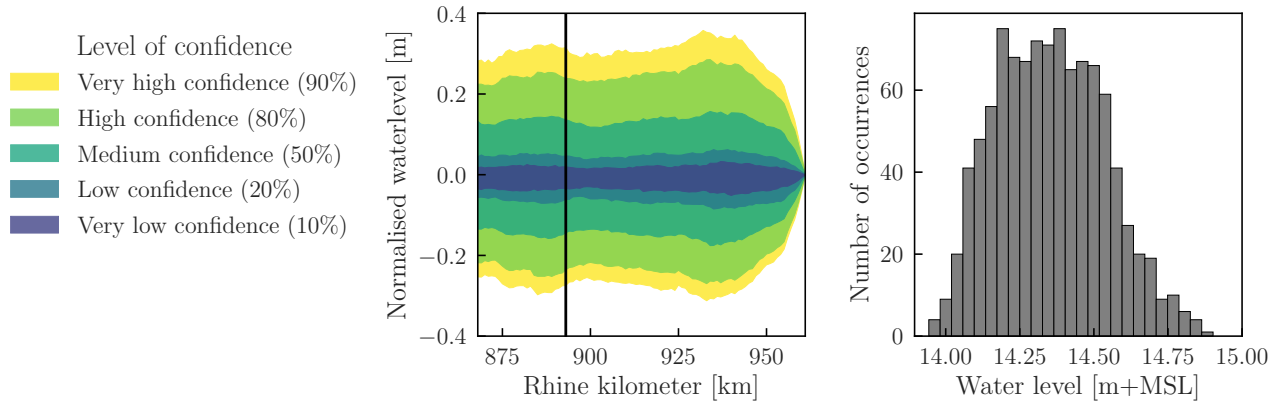
To compare the uncertainty of interventions we calculate the *relative uncertainty*, i.e. the uncertainty relative to the expected effect. For this, we use an adapted version of the coefficient of variation, defined as:

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

where  $MCI_{90,\Delta H}$  is the 90% confidence interval for  $\Delta H$  and  $E(\Delta H)$  the average of  $\Delta H$ . The  $U_{r;90}$  is expressed in percentages. Therefore, a  $U_{r;90}$  of 100% means that the 90% MCI is as large as the average effect. Since both  $MCI_{90,\Delta H}$  and  $E(\Delta H)$  are stochastic,  $U_{r;90}$  is stochastic as well.

### 3 Results

#### 5 3.1 Pre-intervention uncertainty

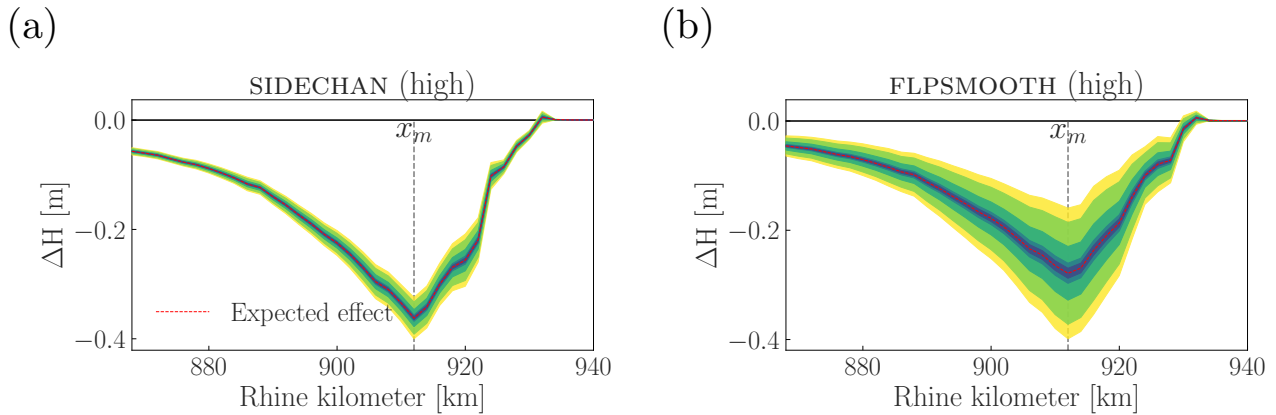


**Figure 5.** Middle panel: Confidence levels of design water levels, normalised by the average water levels at each location. The water flow is from left (upstream) to right (downstream). The black line indicates km 893. Right panel: the histogram of the (not normalised) water levels at km 893.

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 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 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  $x_m$ .

5 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 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

10

15 accounted for in this study.

The novelty of Fig. 6 these 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 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.

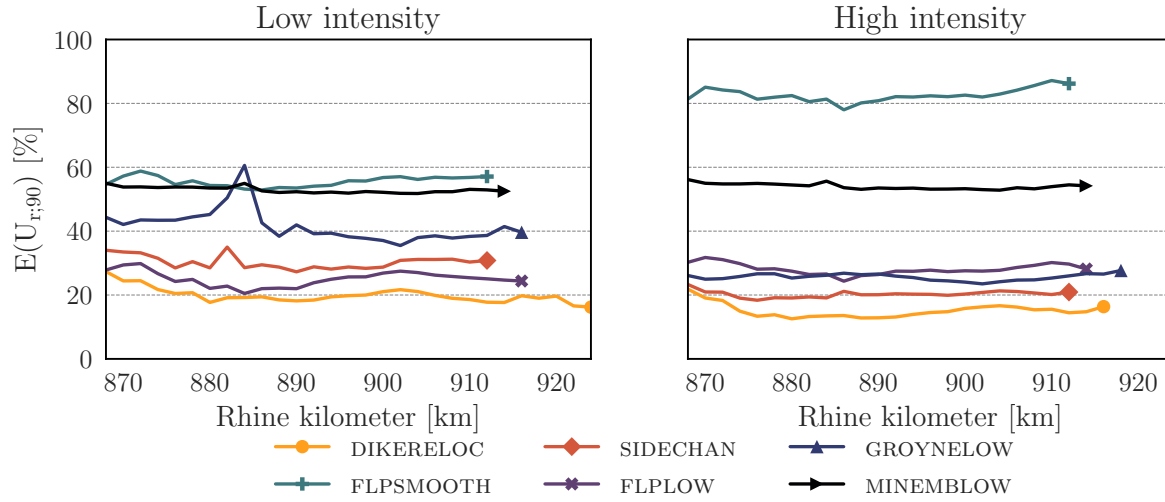
Next, we compute the relative uncertainty ( $U_{r;90}$ ) for each intervention and intensity. ~~Since We first marginalise the uncertainty in  $U_{r;90}$  is stochastic, we first reduce this due to coral estimation~~ by calculating the expected value ( $E(U_{r;90})$ ) for each point along the river (Fig. 7). ~~Results 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  $U_{r;90}$  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 effects. In general, the constant values for  ~~$U_{r;90}$~~   $E(U_{r;90})$  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 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 ( $E(\Delta H)_{x_m}$ ) are accompanied by small uncertainties ( $E(MCI_{90})_{x_m}$ ). 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.

### 3.3 Inter-comparison between interventions

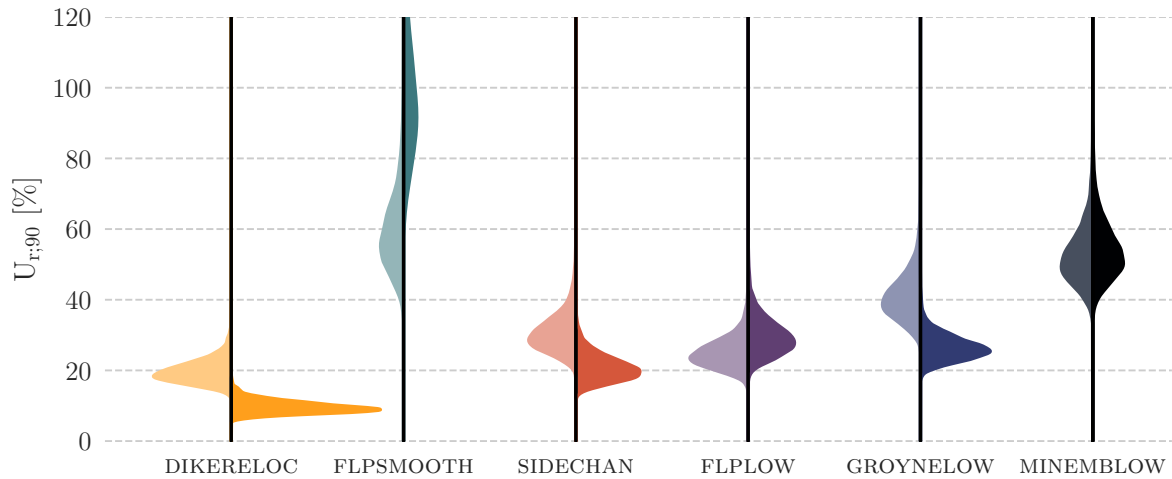
A straightforward way to compare differences between the effect of two or more interventions is to look at the backwater figures. In Fig. 6 we see two diagrams of interventions that result in flood level decrease within the same order of magnitude, but with very different uncertainty. The uncertainty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



**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;  $E(\Delta H)_{x_m}$ = expected effect at  $x_m$ ;  $E(MCI_{90})_{x_m}$ =expected 90% uncertainty at  $x_m$ ;  $\widehat{U}_{r;90} E(U_{r;90})$  = **Coefficient of variation, averaged over Expected relative uncertainty along the river length (average  $\pm$  the standard deviation)**. The rows are sorted from highest expected effect (top) to lowest expected effect (bottom).

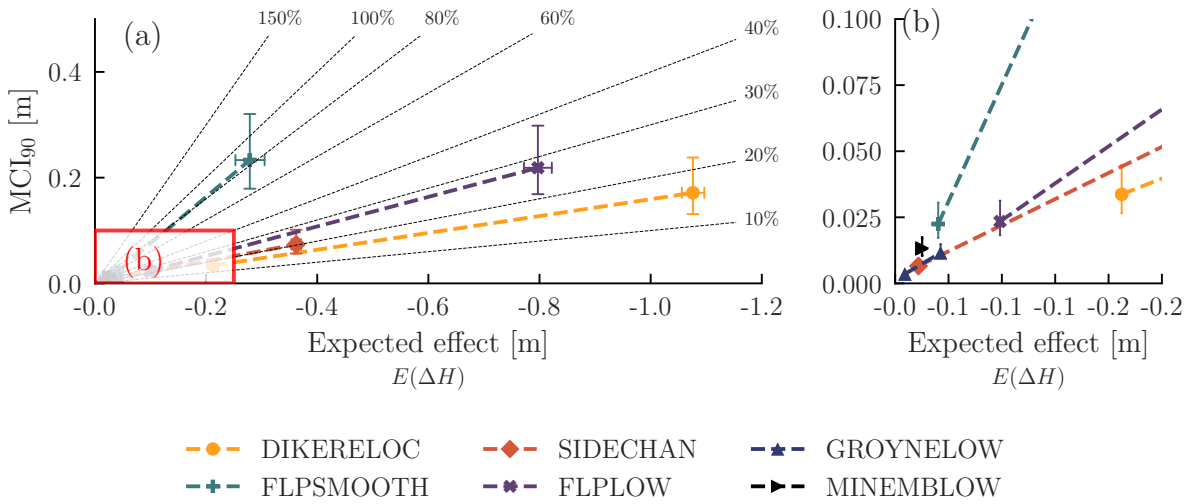
Intervention		$x_m$	$E(\Delta H)_{x_m}$	$E(MCI_{90})_{x_m}$	$\widehat{U}_{r;90} E(U_{r;90})$
DIKERELOC	H	916	-1.08	0.176	15 $\pm$ 2.2%
	L	924	-0.21	0.034	20 $\pm$ 2.3%
FLPLOW	H	914	-0.80	0.224	28 $\pm$ 1.7%
	L	916	-0.10	0.024	25 $\pm$ 2.3%
SIDECHAN	H	912	-0.36	0.076	18 $\pm$ 2.0%
	L	912	-0.02	0.007	20 $\pm$ 2.1%
FLPSMOOTH	H	912	-0.28	0.239	82 $\pm$ 1.9%
	L	912	-0.04	0.023	55 $\pm$ 1.6%
GROYNLOW	H	918	-0.04	0.012	23 $\pm$ 2.6%
	L	916	-0.01	0.004	35 $\pm$ 4.2%
MINEMBLOW	H	914	-0.03	0.014	53 $\pm$ 5.4%
	L	914	-0.03	0.013	52 $\pm$ 5.3%



**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.

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 clearly more uncertain than the other four interventions and (ii) that 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. By linearly interpolating between the ~~two expected values of the~~ low and high intensities, we ~~can estimate~~ 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, FLFLOW, SIDECHAN and DIKERELOC), within the bounds marked by the low- and high intensities. The confidence 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 (FLFLOW) and 21 cm (FLPSMOOTH). These intervals may or may not be acceptable.



**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.

Instead of looking at the expected effect, the confidence intervals can be marginalised by designing for a given exceedance probability of the effect (Figure 10). The desired effect is related to the expected effect through the exceedance probability. ~~For example, figure 10 (a) shows various exceedance probabilities for FLPSMOOTH.~~ To illustrate 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-



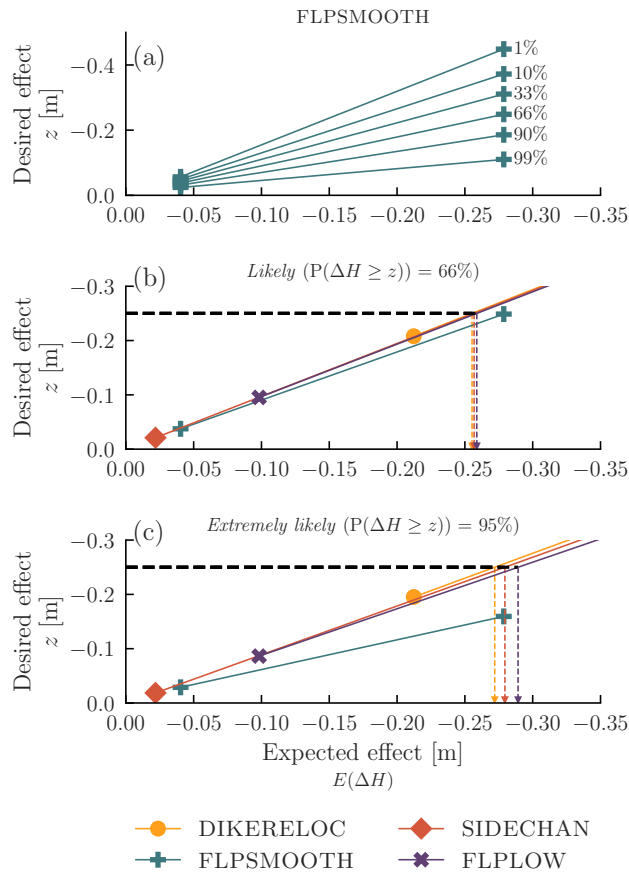
designed depends on the uncertainty of that measure and the exceedance probability. Higher exceedance probabilities such as a 95% likelihood ("Extremely likely") can also preclude some flood mitigation measures (Figure 10 (c)). With this level of confidence, FLPSMOOTH can no longer be considered: none 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

The 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 ( $\epsilon$  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, although the specific results for these interventions is likely not.

~~In studying the effect of uncertainty on intervention design (figures 9 and 10) we linearly interpolated between the high intensity and low intensity interventions. The assumption of linearity is reasonable given the approximately constant  $U_{r,90}$  between the two intensities. However, in practice this should be only seen as a first approximation. Multiple design iteration would be necessary to arrive at a design that meets the requirements.~~

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, ~~reduces the required computer resources~~. 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 this



**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.

the estimation uncertainty by increasing the number of model runs with the post-intervention models (Berends et al., 2018), if (computational) resources allow. ~~In this study, we show that the relative uncertainty  $U_{T,90}$  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).~~

5 ~~Our study focused on providing uncertainty estimations for many different interventions, not to provide an optimal intervention 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.~~

15 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.

25 ~~In literature, there are several examples of probabilistic and risk-based approaches in designing flood adaptation strategies (Lendering et al., 2018; Klijn et al., 2018). Model uncertainty is generally not explicitly quantified as part of such approaches, but could be a natural addition to such approaches. This is especially so when model uncertainty influences intervention design (Hall and Solomatine, 2008; Beven et al., 2018). The quantification and visualisation of uncertainties here may be a step towards this goal.~~

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 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). Given these considerations, there are several probabilistic frameworks 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.

## 5 Conclusions

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

ing 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 benefitted 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			
<b>Empirical distribution</b>					
612 - 637	Alluvial bed				
<b>Uniform distribution</b>					
n/a	Classification map				
<b>Triangular distributions</b>					
		<i>min</i>	<i>mean</i>	<i>max</i>	
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	
<b>Lognormal distributions</b>					
		$\mu_{h_v}$	$\sigma_{h_v}$	$\mu_{n_v}$	$\sigma_{n_v}$
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	<i>Deterministic</i>	<i>Deterministic</i>	-4.68	0.67
1242	Softwood production forest	<i>Deterministic</i>	<i>Deterministic</i>	-4.72	0.66
1243*	Pine forest	<i>Deterministic</i>	<i>Deterministic</i>	-4.18	0.54
1244	Hardwood forest	<i>Deterministic</i>	<i>Deterministic</i>	-3.45	0.77
1245	Softwood forest	<i>Deterministic</i>	<i>Deterministic</i>	-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.
- 5 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.
- 10 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.
- 15 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.
- 20 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](https://doi.org/10.1016/S0079-1946(96)00005-5), [http://dx.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.
- Hall, J. and Solomatine, D.: A framework for uncertainty analysis in flood risk management decisions, *International Journal of River Basin Management*, 6, 85–98, <https://doi.org/10.1080/15715124.2008.9635339>, 2008.
- 25 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 velocity in case of submerged rigid cylindrical vegetation, *Water Resources Research*, 43, <https://doi.org/10.1029/2006WR005625>, 2007.
- 30 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 Management*, 11, 287–299, <https://doi.org/10.1080/15715124.2013.811418>, 2013.
- 35

- 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.
- 10 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](https://doi.org/10.1007/978-981-10-7218-5_53), 2018.
- 20 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.
- 25 Oreskes, N., Shrader-Frechette, 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.
- 30 Pinter, N.: ENVIRONMENT: One Step Forward, Two Steps Back on U.S. Floodplains, *Science*, 308, 207–208, <https://doi.org/10.1126/science.1108411>, 2005.
- 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.
- 35 Pinter, N., Jemberie, A. A., Remo, J. W. F., Heine, R. A., and Ickes, B. S.: Flood trends and river engineering on the Mississippi River system, *Geophysical Research Letters*, 35, <https://doi.org/10.1029/2008gl035987>, 2008.
- Razavi, S., Tolson, B. A., and Burn, D. H.: Review of surrogate modeling in water resources, *Water Resources Research*, 48, W07401, <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.
- 5 Shields, F. D., Coulton, K. G., and Nepf, H.: Representation of Vegetation in Two-Dimensional Hydrodynamic Models, *Journal of Hydraulic Engineering*, 143, 02517002, [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0001320](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.
- 10 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.: Hydraulic 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\(500\)](https://doi.org/10.1061/(ASCE)0733-9429(2002)128:5(500)), 2002.
- 15 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 errors 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.
- 20 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., Bloeker, 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, e1602762, <https://doi.org/10.1126/sciadv.1602762>, 2017.
- 25 Uusitalo, L., Lehtikoinen, 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 Buuren, A. and Warner, J.: From Bypass to Bathtub: Backfiring Policy Labels in Dutch Water Governance, *Environment and Planning C: Government and Policy*, 32, 1000–1016, <https://doi.org/10.1068/c12212>, 2014.
- 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.
- 30 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 Uijttewaai, 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.
- 35 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.
- 10 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 Management*, 6, 302–318, <https://doi.org/10.1111/jfr3.12014>, 2013b.
- 15 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.
- 20 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.