Measuring Confidence in NaFRA Outputs

Stage 3 – Route Mapping

Report – FCPIF00151B00/R3

We are The Environment Agency. It's our job to look after your environment and make it **a better place** – for you, and for future generations.

Your environment is the air you breathe, the water you drink and the ground you walk on. Working with business, Government and society as a whole, we are making your environment cleaner and healthier.

The Environment Agency. Out there, making your environment a better place.

Published by:

Environment Agency
Horizon House, Deanery Road
Bristol BS1 5AH
Tel: 0117 934 4000
Email: enquiries@environment-agency.gov.uk
www.environment-agency.gov.uk

Further copies of this report are available from our publications catalogue: http://publications.environment-agency.gov.uk or our National Customer Contact Centre: T: 08708 506506

E: enquiries@environment-agency.gov.uk.

Executive summary

This report is provided as part of the Measuring Confidence in NaFRA Outputs Project and is tasked with setting out a route map for the future development of methods to support the assessment of the confidence and uncertainty within the NaFRA. It is assumed that the reader is familiar within RASP, MDSF2 and NaFRA.

This report does not attempt to present the Business Case for any of the recommended activities. Although the Route Map presented here should provide a key contribution to a Business Case(s), the Business Case itself is outside of scope.

Acknowledgements

The comments and contributions from Ian Meadowcroft, Mike Steel, Dave Hornby, Selena Peters, Jeff Baldwin and Ollie Snowdon of the Environment Agency are all gratefully acknowledged. Inputs and advice from Rob Lamb (JBA) and review comments from Neil Hunter and Andrew Gubbin, both of JBA, are also acknowledged.

Contents

1 Introduction

1.1 Background

The Confidence Indexing (CI) method (as implemented 2012/13) provides a significant step forwards in understanding and communicating the confidence within the NaFRA. The approach is based on a structured process of scoring and weighting the quality of both the input data and model performance. As such it necessarily relies upon many significant assumptions to derive a quantified estimate of the chance that the flood likelihood band associated with a particular Impact Cell has been correctly identified.

Prior to the development of the CI a number of previous studies have attempted to provide a fully quantified assessment of uncertainty using a probabilistic framework alone. None, however, were perceived to provide credible results. This does not imply that a probabilistic approach has no place, but rather that its limitations should be recognised and the results used alongside non-probabilistic approaches that rely more on structured judgment and sensitivity analysis.

The Route map presented here therefore recognises that achieving a credible estimate of confidence and uncertainty is a significant challenge and unlikely to be achieved through a single approach.

1.2 Report aim and guiding principles

The aim of this route mapping activity is to set out the key activities required to achieve a more robust, transparent and defensible assessment of the uncertainty within NaFRA and the key drivers of that uncertainty. To help meet this overall aim the Route Map will need to support the Environment Agency in:

- Developing the evidence to support/update the Confidence Indexing approach.
- Incorporating known uncertainties (associated with the input data and models used to support NaFRA) in a defensible quantified analysis.
- Assessing the potential for gross errors (arising from the lack of representativeness or incompleteness of the data and models used to support NaFRA).
- Implementing the various strands of uncertainty efficiently, reusing information and avoiding duplication of analysis streams or processes.
- Appropriately communicating a more comprehensive understanding of uncertainty and confidence.

For it to be credible and promote buy-in the Route Map will also need to show how it:

- Builds upon the lessons from the development and implementation of the Confidence Indexing approach.
- Uses ground truthing and observations to build an evidence-supported case of the credibility of the results.

- Is honest about what is included and what is excluded from the analysis and takes this into account when expressing the uncertainty in both single and collective statements (i.e. uncertainty at any given location or aggregated output).

- Uses information relating to the typology of the Flood Areas to highlight the completeness (or otherwise) of the modelled representation of the real system.

- Uses the results of the NaFRA benchmarking activities to support the quantification of uncertainty in the output results introduced through errors in the input data, model components and model structure.

2 Lessons from previous studies

The Confidence Indexing approach has been largely accepted as a ‘success’ (based on ad hoc feedback to the authors). It has delivered a first national assessment of confidence and been widely accepted as a pragmatic approach given the timescales and resources available. More specifically perhaps, the key reasons for the success of the Confidence Indexing method were:

- The underlying business imperative to assess the confidence in the NaFRA banding.
- The conceptual simplicity, yet intuitive nature of the approach.
- The ease of integration of the approach with other sources of data (for example data quality flags).
- The associated provision within familiar software (MDSF2) and the provision of additional training.

This does not mean, however, that the CI approach will continue to be viewed as credible in its current form. The scientific Peer Review undertaken at the time acknowledged the constraints of time and, in this context, accepted the approach as a good step forwards.

Prior to the CI approach all of previous attempts to assess confidence/uncertainty in NaFRA/MDSF2 have largely failed to make a significant impact on the Environment Agency’s business. The reasons for this are numerous (as discussed in Table 2.1) but can be distilled into six main points:

- The lack of a clear business need and capacity to use the information.
- A lack of understanding of the approach, perceived to be a ‘black-box.
- An inability to contribute perceptions of confidence to the analysis process in a structured objective way.
- A lack of credibility in an approach that relies on a probabilistic assessment given the potential for gross model errors (arising from incomplete descriptions or systematic errors).
- A lack of understanding of how best to aggregate uncertainties spatially.

Past study	Summary of study and lessons
NaFRA 2009 Data Improvements, Uncertainty and Sensitivity Analysis	Case study analysis undertaken in approximately 200 fluvial Flood Areas in four catchments of different characteristics, making use of the Latin Hypercube: One-At-a-Time (LH-OAT) and RASP-SU (RASP-Structured Uncertainty) approaches. • The most important model inputs and their associated uncertainties were identified by a combination of expert judgement and desk studies. The uncertainties within these were individually specified according to how the inputs were obtained (e.g. from 'best of breed' approaches to approaches requiring 'heroic assumptions'). • The sensitivity analysis has clearly demonstrated that three datasets – water level, defence crest level and in-river available volume/volume capping – were the most important input variables in determining the uncertainty associated with the model outputs. • There were concerns regarding the selection of the 'most important inputs' and the definition of uncertainties associated with these (as these were based on a combination of expert judgement and desk truthing studies carried out as part of this project and as such are themselves subject to uncertainty). • A better understanding of the input uncertainties of greatest importance would help improve the credibility in the approach. • No consideration was given to uncertainties in model structure nor the NaFRA validation process and Area Output Review that are applied after modelling but before the model results are published. • Initially, results were assessed for a high (80%) confidence level. As the analysis unfolded, results suggested that an 80% confidence level in results was rarely achievable. (based on NaFRA 2009 Data Improvements, Uncertainty and Sensitivity Analysis - Uncertainty and Sensitivity Analysis, A Case Study Summary, Halcrow 2011)
Exploring the sensitivity of RASP HLM+ to variations in input data and model parameters	Pilot sites used to test the RASP HLM+ via perturbation of input data and model parameters. This study highlighted that: • Appropriate resources and technologies should be devoted to evaluating loads and understanding their uncertainty. • Risk is very sensitive to crest level, condition grade, property floor space / damage functions and the number of events in the loading curve. • Defence levels should be linked to loading conditions using a common datum • Results are very dependent on the specifics of the site being tested. • Sensitivity analysis on a limited set of pilot sites provides

Past study	Summary of study and lessons
	useful background, but with limited specific detail. It was unclear if the prioritisation of the data uncertainties derived was applicable to all Flood Areas and without a quantified assessment, determining the value of improving the data was difficult to determine. (based on Science Report SC050064/SR - Exploring the sensitivity of RASP HLM+ to variations in input data and model parameters, HR Wallingford 2006)
Understanding confidence in NaFRA 08	Variance based sensitivity analysis applied to fluvial pilots: • Uncertainty and sensitivity analysis is appropriate for some but not all types of uncertainty. It reflects accuracy of data and model parameters but not major errors in data or model, or missing data. • The method used assumes independence between the input variables and parameters, which may not be appropriate but could be easily introduced where appropriate. • The uncertainty in the overtopping volume is significant and varies greatly between defences. • The uncertainty in flood depth within an Impact Zone increases with return period and is sensitive to the uncertainty in the overtopping volume and quality of the ground data. • Data improvements should be targeted to priority areas - for example to reduce the uncertainty in high consequence areas. • The sensitivity analysis demonstrated how to identify individual data priorities for each FRMS. • The sensitivity analysis also highlighted the importance of the 1:10-1:50-year return period events in establishing a reliable estimate of depth v probability and risk. • A much closer link should be established between the uncertainty work and the gathering of specific supporting datasets to inform the assessment of the uncertainty in the input variables and parameters. (based on HR Wallingford Report EX 5953 – Understanding and communicating our confidence in the National Flood Risk Assessment 2008 – A trial study, HR Wallingford 2009)
LH-OAT application	An alternate method for Uncertainty Analysis and Sensitivity Analysis has been developed through the Environment Agency R&D project “SC090008, Work Package1: Validation of Probabilistic Flood Models” with similar issues of credibility as the RASP-SU. (see NaFRA 2009 Data Improvements, Uncertainty and Sensitivity Analysis - A Case Study Summary, Appendix A Summary Note I5, Halcrow 2011)
Validation of probabilistic flood	• The literature review of validation of probabilistic methods both within flood risk and elsewhere has provided a good summary of the current state of the art. Whilst there is

Past study	Summary of study and lessons
models SC090008/WP1 (2010)	much literature on uncertainty analysis, there is much less literature on other aspects of the validation of probabilistic methods. • The previous evaluation work on the RASP HLM+ engine provides much useful information, but has been focused on sensitivity and uncertainty analysis, rather than validation against independent data. • The development of a High Level Conceptual Framework (HLCF) must consider this requirement and place a particular emphasis on user focused communication of fitness for purpose aligned with a capacity building programme within the user community. • Computation times may be a serious issue for the analysis required to understand parameter uncertainty. Reducing the number of parameters subject to uncertainty analysis to only key variables, and neglecting the rest, if done carefully, may reduce computation times significantly. • Different types of model outputs may exhibit very different behaviours, as shown through the validation of depth-probability and risk outputs. • Probabilistic models are often used to represent rare events such as breaching, and these are very difficult to validate due to lack of observed data. The rarity of defence failure cases means that validating defence fragility curves may be extremely difficult. Some data were collected on defence failures during the summer 2007 floods (Royal Haskoning report, 2008) • Sources of validation data are limited. In particular, information on damages experienced during real floods is difficult to access and interpret because of the format in which they are made available. A data set describing some form of long term average losses over a large scale (catchment or national) would be a great asset in validating damage estimates. • Direct analysis of the most useful model outputs may provide the most convincing evidence for the success or otherwise of a model in generating results that are fit for purpose, but the lack of validation data of this type means that this will rarely be achievable. Instilling user confidence will be more difficult using indirect data. • Model validation generally relies on validation for a few test sites, and extrapolation of the error to untested sites. There are two key issues here: • How applicable are results from one test site to another? Evidence from the other studies has shown that

Past study	Summary of study and lessons
	confidence in results can be highly variable spatially; and • How do we aggregate these results to larger scales? Do the errors add up, or cancel out or are other statistical methods used to upscale the results? (from Validation of probabilistic flood models SC09008/WP1, Phase 1 Final Report, Environment Agency, 2010)

Table 2.1 Summary of lessons from past studies

3 Uncertainty of interest

There are many types of uncertainty. The sources of uncertainty that are of interest here can be classified into two forms:

- **Quantifiable knowledge uncertainty (QKU)** arising from errors in crest levels or specific model components (such as the Rapid Flood Spreading Method, RFSM) that can be credibly described in probabilistic terms.

- **Gross error potential (GEP)** arising from either an incompleteness of the model structure (e.g. exclusion of pumps when pumps are known to control flooding in a particular area) or systematic errors (such as an embedded bias towards overestimating the chance of a breach) that cannot be credibly expressed in probabilistic terms.

These two classes of uncertainty are illustrated in Figure 3.1. By understanding these two aspects of uncertainty an overall understanding of the confidence in the output results, and hence the appropriate means of communication and business response, can start to be understood.

Each class of uncertainty is discussed in more detail below (Section 3.1 and 3.2) together with the type of uncertainties excluded from the Route Map (Section 3.3).

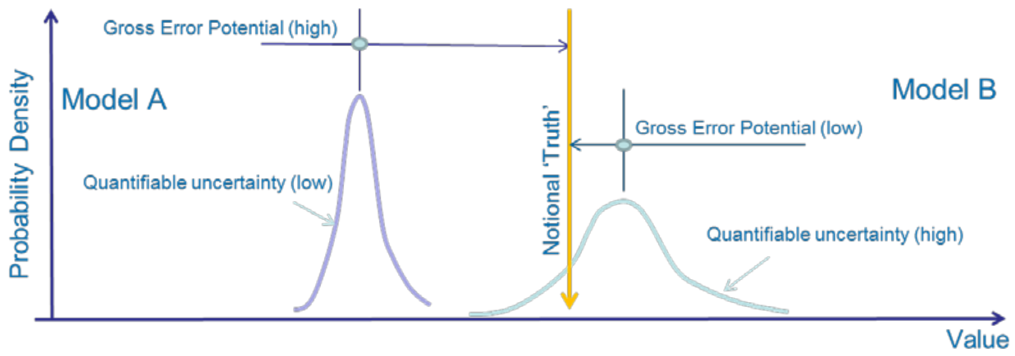

Figure 3.1 Quantifiable knowledge uncertainties and gross errors - their impact on results

3.1 Quantifiable knowledge uncertainties

In some cases it is possible to make quantified probabilistic statements regarding our lack of sureness in the data and models used to support NaFRA. For example, a crest level derived from low level Light Detection and Ranging (LIDAR) or a 1:100-year water level derived from a local detailed study. These 'knowable' uncertainties are referred to here as *quantifiable knowledge uncertainties* (QKU) and arise from errors in:

- *Data* – reflecting the quality of data inputs such as crest levels, water levels, fragility etc.

- *Risk model (e.g. RASP)* – reflecting the errors within the individual components (e.g. the RFSM) and the framework of analysis (e.g. sampling approaching)

- *Uncertainty model* - regarding the errors introduced through the structure of the uncertainty analysis itself, including:

  - *Correlation of uncertainties within the inputs* - for example between (i) individual datasets, such as the crest levels of adjacent defences (if one is higher than expected is it likely the next one is too?) or in water levels at a point and adjacent points (if one return period water level is low so is another?), and (ii) two or more datasets (such as tidal levels and fluvial levels near the tidal limit).

  - *Spatial aggregation of the uncertainty in the results* – the aggregation from Impact Cells to Impact Zones to Flood Areas to national figures. Although probability of inundation is calculated at the Impact Cell scale, errors in the digital terrain model (DTM) or inflow volume can have Flood Area wide implications on flow paths, depths and probabilities. The correlation between uncertainties is difficult to determine. For example, it is unlikely to be appropriate to sum the percentile values across all Impact Cells (as this may lead to an incredibly large uncertainty in the aggregated estimate). Equally, it is unlikely to be appropriate to assume uncertainties between ICs are fully uncorrelated or sum through random sampling (as this may lead to incredibly small uncertainty in the aggregated estimate as errors cancel out). The most appropriate approach will need careful consideration.

3.2 Gross error potential

Gross errors have the potential to arise due to fundamental mismatches between the modelled system and reality. It is often impossible (meaningless) to make probabilistic statements regarding the influence such a mismatch may have on our confidence in the output results. For example, the model may be incomplete (e.g. it simply does not include important features of the real flood system, for example the influence of groundwater or pumped drainage) or may lack the physics to represent particular processes (e.g. the influence of momentum in flood spreading when simulating a breach or in steep floodplains). In some situations these differences may have a profound influence on flood risk in others less so.

The Route Map outlined in the following sections attempts to provide credible (non-probabilistic) insights into the potential gross errors by considering known mismatches between:

- *Physics* – The physics included in the model components and those that are of importance in a particular flood area (e.g. momentum within the spreading models).

- *Risk model structure* – The incomplete nature of the model structure (e.g. pumps are currently excluded from RASP models but may be an important control in a particular Flood Area).

- *Uncertainty model structure* – i.e. individual uncertainties are considered to be random errors.

- *Systematic errors* – i.e. the DTM is 1m lower, or all water levels 0.5m higher.

It should be noted that the focus here is on fluvial and coastal flooding only – not flooding per se. Uncertainty introduced due to this limited remit is excluded.

3.3 Uncertainties excluded from the Route Map

No attempt is made to address:

- *Aleatory uncertainties* – those arising from the randomness in nature, driving a range of return period storms for example, are explicitly addressed within the RASP modelling approach (explicitly modelling a range of return period events) are not addressed further here.

- *Uncertainty in future change* – those that depend on exogenous change (out of control of the flood risk manager) such as climate change scenario, funding etc. The uncertainty in the change of flows or sea level, for example, arising under a given scenario could be considered a quantifiable knowledge uncertainty and hence is covered here.

4 Potential approaches

4.1 Quantifiable knowledge uncertainty

The most common and convenient approach to assess the quantifiable uncertainty is through some form of ***probabilistic approach***. Typically errors in the input data and model components are described by a probabilistic distribution and associated parameters (based on ground truthed evidence or judgement) and then propagated through the risk analysis model yielding probabilistic information on the uncertainty in the output results.

- *Advantages* – Provides quantified uncertainty estimates based on a transparent (if somewhat complex) approach.

- *Disadvantages* – Can be perceived as a black-box and (incorrectly) assumed to capture all uncertainty. Fails to capture potential gross errors arising from the incompleteness in physics/model structures or systematic errors in the data/model (Table 2.1).

4.2 Gross error potential

By their very nature potential gross errors do not lend themselves to probabilistic description. This reflects our fundamental lack of knowledge regarding the degree to which they may exist in any flood system and the influence they may have on the assessed risk. This does not mean, however, that we are in total ignorance. As introduced earlier many of the drivers of potential gross errors are known (e.g. the lack pumps in the system model). It is therefore possible to explore potential gross errors (and make qualitative statements regarding their likely impact) through non-probabilistic approaches, such as:

What-if testing (model based scenarios / sensitivity testing)

The user perturbs the inputs to the model to explore the sensitivity of the output results to changes in the input. For example, what-if the fragility curves are assumed to be twice as strong.

- *Advantages* – Simple, user driven to explore impact of potential errors

- *Disadvantages* – Fails to provide an understanding of how likely errors are. As with probabilistic approaches it fails to capture incompleteness within the model or systematic errors in the data/model.

Structured qualitative description and multi-criteria assessment

A structured narrative is developed for each flood system typology and the likely impact of potential gross errors on the uncertainty in the output. The narrative would be developed based on multiple evidence strands. The CI approach is one example of this type of approach.

- *Advantages* – Simple and (potentially) comprehensive in scope (i.e. complete inclusion of available knowledge)

- *Disadvantages* – Non-quantified

4.3 A comprehensive uncertainty assessment

To establish a comprehensive understanding of uncertainty and its potential impact on decisions, those uncertainties that can be quantified should be and the potential for gross errors understood. The approach proposed within this Route Map is based on this combined approach.

- *Advantages* – (Potentially) comprehensive in scope and defensible. Applicable to the widest possible set of business uses. Supports a more honest, defensible, description of uncertainty that includes a qualitative and quantitative component.

- *Disadvantages* – A combination of methods required and potentially more complex communication

A combined assessment could, for example, be used to populate an ***'uncertainty matrix'*** as shown in Figure 4.1.

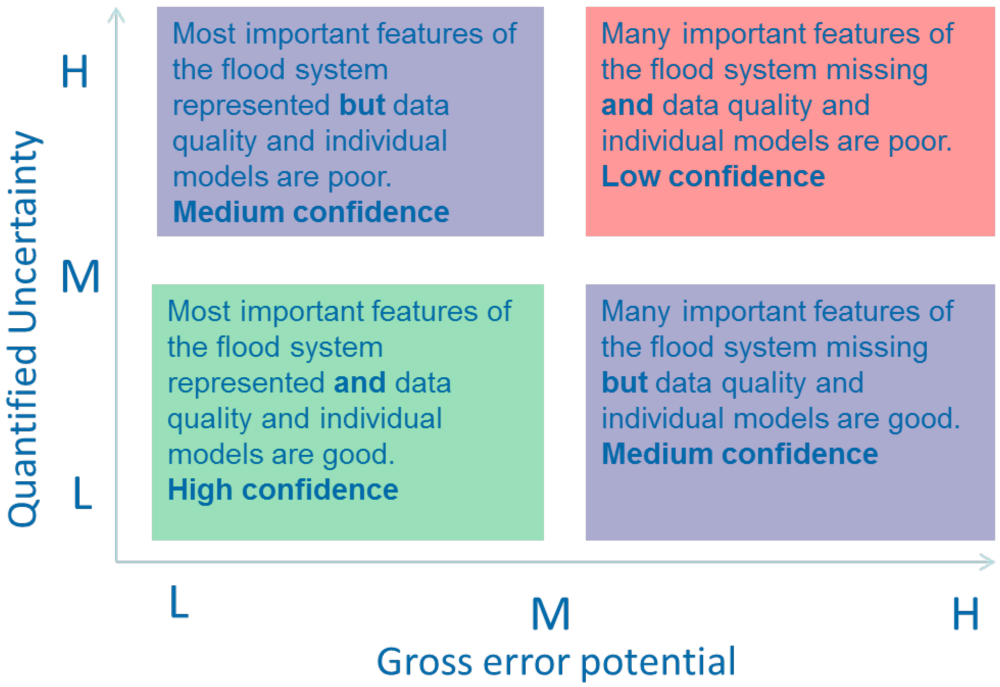

Figure 4.1 An 'uncertainty matrix'

Presentation as a matrix (rather than a single measure) strengthens the idea that QKU and *potential gross errors* are distinct uncertainty types – both of which are important and both will affect how the information is used and appropriately responded to. For example, the bottom left of Figure 4.1 shows a low gross error potential (i.e. the model used is considered to represent all of the key processes that influence flood risk in a particular area) but with a low quantified uncertainty (i.e. the input data, such as crest levels, and/or the models used are good). Overall this implies a 'high degree of confidence in the results'. This can be contrasted with the top right where the gross error potential is high (as many important features of the flood system are missing from the model, e.g. pumps) and the input data quality / individual model components are poor. Overall this implies a 'low confidence' in the results.

Figure 4.1 is of course for illustration and it will be necessary to link these types of descriptions to specific management activities before taking forward into operation.

5 Overview of the Route Map

The Route Map has been divided into five activities in order to implement a combined assessment of uncertainty (including insights into both QKU and potential gross errors).

Each activity is designed to provide a contribution to an improved understanding of confidence and its use. Each activity is structured to complement (not duplicate) each other and as a whole will provide a credible and defensible assessment of confidence and uncertainty that addresses a wide range of business needs.

The five activities are listed below and discussed in more detail in the following sections.

- **Activity 1 – Explore and map business needs, responses and useful measures of uncertainty** – seeks to better understand the specific business needs; the single and collective measures of uncertainty and confidence that would be of most use to the Environment Agency and its partners (including the public), and how the business should/can respond to this additional information.

- **Activity 2 – Improve the Confidence Indexing approach** – seeks to improve the credibility of the Confidence Indexing approach.

- **Activity 3 – Assess quantifiable knowledge uncertainties** – seeks to address the QKU (within a probabilistic framework) and provide quantified estimates of uncertainty.

- **Activity 4 – Asses the gross error potential** – seeks to provide insight into the potential for gross errors in the results (arising from incompleteness or lack of representativeness of the approach in different typologies).

- **Activity 5 – Validate, review and update approaches** – establishing a process of validation to help establish if the estimates of uncertainty are of the 'right order'. As a minimum this should include a process to audit the specific uncertainties included and those excluded from the analysis. A general process of on-going review and improvement will also be important to reflect new evidence and methods as they emerge.

A high level description of the interaction between these activities is shown in Figure 5.1.

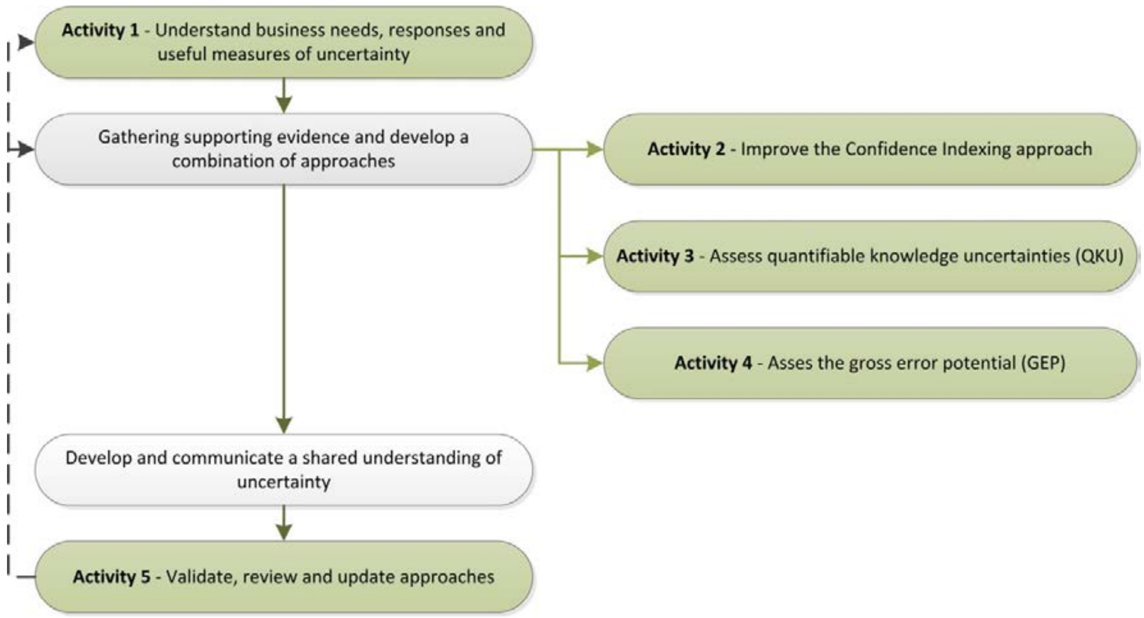

Figure 5.1 High level Route Map of activities

6 Activity 1 – Business responses and useful measures of uncertainty

6.1 Introduction

Uncertainty has always been a central consideration in the assessment of risk and the decisions taken. To date, due to the lack of credible and useable evidence on uncertainties, these considerations have typically been based on ‘expert judgement’. Equally, take-up of uncertain information has been hampered by the complex language often used to express uncertainty and confidence. Consistent, well-understood and understandable semantics are only just starting to become established in practice (although a significant body of academic literature exists).

This first activity within the Route Map is therefore to explore:

- How the Environment Agency and its partners would use new information on uncertainty to make better choices
- The capacity different partners have to take it on board
- The measures of uncertainty and confidence that are perceived as most useful
- The type of questions the information would help answer
- How best to communicate the results.

6.2 Supporting wider use of the Confidence Index data across the business in the short term

The analysis undertaken in 2012/13 using the CI approach provides a new understanding of the level of confidence in the assessed likelihood of flooding. Although a first step this analysis provides a number of valuable insights that could be used across Environment Agency business activities (and perhaps more widely) to support better decision making, including for example:

- Investment planning
- Planning and development control
- Communication with insurers and other stakeholders
- Communications on WIYBY, etc.

6.3 Scoping the business needs, responses and communication in the medium to longer term

The Confidence Indexing approach does not provide, on its own, a complete understanding of the uncertainty within present day and future risks. Equally, although it provides some guidance on the sources of uncertainty in any Flood Area more could be done to highlight the key drivers of uncertainty and guide the most appropriate responses to it (e.g. would data or model improvements help or are these good enough).

This task would scope out the medium to longer term business needs, responses and opportunities for communicating uncertainty and confidence in NaFRA. The proposed scoping activity would look more broadly than simply the CI to help understand the benefits of assessing uncertainty in a consistent way across the business, including:

Business needs and drivers – Understanding the longer term drivers for information on confidence and uncertainty across the Environment Agency business areas, its partners and the public will be a vital first step.

Appropriate decision responses to uncertainty – The availability of increasingly rich information on uncertainties will challenge existing decision making across the full range of business activities. A key activity within the Route Map is to understand this, including:

- Developing new protocols on decision making and the justification/targeting of investment
- Developing a more structured understanding of 'what is good enough' for different uses
- Understanding what the new information can and cannot be used for.

Appropriate measures and descriptors of confidence and uncertainty – With new information comes a challenge of communication both within and external to the Environment Agency. 'What' is communicated as well as 'how' it is communicated are equally important questions and both will need to be addressed, including:

- *Means of communicating uncertainty?*

  There will be a need to develop a common message and syntax across the Environment Agency business using, for example, a combination of quantified estimates, narratives, star-rating, etc. Standardising the visualisation tools, such as maps and tables, matrices etc, will also be important.

- *What spatial scale is appropriate to represent uncertainty?*

  Is it appropriate to express the uncertainty at an individual property, Impact Cell or flood defence or should the uncertainty be expressed at a Flood Area or SAMP scale? Input data can be sensibly shown at their native scale (defence parameters at defence scale, DTM level at Impact Cell (IC) scale, etc), but aggregating quantified uncertainties in probability and damages by Impact Zone (IZ), Flood Area or larger scale is problematic and will need careful consideration (in terms of how the results would be used and how to credibly calculate them).

- *Specific responses and guidance*

Could more credible information on uncertainty help provide clearer messages to the Environment Agency partners? For example, it is often said that the Environment Agency and its partners should consider that the uncertainty in the depth of flooding data for a given return period is 300mm. This is an important assumption and influences a number of decisions (e.g. development). Two questions arise: (i) Is the uncertainty within NaFRA within this tolerance? (ii) Is that level of accuracy good enough?

6.4 Recommendations for taking Activity 1 forward

The recommended approach for taking Activity 1 forward is to undertake a single study that addresses both the short term use of the CI and sets out the longer term business responses. This would be a relatively small study (max. six months) including a combination of consultant and Environment Agency staff time.

The study would include as a minimum:

- A workshop with the Environment Agency and its partners to scope the measures and scales of interest
- A desk study to develop new protocols of decision making and fabrication of key measures that could be obtained at different scales
- A report

Likely external cost: £30-50k.

7 Activity 2 – Improve the confidence indexing approach

7.1 Introduction

The application of the CI approach in 2012/13 was the first uncertainty assessment applied at a national scale. As with all first applications, the experience of implementation provides an opportunity for critical review and improvement. Activity 2 focuses on capturing and implementing these lessons.

7.2 Recommendations for taking Activity 2 forward

The CI approach provides a useful step forward and it is not proposed to revisit the basic methodological framework. Instead a series of steps to improve the existing approach are proposed below in Table 7.1.

CI#2.1	**Gathering of feedback from Environment Agency staff**			
Gather feedback from the various local teams that contributed to the NaFRA Local Review exercise from November 2012 through to early April 2013. The review should include: (i) understanding of the CI (star-rating) and how it is established; (ii) how they see this additional information being used; (iii) the usability of the tools and guidance, and (iv) the scoring method associated with the local evidence and priorities for potential improvement. It is recommended that this is undertaken sooner rather than later to ensure the lessons are captured while fresh and momentum is maintained.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
High	Low	4m	None	**High**

CI#2.2	**An interrogation of the results**			
The results for key typologies (guided by the analysis undertaken for the Benchmarking studies) should be reviewed to: (i) identify representativeness of the typologies (too many, too few, missing); (ii) identify the typologies with greatest / least uncertainty; (iii) identify key drivers of uncertainties in different typologies, and (iv) highlight specific technical and software issues (building on the feedback from CI#2.1). This is an important step before finalising any updates of the approach and software.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Moderate	4m	Logically completed in parallel with #CI.1	**High**

CI#2.3	Update of the method accordingly			
The update of the method will be contingent on the findings of tasks 2.1 and 2.2 above. It is possible however to identify those aspects that are likely to require modification in order to improve the credibility and utility of the weighting and scoring approach (without modifying its basic structure). These are captured in Annex 1.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Mod	3m	On CI#2.1 and 2.2.	Depending on findings from above

CI#2.4	Update the software as necessary			
The update of the software will be contingent on the findings of Task 2.1-2.3 above. Any updates that are required will also need to be incorporated into an appropriate release of MDSF2 and supported by associated activities of user acceptance, manual updates and any necessary training.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Mod	4m (could be more if additional typologies are introduced for example – or less if very little changes)	CI.#2.1 and 2.3.	Depending on findings from above

Table 7.1 Recommended activities to improve the Confidence Indexing approach

8 Activity 3 – Assess the quantifiable knowledge uncertainties

8.1 Introduction

Activity 3 seeks to provide a probabilistic assessment of the uncertainties arising from quantifiable errors in the input data and model components (i.e. largely the same issues as assessed through an expert scoring and weighting method by the CI).

Various attempts have been made to quantify the uncertainty in NaFRA using probabilistic techniques. None of these previous approaches, however, have been successfully taken up by the Environment Agency. This reflects a number of issues concerning the lack of agreement regarding the methods themselves, a lack of evidence to express the uncertainty in the input data, the model components and model structure used. Perhaps the single most important factor has been a lack of a widespread business demand to support the development and testing of the approaches. This is no longer the case and the demand for information regarding confidence and uncertainty has significantly matured in recent years[1].

8.2 Benefits of moving towards a quantified approach

A quantified approach is a natural extension of the weighting and scoring approach developed in 2012 in support of the Confidence Index. A quantified approach however has a number of potential advantages over the Confidence Indexing method. For example, once implemented within RASP (such as RASP-SU or LH-OAT) fewer assumptions would need to be made regarding the way input data and model uncertainties combine to influence the NaFRA result within any given Impact Cell. Specifically a probabilistic assessment of QKU embedded within RASP would remove the need to:

- *Predefine the influence of individual data and model uncertainties on confidence* – this would no longer be required since it is implicit in the quantified methodologies and can be output explicitly as parameter of sensitivity in the results.

- *Define a Flood Area Data Quality Score (DQS)* – this would not be needed, as system scale effects would be explicit in the approach.

- *Map confidence scores to uncertainty* – this would not be needed as a quantified uncertainty by IC would be directly returned by the calculation.

---

[1] The Agency has various drivers for moving towards some form of quantified estimates of uncertainty (e.g. NAO and the review by Prof. Paul Bates) and recurring high level questions regarding how good the results are and what they can reliably be used for.

A quantified approach would also provide a number of important additional insights:

- *Uncertainty in the outputs*: including the uncertainty in depth v probability relationship, Expected Annual Damages and other outputs.
- *Attribution to drivers*: enabling the contribution that the uncertainty in a particular input data source or model component makes to the overall uncertainty in the output to be quantified.
- *Ranking of drivers*: enabling individual input data sources to be ranked according to how sensitive the output is to their associated uncertainty.

Table 8.1 and Table 8.2 provide an initial summary of the opportunities provided by the moving towards a quantified approach for each business use and business group in the Environment Agency. These tables are based on the teams view; rather than consultation with the Environment Agency at this stage but hopefully still provide a useful overview.

Business use	**Opportunity provided (subject to understanding of limitations)**	**Information on uncertainty of interest**			**Added value**
		Output	**Attribution to drivers**	**Ranking of drivers**	**(H/M/L)**
National policy choices (e.g. LTIS)	Quantified uncertainty in present day and future risks (given a future scenario).	***	*	**	H
Strategic planning choices (e.g. SMPs, CFMPs SAMPs)	*As above.*	***	**	**	H
Local investment choices (e.g. through project and scheme appraisals)	*As above.*				
Influencing spatial planning (e.g. development control and local structure plans)	Quantified uncertainty in probability, depth and velocities in the present day and future cases.	***			M
Forecasting and warning	Quantified uncertainty in probability, depth and velocities for forecast events.	***			H

Business use	Opportunity provided (subject to understanding of limitations)	Information on uncertainty of interest			Added value
		Output	Attribution to drivers	Ranking of drivers	(H/M/L)
Detailed design choices (e.g. freeboard, material types)	Quantified uncertainty in a full range of issues (e.g. failure probability, risks, flood probabilities, depths and velocities).	***	***	***	H (long term opportunity, unlikely to be possible/practical in medium term)
Minor investments (e.g. maintenance)	None – but better directed based on SAMPS above.				
Partner and public communications	Spatially differentiated communication of confidence and what specific information can be used for.	***			H

Table 8.1 Business use - potential opportunities provided by a probabilistic approach

Business group	Opportunity provided	Information on uncertainty of interest			Added value
		Output	Attribution to drivers	Ranking of drivers	(H/M/L)
Investment and planning	To communicate level of confidence / uncertainty in national / regional assessments, to test robustness of conclusions, and to highlight areas of models / data that need to be improved.	***	**	*(*)	H
Asset management	Better targeting of monitoring, data collection and interventions.	***	**	*(*)	H

Business group	Opportunity provided	Information on uncertainty of interest			Added value
		Output	Attribution to drivers	Ranking of drivers	(H/M/L)
Incident management	Better targeting of responses (directing resources internally, promoting forecast model improvements).	***	**	*(*)	H
Modelling and mapping	Better targeting of responses (directing resources internally, promoting data and model improvements).	***	**	*(*)	H
Engagement	Better understanding on information use.	***			H
Partnership and external funding	Support to improving data and models, and adopting adaptive strategies.	***		*	L
Strategic overview	Better understanding of year-on-year risk change, on-going improvements and priorities for action.	***	**	*	H

Table 8.2 Business group - potential opportunities provided by a probabilistic approach

8.3 The feasibility of taking forward a probabilistic method

Recent years have seen various improvements that make a credible probabilistic assessment of the quantifiable knowledge uncertainties a much stronger possibility. The enablers of the approach and the barriers that persist include:

Enablers

- *Experience* – The various previous attempts (Table 2.1) and the recent successful application of the Confidence Indexing approach provide a much more 'rounded' view of what is feasible and what is required.

- *Benchmarking* – Through the on-going benchmarking studies a more structured understanding of the performance of the system models is starting to emerge. This understanding provides a key support in describing the uncertainty introduced by the modelling process (components and structure) in different typological settings.

- *Ground truthing* – Some additional work has been undertaken to provide better evidence on the various model inputs. See Annex 2 for an overview.

- *Computational power and work flows* – Access to higher specification computing (including distributed computing) continues to reduce runtimes. NaFRA/MDSF2 work flows continue to be optimised and managing and running multiple runs (selecting parameter values and modifying databases) is increasingly straightforward (a prerequisite of a probabilistic approach).

Barriers

- *Ground truthing* – A comprehensive set of values is still not available. A further work on this aspect will be of fundamental importance.

- *Incompleteness* – The basic structures of the RASP calculation have remained unchanged for some time. The issues associated with the incomplete representation of the flooding system that RASP provides (in some typologies) therefore remains a key barrier to probabilistic assessment (hence the proposed the development of a more qualitative descriptor of gross error potential – see Activity 4).

- *Difficult science questions* – Careful consideration will be needed to determine how best to represent individual and correlated uncertainties as well as spatially aggregate uncertainties from Impact Cell through to national scales.

- *New methods and software will be required* – Past uncertainty models (RASP-SU/LH-OAT) have been written for NaFRA 08. These will need to be developed to reflect the new methods and integrated into the MDSF2 structures.

- *New communication tools* – Communicating uncertainty remains a challenge – developing clear and understandable messages will be vital.

- *Resources required* – Achieving a quantified assessment of uncertainty is not a straightforward goal; it requires science, development and implementation effort. This will require appropriately skilled scientific resources, advanced computation resources and focused communication resources.

8.4 Recommendations for taking Activity 3 forward

Table 9.1 summarises the activities required to take forward a probabilistic assessment of the QKU. It is recommended that Activity QKU#3.1 is complete prior to initiating others. Once complete Activities QKU#3.2-3.6 would be appropriately commissioned as a single project with a Gateway decision to decide whether or not to implement within MDSF2 (QKU#3.7).

Likely external cost of QKU#3.1: £50-100k

The potential development of a probabilistic approach to assess the quantifiable knowledge uncertainties is discussed in more detail in Annex 2.

<table>
<tr><th>QKU#3.1</th><th colspan="4">Scoping stage and business case</th></tr>
<tr><td colspan="5">This first task scopes the alternative approaches for the probabilistic propagation of uncertainties through the RASP engine (as enacted in MDSF2). It is likely that no single approach will be uniquely preferred for all business needs and consideration should be given to a hierarchy of approaches; allowing the user to turn on and off parameters, and to choose the outputs of most interest/use. The runtime and supporting data and development needs of each approach will need to be explored to support a detailed business case for taking forward a quantified approach.

To ensure a credible scoping stage it is recommended that it is supported by comparisons at selected pilot sites of LH-OAT and RASP-SU using the latest RASP engine and common distributions using the latest available ground truthing results (see Annex 2). This will therefore require some (limited) method and model development / testing.</td></tr>
<tr><td>Environment Agency resources</td><td>External resources</td><td>Duration</td><td>Dependence</td><td>Priority</td></tr>
<tr><td>Low</td><td>Medium</td><td>8-12m</td><td>None</td><td>High</td></tr>
</table>

<table>
<tr><th>QKU#3.2</th><th colspan="4">Quantifying and expressing uncertainty in data inputs</th></tr>
<tr><td colspan="5">Of fundamental importance to the credibility of a quantified approach (notwithstanding the associated issues of potential gross errors addressed in Activity 4) is the credibility of the evidence used to establish the probability distributions associated with the input data. Within this task the existing ground truthing will be extended to provide a credible and comprehensive evidence base. This will, for example, enable uncertainties within a crest level with DQS of 2 to be determined and distinguished from a crest level with a DQS of 3 and so on.

This task will result in a structured presentation of the evidence that supports a probabilistic description of the errors associated with different DQSs (including both distribution type and distribution parameters). It will also need to describe any correlations between errors (supported by observation from the field and / or argument).</td></tr>
<tr><td>Environment Agency resources</td><td>External resources</td><td>Duration</td><td>Dependence</td><td>Priority</td></tr>
<tr><td>Low</td><td>Medium</td><td>6m</td><td>QKU#3.1 and CI #1-3.</td><td>High</td></tr>
</table>

QKU#3.3	**Quantifying and expressing uncertainty in model performance (from benchmarking studies)**			
The results of the on-going MDSF2 Benchmarking studies will help provide quantified descriptions of the model uncertainties in different typologies. This task will use those results, together with expert evidence where needed, to construct appropriate probabilistic descriptions (distributions and parameters) that reflect the uncertainty introduced by each model component(s). It will also need to describe any correlations between errors. This task will need to be carefully thought through to ensure these aspects are appropriately introduced to the uncertainty analysis.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Medium	6m	QKU#3.1	**High**

QKU#3.4	**Method development, prototyping and initial piloting**			
In this task existing approaches (LH-OAT and RASP-SU) would be reviewed in detail (building on the scoping work) and developed as appropriate; bringing them together into a single hierarchical approach (rather than two distinct approaches) or, if necessary, replacing them. The resulting method would then be enacted in prototype code and applied to a set of common pilots (most logically those used within the benchmarking studies). The key outputs would be the prototype models, a science report and recommendations for implementation. The results of the analysis undertaken on the pilot sites will also provide valuable insights to support future updating of the Confidence Indexing approach (specifically the weighting given to individual Data and Model Performance attributes in different typologies and the mapping of the Confidence Index into a quantified estimate of uncertainty).Note: The LH-OAT and RASP-SU applications together with initial options for enhancement are discussed in more detail in Annex 2.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	High	12-18m	QKU#3.1-3 (results would support the CI approach)	**High** (the added value of the previous tasks are not fully realised until this step is complete)

QKU#3.5	**Develop communication approaches and prototype**			
Given the much richer picture of uncertainty available from the quantified applications this task would consider how best to communicate this information to different business uses.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Low	3m	QKU#3.4	Importance is high once previous tasks complete

QKU#3.6	**Extended piloting and develop credibility with uses**			
The prototype codes, methods and communication approaches would be trailed more extensively and finalised. The results of this analysis would support: Further improvement to the Confidence Indexing approach Prepare the way for implementation in MDSF2 if considered appropriate				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Med	6m	QKU#3.4	Importance is high once previous tasks complete

QKU#3.7	**Update MDSF2 software and manuals**			
If considered appropriate at this stage the methods could be embedded into the MDSF2 software and associated manuals and training provided.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	High	6-10m	QKU#3.4	To be determined based on need at time

Table 8.3 Recommended activities for taking forward an assessment of quantifiable knowledge uncertainties

9 Activity 4 – Assess the gross error potential

9.1 Introduction

Activity 4 seeks to develop an indicator of the 'Gross Error Potential' (GEP). The GEP reflects the degree to which the modelling processes within NaFRA/MDSF2 fails to represent the reality of any given typology (including the potential for systematic errors). The GEP is not envisaged as a quantified parameter but rather a qualitative statement (for example from high to low) that reflects a combination of:

- *A statement of fact* – reflecting 'known unknowns' (i.e. those components of each typology that are excluded, or only superficially included, within the current modelling approach).

- *A judgement of error potential* – reflecting 'unknown unknowns' (i.e. systematic errors in ground level or damage functions or simply 'wrong' analysis).

The assessment of the GEP is likely to follow a similar process to that developed in support of the CI (i.e. structured expert judgement supported by evidence where possible).

A good starting point for the assessment of the GEP is the typology of the Flood Area. The typology, developed as part of the Confidence Indexing method, attempts to provide a complete characterisation of the different physical settings in all Flood Areas. This includes aspects that are not currently represented in the flood risk model. For example, 'complex defences' highlights that a Flood Area may include (or be influenced by) multiple point assets such as pumps or barriers, neither which are explicitly represented within MDSF2 or the NaFRA toolset. Equally, the typology highlights those Flood Areas where the important physical processes may be missing or the underlying assumptions within RASP may be wrong. For example, in complex drainage channels, where the independence of flood areas may be invalid or where the valley may be steep and momentum in the flow is important (and currently poorly represented within the RFSM). Coupled with the understanding for the benchmarking activities (and in particular comparisons to more complete system models where available) the most important emissions from the current system model could be highlighted and reflected in a descriptor of the GEP.

9.2 Recommendations for taking Activity 4 forward

The recommended approach is to first scope in more detail the GEP descriptor (GEP#4.1) before moving forwards with the method and associated software developments (GEP#3.2 and #3.3).

The likely external cost for GEP#4.1: £30-40k.

GEP#4.1	**Scope**			
The first stage would be to scope the types of uncertainty that cannot credibly be described using probabilistic approaches and have the potential to cause gross errors in the results of NaFRA / MDSF2. For each source of gross error (both statements of fact and judgment based suggestions) would be described and the way in which they may manifest themselves within the results explored, including those sources where a detailed understanding of the issue is unknown. Understanding the relationship between the potential for gross error and the business use will also be a key consideration.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Med	3-6m	Understanding of the latest typologies. Insights would be useful additions to both the CI and QU approaches.	**High**

GEP#4.2	**Develop method and pilot**			
If considered appropriate a GEP descriptor would be developed in this task. This would be a process similar to the Confidence Indexing approach and could be viewed as the second dimension on an 'uncertainty matrix' introduced in Figure 4.1. One axis of the matrix would represent the quantified uncertainties (expressed as a CI or distribution) and the second axis representing the 'gross error potential'.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low/Med	Low/Med	4-6m	GEP#4.1 and Activity 1	**High**

GEP#3	**Update MDSF2 software**			
If considered appropriate a descriptor of the GEP would be relatively easily implemented in MDSF2 in a similar manner to the CI. Associated manuals and training would need to be updated.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Low/Med	3m	GEP#4.2	**High**

Table 9.1 Recommended activities for taking forward an assessment of the 'gross error potential'

10 Activity 5 – Audit, validate, review and update

10.1 Introduction

Validation through some form of comparison with observations or alternative approaches (or sources of evidences) will be vital to add credibility to the overall approach. This is not straightforward – in conceptual and in practical terms. This does not however mean it is impossible.

Activity 5 seeks to develop and implement a process of audit, validation, review and update of the supporting methods, software and the evidence on which they are based.

10.2 Recommendations for taking Activity 5 forward

Two sub-activities are proposed within Activity 5 (Table 10.1). The first focuses on developing a framework for validating the results of the uncertainty analysis (across all methods). The second activity focuses on the process of update and review.

It is recommended that Validate and Review (VR) #5.1 is completed as soon as reasonably practicable.

Likely external cost: £40-60k.

The priority associated with VR#5.2 will reflect the degree to which an understanding of uncertainty is adopted within the Environment Agency business processes.

VR#5.1	**Audit and validation framework**			
This task will develop a process through which the credibility of the results can be established. In developing this process the framework set out in the recent R&D project covering the 'validation of probabilistic models' should be considered in detail. Central to the success of this process will be an understanding of the different business uses of uncertainty information (Activity 1) and the appropriate means of validation in the context of that use.				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Med	4-6m	Activity 1 (Most logically follows the review of the CI approach and the scoping of the approaches in support of the quantified approach and assessment of the potential gross error).	**High**

VR#5.2	**Update and review**			
This task reflects upon how the methods and information on uncertainty is actually being used within the business, the barriers to wider take-up (if any) and the enablers where take-up is best. It will need to set out lessons learnt and prioritised actions for on-going improvements to the methods, software, training or supporting evidence (if any).				
Environment Agency resources	*External resources*	*Duration*	*Dependence*	*Priority*
Low	Low	3m	Assumes an on-going process of confidence assessment	**High**

Table 10.1 Recommended activities to support the process of audit, validation, review and update

11 Summary of priority recommendations

The Route Map outlined in the previous chapters seeks to provide an appropriately complete assessment of the uncertainty within NaFRA. Each activity contributes to the development of a common understanding of the uncertainties within NaFRA, with insights gained from one improving the credibility of the estimates from another. MDSF2 provides the common software framework within which each approach is implemented, with the user choosing the approach that best reflects their needs (should they require anything above the base level of uncertainty assessment that the Environment Agency adopts). The relationship between the various activities (and the tasks within them) is shown in Figure 11.1.

The complete Route Map may take a number of years to implement. This does not mean, however, that progress cannot be made in a much shorter time. Significant measured steps forward can be made by progressively completing various prioritised tasks, as follows:

Priority 1 **– Activity 2 - Improve the Confidence Indexing approach (CI#2.1-2.3)** – Capturing the lessons from the first application of the Confidence Indexing method while the approach and the results easily recalled and captured provides an important 'quick win'. This will improve the credibility of the Confidence Indexing approach in the short term and lay the foundation for the other activities proposed in the Route Map.

Priority 2 **– Activity 1 Explore and map business needs, responses and useful measures of uncertainty** – Developing a clear understanding of business needs and the specific measures of uncertainty and confidence that best support these needs is a fundamental requirement. This activity is therefore a vital precursor to the development of additional data gathering, methods or software.

Priority 3 **– Activity 3 – Assess the quantifiable knowledge uncertainties (QKU#3.1 Scoping and business case)** – A quantified assessment of uncertainty that avoids many of the assumptions made in the CI approach and provides a much richer picture of uncertainty and the drivers of uncertainty is a natural extension of the CI approaches. Scoping the requirements for, and approach to, a probabilistic assessment of quantifiable knowledge uncertainties is the first step towards this.

Priority 4 **– Activity 4 – Assess the gross error potential (GEP#4.1 Scoping)** – A persistent shortcoming of all previous attempts to assess confidence in NaFRA (including the CI approach) has been the emission of potential gross errors. Scoping the requirements for, and approach to, providing insight into the potential gross errors arising from an incompleteness or lack of representativeness of the modelling approach in different typologies is the first step towards this.

Priority 5 **– Activity 5 – Audit, validation, review and update of approaches (VR#5.1 Validation framework)** – Establishing a process of audit and validation (or at least collaboration) of the results in the context of different business uses will be important to develop a wider acceptance of the approaches used the results provided.

Priority 6 **- Activity 3 – Assess the quantifiable knowledge uncertainties (QKU#3.2-3.6)** – This Activity seeks to develop and apply a quantified approach to a series of pilot studies. The results will support the on-going development of the Confidence Indexing approach (by providing detailed assessment at pilot sites for a variety of typologies where the highest quality benchmarking data exist). These activities will also establish an equivalent probabilistic assessment capable of providing quantified estimates of uncertainty in the results and their sensitivity to errors in the input data and supporting models. Once piloted the decision to embed the approach within MDSF2 can be made.

Priority 7 **– Activity 4 Assess gross potential error (GEP#4.2-3 Develop method and implement).**

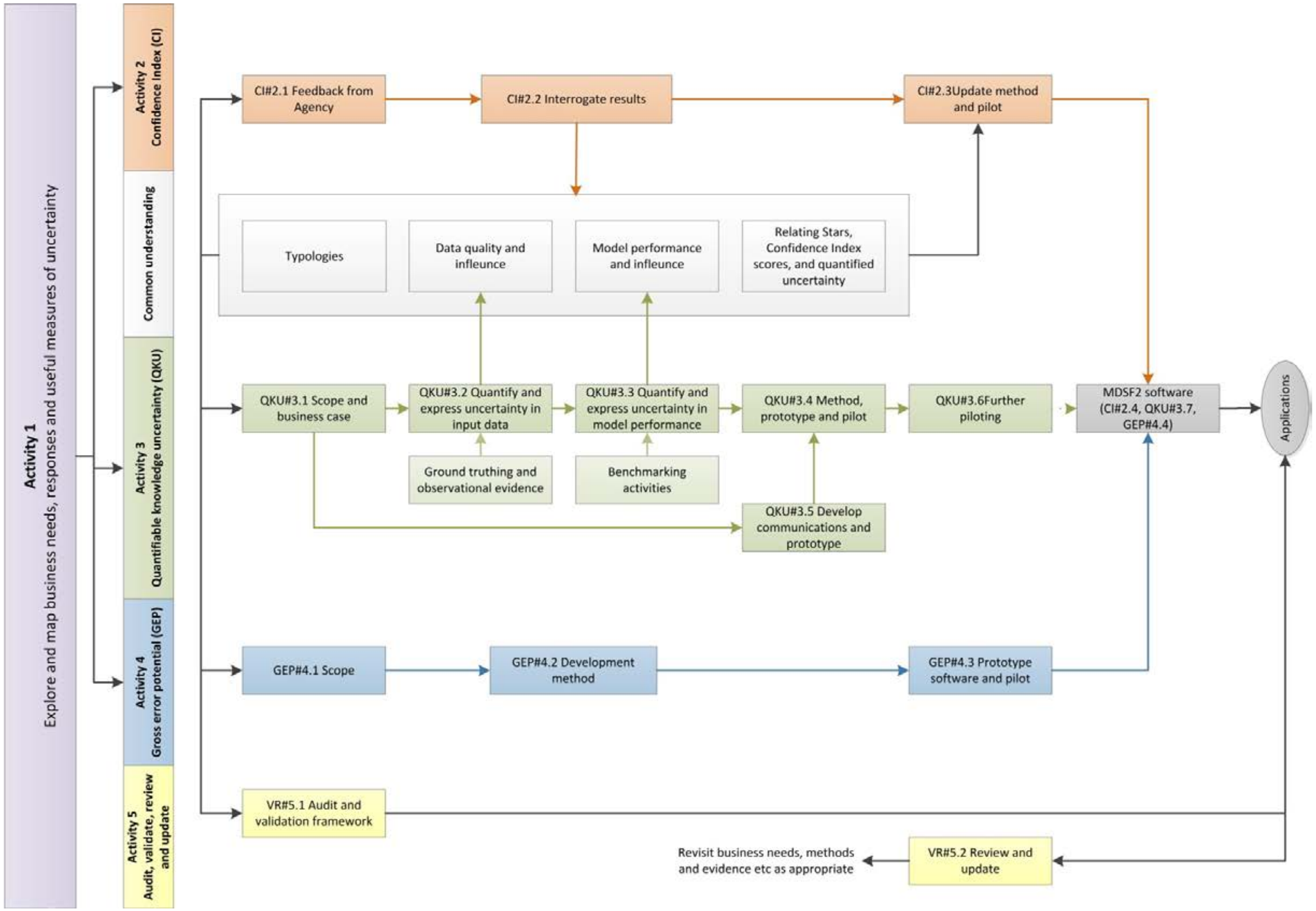

Figure 11.1 Detailed Route Map of activities

Annex 1 – Future enhancements to the weighting and scoring approach

Each step within the current approach (as outlined in the workflow contained in the Confidence Score Method Report) has been reviewed and specific recommendations made where considered necessary and practical to do so. These are summarised in Table A1.1.

Proposed activity	Rationale	Outcome	Priority	Resource effort
Review of typologies to fine-tune definitions in light of outlier and confidence analysis.	Disaggregation of most significant typology groups (if necessary/sensible) Aggregation of insignificant typologies Confirmation / modification of typology boundaries (i.e. >2m ht)	More focused typologies supporting clearer messages (i.e. targeted data and method improvements)	High	Low
Review of individual Data Quality Scores (i_DQS) using ground truth comparisons	Identify any redundancies or missing sources within the i_DQS Confirm relative ranking of the sources of individual data. Confirm the relative ranking between data scores with comment dimensions (i.e. mOD, inflow)	A more robust evidence in support of the i_DQS	High	High (requiring further ground truthing)
Review of the definition of the Flood Area Data Quality Scores (fa_DQS)			Not possible without quantified uncertainty analysis	
Review the Data Quality Index (DQI)	Update based on any revision to the typology definitions (above) Use experience from first application to review parameter influence on DQI.	A more robust evidence in support of the DQI	High	Low
Review the individual Model Performance Scores (i_MPS)	Current table assigns a default score of 3 to most model components. Requires review.	A more robust evidence in support of the MPS	Not possible / meaningful without quantified uncertainty	

Proposed activity	Rationale	Outcome	Priority	Resource effort
			analysis	
Review Flood Area Model Performance Scores (fa_MPS)			Not possible without quantified uncertainty analysis	
Review the Model Performance Index (MPI)	Update based on any revision to the typology definitions (above) Use experience from first application to review parameter influence on MPI.	A more robust evidence in support of the MPI	High	Low

Table A1.1 Potential short term improvements to the scoring and weighting approach (to be confirmed following review of results)

Review the assessment of the Confidence Index	**Review the assumption of equal weighting**	**A more robust evidence in support of the CI calculation**	**Not possible / meaningful without quantified uncertainty analysis**	**Resource effort**
Review the mapping of the Confidence Index to uncertainty in a given band	The linkage between CI and the quantified uncertainty is a central assumption (including the assigned SD, assumption of a normal distribution etc). These require review based on experience of first application and quantified analysis when available.	Greater credibility in the relationship between CI and uncertainty in a band and responds to a key query in previous Peer Reviews.	High	Low (review) High (quantified analysis)
Review the approach to combining joint fluvial and coastal Impact Cells	This relates only to NaFRA 08 models. No further review			
Review the approach to assigning the % correct to each likelihood band.	This would be included in above.			
Review the mapping between the	Review based on experience of first application and updates	A more credible evidence in	High	Low

Review the assessment of the Confidence Index	Review the assumption of equal weighting	A more robust evidence in support of the CI calculation	Not possible / meaningful without quantified uncertainty analysis	Resource effort
local evidence parameter score, the overall local score and the Star Rating	to the automated method above.	support of the mapping of local evidence to stars		
Undertake detailed case studies (to support the approach)	Exemplar sites can be a powerful means of gaining credibility. There was no time for this in the development and is proposed here.	Support to all above	High	Med
Update reports accordingly				

Table A1.2 Potential short term improvements to the scoring and weighting approach (to be confirmed following review of results) - continued

Annex 2 – Future development of probabilistic approach to assess quantifiable knowledge uncertainties

Comparison of past approaches

In particular two approaches have been explored and initial trials undertaken, namely the so-called LH-OAT and RASP-SU applications. These approaches are compared in Table A1.2.

Existing evidence to support quantified analysis

A quantified uncertainty analysis relies upon credible evidence on the uncertainty associated with the input data used, the performance of the models it runs and any errors introduced by the structure of the model used. Each of these aspects is discussed below.

Note: Gathering credible evidence on all of these aspects is critical for the quantified approach but would also support improvement of the weighting and scoring approach.

Gathering evidence on the uncertainty in the input data

Ground truthing has been undertaken for some parameters. These are summarised and discussed below. Modification to these values is straightforward regardless of the approach to the analysis taken.

Opportunities for improving LH-OAT and RASP-SU

The previous versions of LH-OAT and RASP-SU included various sources of uncertainty. These are highlighted in Table A2.3 below together with an indication of the effort required to implement additional aspects.

Issue	**RASP-SU**	**LH-OAT**
Overview		
Acronym	Risk Assessment for Strategic Planning - Structured uncertainty	Latin-hypercube - One at a time
Basic description	Sub-divides the modelling process into a series of stages and applies a staged variance-based sensitivity analysis.	Individually perturbs the inputs to the RASP model and records variation in outputs.
Sampling	Monte-Carlo	Orthogonal Latin Hypercube
Control over convergence rules	Yes	Yes
Runtime	Long (note: depends on number of variables and convergence parameters)	Long (note: depends on number of variables and convergence parameters)
Control over input distributions and	Yes	Yes

Issue	RASP-SU	LH-OAT
parameters		
Aspects included		
Data uncertainties	Yes	Yes
Model component uncertainties	Yes	No
Joint uncertainties	Yes	No
Model structure uncertainties	No	No
Model requirements	Existing NaFRA 08 dataset Input data distributions Input model performance distributions	Existing NaFRA 08 dataset and model Input data distributions
Results		
Uncertainty in probability for a given depth or vice versa (by IC)	Yes	Yes
Uncertainty in damages for a given depth (by IZ)	Yes	Yes
Uncertainty in EAD (by IC)	Yes	Yes
Uncertainty on inflow volume (by defence)	Yes	Yes
Attribution of the uncertainty to input data and model components	Yes	No
Ranking of input data by influence on output uncertainty	Yes	Yes
Current implementation	Reconfiguration of NaFRA 08 toolset	Wrapper around NaFRA 08 toolset
Validation	Limited	Limited

Table A2.3 A comparison of RASP-SU and LH-OAT

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
Source terms									
Water Level (in-river)	Fluvial	Data sources vary from NFCDD 1D, 2D models through to generalised JFlow. Is linked to each asset	Normal	n/a	n/a	0 m	Varies by data source and RP		NaFRA 2009 I9 Ground truthing (Tables 5-8)
Water level (tidal)	Tidal	Tidal water level loadings are spatially interpolated between fluvial water levels and coastal loadings. Is linked to each asset					1.50m	It is difficult to identify the source and presumably it has overall high uncertainty. Exception is where 2D model grids or 1D model nodes are available as source data	NaFRA 2009 desk study I9 - Ground Truthing (pg15)
Water level (coastal)	Coastal wave and water level combinations	Varies by JP region only, is not asset-specific					n/a	Should be same approach as for the Fluvial tables	None
Overtopping rates	Coastal Overtopping volumes	Related to asset based on its toe level (beach level) , crest level and structure type (vertical, embankment, beach) and the	Normal				n/a	SD should vary by RP and defence exposure. NCERM defence orientation considered by use of NFCDD defence exposure parameter to differentiate	None

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
		JP region						between defences on open coast and those in more sheltered locations.	
Pathway (asset) terms									
Crest Level		Data sources vary from measured through to estimate from a 'default' SoP	Normal	n/a	n/a	0 m	Varies 0.1-1.2m	SD varies by data source from 'measured' through to estimated from SoP and 'unidentifiable source'.	NFCDD Data Quality Flag
Standard of Protection (SoP)	Indirectly included under Crest Level above	Sourced from NFCDD through to FRMS consequence rating related 'default'						SoP is not directly used by the model, but translated into Crest Level where this is missing or inconsistent with other input data (water levels, ground levels)	
Condition Grade	Drives the defence fragility (probability of breach at loading x) Number of categories is limited to 5.	Mostly sourced from NFCDD overall condition from last inspection. Default CG = 3 for majority of high ground/natural bank/ channel or where missing.	Variable				mostly a max of +/- 1 category SD	Needs revision: 64% chance that the condition grade is correctly assigned. 34% chance +/- 1 in error 2% chance +/- 2 in error These agree well with general experience from NCERM project (expert judgement)	I9 ground truthing:"condition grade CDF's as shown in Figure 25 through to Figure 29 should be used within the NaFRA uncertainty and sensitivity analysis " Figs provide cumulative probability of "Inspector" (NFCDD record)

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
								that there is a relatively small chance that the assigned CG is out by more than 1 category.	agreeing/not agreeing with "Master" The % in this table is based on the data in I9 note from project NaFRA 2009 Data Improvement, Uncertainty and Sensitivity (Halcrow, 2011).
(Asset) Ground Level	Represents the breach scenario crest level	Estimated from the DTM.	Normal			0 m	Varies 0.2 – 0.5m	Variation depends on source of DTM (LiDAR, SAR or unidentifiable)	
Toe level	Relevant for coastal assets only	This is the beach level at its interface with the defence structure, not the toe level of the defence structure itself	Normal			0 m	1 m	For NCERM model the NFCDD parameters re foreshore level, beach volatility, composition and toe dependency of defence structural form were included to obtain a measure of the foreshore characteristics @ each defence as well as nature of the defence itself; these factors were then assessed to determine toe failure probability. Further	NFCDD parameters re foreshore and defence type. NCERM and ABI processes/ data.

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
								sub-division of this category could be considered in the future to allow a range of SDs to be attributed, for egg an embankment on a stable accreting foreshore currently at a high level will have a low SD, whilst a vertical defence on an eroding foreshore currently at a low level will have a high SD.	
RASP Asset Class Type	Drives fragility curve selection – no uncertainty statements found							Derivation of the RASP Asset Class Type is heavily dependent on NFCDD asset detail, use of defaults and input data review overrides. Requires assessment of impact of misclassification in relation to fragility curves (or a group of classes with distinctive fragility curves)	The finding from the sampling study suggests the RASP classification routine in the pre-processing tool is very effective, i.e. 99.73% correct. Hence there is negligible influence of the misclassification of the RASP classification routine, and no further study is required (from project NaFRA 2009 Data Improvement, Uncertainty and

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
									Sensitivity (Halcrow, 2011).
Fragility		Differentiated to RASP Type or Group	Uniform	LB fragility curve	UB fragility curve	BE fragility curve	n/a	Use of Lower and Upper bands discontinued since NaFRA 2005. It is recommended that the fragility is varied between the upper and lower bounds of the fragility curves using a uniform distribution, and that as an extreme case the breach probability is set to zero for some runs. (from project NaFRA 2009 Data Improvement, Uncertainty and Sensitivity (Halcrow, 2011).	RASP 2002/EX5953 Given the findings of Simm et al. (2009), this will show the potential effect of a hypothesised overestimation of breach probability in the standard fragility curves. It will also provide further insight into the relative importance of overflowing and breaching processes
Defence Length									
Overtopping volumes (by return period and origination of defence)	Coastal assets	Currently the certainty is described in narrative: 'good, or less certainty, more uncertainty' by RP for:				n/a	n/a	Peer review (2013): Defence normal revise to normal +/- 15deg, within 90 degrees of mean wave direction and greater than 90 degrees of mean	See Model parameters below From NCERM: defence orientation is considered by use of NFCDD defence exposure parameter to differentiate

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
		Defence normal within 15 degrees of mean wave direction greater than 15 degrees of mean wave direction						wave direction	between defences on open coast and those in more sheltered locations.
Tributary volume			Factor = Power(10,X)			0	0.3	Based on an analysis of peak flows, NaFRA 2009/10 desk truthing has suggested a normal distribution with a SD factorial error of 2."NFCDD is considered to be a source of validated and calibrated ground truthing data....calculated mean difference between NFCDD peak flow estimates and RRC peak flow estimates [...] notable that the NFCDD provides significantly higher estimates of flow than the RRC data,	

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
								but that this difference reduces with event severity" SD 0.3 taken from 90%ile I9 - Ground Truthing (Fig17/Tbl15)	
Model parameters									
Breach width multiplier	Breach width is very lowly ranked in terms of its importance in estimating model outputs		Triangular	0.5*BE	2*BE			BE (soft assets) = 0.2, BE(hard assets) =0.1 EX5953: -50%/+100%) (HR Wallingford, 2009. Multi-channel river investigation, report EX5984) As surrogate indicators of defence fragility, this would tend to suggest that the output uncertainty is dominated by the uncertainty in input data associated with the non-failure mode of the model, i.e., overflow rather than breaching. (from	It is difficult to quantify the model error in estimated volumes (breach and overflow). The uncertainty in overflow duration may be of the order of ±50%. The use of a single weir coefficient of 1.7 may add another error of the order of 40%. Breach width is assigned an essentially arbitrary fraction of the defence length, which may also be considerably biased. Along with other sources, this could

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
								NaFRA 2009 Sensitivity and Uncertainty Analysis – Case Study Summary Report)	easily give a total factorial error2 [in estimated volumes] of at least 2. (from NaFRA 2009 Sensitivity and Uncertainty Analysis – Case Study Summary Report, 2011)
Hydrograph multiplier			Normal			0.55	0.039	Mean value: Discharge is not constant over duration, so a hydrograph multiplier of 0.55 is used to adjust the duration multiplied by the discharge (HR Wallingford, 2009. Multi-channel river investigation, report EX5984)	
Over flow duration			Normal					The uncertainty in overflow duration may be of the order of ±50% (from NaFRA 2009 Sensitivity and Uncertainty Analysis – Case Study Summary Report)	
Weir equation co-efficient			Uniform distribution	1.0	1.7			Current good practice suggests use of weir	

[2] Factorial error 2 is equivalent to errors in the range +100% (i.e. multiplied by two) to -50% (i.e. divided by two).

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
								coefficients in the range of 1.7 – 1.0, depending on the state of the defence, vegetation etc – for flow over natural ground, a weir coefficient lower than 1.0 may be applicable (see ISIS User Guide). (from NaFRA 2009 Sensitivity and Uncertainty Analysis – Case Study Summary Report)	
Spreading of inflows									
RFSM - Error (depth) by Impact Zone			Normal			0m	0.33m	Error in RFSM depths (conditional on correct inflow volumes) will be of the order of ±0.5m. This is likely to be lower than the error in breach/overflow volumes (from NaFRA 2009 Sensitivity and Uncertainty Analysis – Case Study Summary Report) Further calibration of the friction and momentum terms	HR Wallingford 2007A. NaFRA 2007 Development WP4 Flood Spreading, report EX5490 - explains the previous parametric method (used in NaFRA 06) and RFSM, and compares RFSM with TUFLOW models for test sites on the Thames, in Boston and Carlisle Further testing for Carlisle and a test site on the river Lee

Input data	Existing statements on uncertainty							Rationale	Supporting evidence
	Comment	Data source	Distribution type	Lower bound	Upper bound	Mean	SD(σ)		
								may be required to provide increased confidence in its applicability across all UK catchments AND ...does not include pathway flooding terms... (SC090008 Tech Report, pg 14)	in London are reported in HR Wallingford, 2007B. NaFRA 2007 Main Phase RFSM flood spreading. Interim Report EX5677

Table A2.4 Existing evidence on data and model component

Input Data	Implemented in previous runs		Difficulty of future inclusion / change		
	RASP-SU	LH-OAT	Easily	More Work	Much more work
Missing properties	✓				✓
Property type				✓	✓
Property ground level	✓				✓
Property threshold – residential and NRP	✓	✓			
Property floor area (m2)	✓				✓
Presence of a basement or not (% of properties within an Impact Zone)	✓				✓
Damage given depth above property floor levels per m2 of floor area	✓		✓		
Defence crest level	✓	✓			
Defence toe level	✓			✓	
Defence ground Level	✓	✓			
Fluvial water level loadings (varying by return period and source)	✓	✓			
Tidal water level loadings (varying by return period and source)			✓		✓
Coastal loading (wave and water level combinations)			✓		✓
Breach width multiplier	✓	✓			
Hydrograph multiplier	✓	✓			
Over flow duration	✓				
Weir equation co-efficient	✓				
Defence fragility	✓			✓	
Condition grade		✓	✓		
Overtopping					✓✓

Input Data	Implemented in previous runs		Difficulty of future inclusion / change		
	RASP-SU	LH-OAT	Easily	More Work	Much more work
volumes (by return period and origination of defence)					
Tributary volume		✓	✓		
Volume capping			✓		
RFSM - Error (depth) by Impact Zone				✓ ✓ (level versus volume curves)	✓ ✓ ✓ (pathways)
Defence system state	✓				
Other comments	LH-OAT - Other parameters and assessments can be added when access to the RASP engine low-level function calls are available				

Table A2.5 Input uncertainties - current implementation and future development potential (key: inclusion in RASP-SU ✓; inclusion in LH-OAT ✓)

List of abbreviations

CI	Confidence Indexing (i.e. the approach implemented within NaFRA/MDSF2 2012/13)
DQS	Data Quality Score
DTM	Digital Terrain Model
GEP	Gross error potential - Uncertainty arising from a fundamental mismatch between the model representation of flood system and reality. The influence of this mismatch on the confidence in the results cannot be credibly described in probabilistic terms.
Ground truthing	A process of comparing different data sources to 'known' best of breed estimates available at the same location.
HLCF	High Level Conceptual Framework
IC	Impact Cell
IZ	Impact Zone
LH-OAT	Latin Hypercube-One-At-a-Time.
LIDAR	Light Detection and Ranging
MDSF2	Modelling Decision Support Framework 2 (the toolset that enacts the RASP methods).
NaFRA	National Flood Risk Assessment.
QKU	Quantifiable knowledge uncertainty - Uncertainty arising from errors in the input data or model performance that can be credibility described in probabilistic terms.
RASP	Risk Assessment for Strategic Planning (methods underlying MDSF2 and NAFRA).
RASP-HML+	RASP-High Level Method plus
RASP-SU	RASP-Structured Uncertainty
RFSM	Rapid Flood Spreading Method
SAMP	System Asset Management Plans

Would you like to find out more about us, or about your environment?

Then call us on

08708 506 506* (Mon-Fri 8-6)

email

enquiries@environment-agency.gov.uk

or visit our website

www.environment-agency.gov.uk

incident hotline 0800 80 70 60 (24hrs)
floodline 0845 988 1188

* Approximate call costs: 8p plus 6p per minute (standard landline).
Please note charges will vary across telephone providers

Environment first: This publication is printed on recycled paper.