delivering benefits through evidence

Developing a prototype tool for mapping flooding from all sources Phase 1: Scoping and conceptual method development

Project: SC080050/R1

Flood and Coastal Erosion Risk Management Research and Development Programme

The Environment Agency is the leading public body protecting and improving the environment in England and Wales.

It's our job to make sure that air, land and water are looked after by everyone in today's society, so that tomorrow's generations inherit a cleaner, healthier world.

Our work includes tackling flooding and pollution incidents, reducing industry's impacts on the environment, cleaning up rivers, coastal waters and contaminated land, and improving wildlife habitats.

This report is the result of research commissioned by the Environment Agency's Evidence Directorate and funded by the joint Environment Agency/Defra Flood and Coastal Erosion Risk Management Research and Development Programme.

Published by:
Environment Agency, Rio House, Waterside Drive, Aztec West, Almondsbury, Bristol, BS32 4UD
Tel: 01454 624400 Fax: 01454 624409
www.environment-agency.gov.uk

ISBN: 978-1-84911-186-7

Email: fcerm.evidence@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.

Author(s):
Matt Horritt, Luke Lovell, Jon Wicks (Halcrow)
Steve Buss (ESI)

Dissemination Status:
Released to all regions
Publicly available

Keywords:
Flood mapping, all sources

Research Contractor:
Jon Wicks
Halcrow Group Ltd
Burderop Park
Swindon
SN4 0QD
Tel: 01793 812479

Environment Agency's Project Manager:
Jim Walker, Evidence Directorate

Theme Manager:
Suresh Surendran, Modelling and Risk Theme

Collaborator(s):
ESI

Project Number:
SC080050

Product Code:
SCHO0310BSGQ-E-P

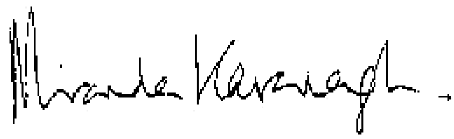
Evidence at the Environment Agency

Evidence underpins the work of the Environment Agency. It provides an up-to-date understanding of the world about us, helps us to develop tools and techniques to monitor and manage our environment as efficiently and effectively as possible. It also helps us to understand how the environment is changing and to identify what the future pressures may be.

The work of the Environment Agency's Evidence Directorate is a key ingredient in the partnership between research, policy and operations that enables the Environment Agency to protect and restore our environment.

The Research & Innovation programme focuses on four main areas of activity:

- **Setting the agenda**, by informing our evidence-based policies, advisory and regulatory roles;
- **Maintaining scientific credibility**, by ensuring that our programmes and projects are fit for purpose and executed according to international standards;
- **Carrying out research**, either by contracting it out to research organisations and consultancies or by doing it ourselves;
- **Delivering information, advice, tools and techniques**, by making appropriate products available to our policy and operations staff.



Miranda Kavanagh

Director of Evidence

Executive summary

This report from Phase 1 of the 'Developing a prototype tool for mapping flooding from all sources' Science project covers the project requirements, consultation and literature review, review of sources of flooding, method development and prototype software tool specification. The key findings are:

- The EU Floods Directive and the Pitt Review are seen as the key project drivers, in particular facilitating delivery of the Environment Agency's expected responsibilities in (i) flood mapping at a national scale, and (ii) provision of tools, techniques and guidance to help local authorities with local flood mapping (including 'combined consequences' mapping).
- All sources of flooding have been considered by this project. Existing datasets for fluvial (where mapped by the Environment Agency), coastal, estuarine and surface water sources are currently suitable for inclusion in a map of all sources of flooding.
- The current status of datasets for the following sources is such that inclusion in an 'all sources' map is unlikely to be feasible: reservoir breach, some minor fluvial watercourses, groundwater, canal breach and overtopping, sewer failure and capacity exceedance, point infrastructure failure (pumps, barriers etc.), tsunami, and water supply infrastructure failure (pipe bursts). Either the necessary datasets, or the underlying science required to produce them, are not available for these sources.
- A method is proposed that is able to generate a probabilistic flood map presenting a measure of the probability of flooding caused by the included sources. The method focuses on the use of pre-existing flood map data, which are combined to provide 'all sources' flood mapping data (rather than an approach in which inputs of water from multiple sources are routed along selected pathways as a fundamental part of the method). The proposed method is therefore reliant on the availability of suitable pre-calculated flood inundation data from individual or previously combined sources. However, the method will identify situations in which new, local integrated (i.e. including flooding from more than one source) modelling is required. The method is generic and can be used for national-scale and local-scale mapping. The method is consistent with the RASP concepts of a system-based, risk-based hierarchical approach and is compatible with the RASP-related products such as NaFRA and MDSF2.
- The method is not constrained to only include currently available sources – any source of flooding can be included where/when suitable data exist.
- A prototype software tool is specified for implementing the method and will be coded in Phase 2.
- The Phase 2 project deliverable will not be 'finished' software suitable for direct operational use, nor will the project be generating new flood maps. Thus, further work will be required before the benefits of the research are fully realised.

Contents

1	Introduction	1
2	Requirements	3
2.1	Project drivers	3
2.2	Ultimate outcomes facilitated by the project	3
2.3	Requirements	4
2.4	Limitations	5
3	Approach to the project	7
4	Review of relevant projects and reports	10
5	Initial consultation	11
5.1	Sources	11
5.2	Methodological approach	11
5.3	Current data sources	12
5.4	Future data sources	13
5.5	Key challenges	14
5.6	Communication	15
5.7	Relative risks from all sources	15
5.8	Project workshop	16
6	Sources of flooding	17
6.1	Review of sources	17
6.2	Priority sources	19
7	Datasets	22
7.1	Dataset details	22
7.2	Gaps	31
7.3	Summary of suitability	31
7.4	Suggestions for overcoming data issues	33
8	Proposed method	34
8.1	Introduction	34
8.2	Proposed method	37
8.3	Representing source interaction and dependency	44
9	High level design andinitial piloting	53
9.1	User requirements	53
9.2	Mock-up demonstrations	54
9.3	Initial piloting	57

10 Specification for prototype software 78
10.1 Non-functional requirements 78
10.2 Functional requirements 79
10.3 Technical specification 81
10.4 Specification for the prototype software tool 86
10.5 Reuse of MDSF 2 code 89
10.6 IPR needs 89

11 Benefits realisation 90
11.1 Introduction 90
11.2 Alignment with high level Science outcomes 90
11.3 Actions to realise the benefits 90
11.4 Links with other Science projects 91

12 Conclusions and recommendations 93
12.1 Recommendations for phase 2 94
12.2 Recommendations for future Environment Agency science 95

Abbreviations 97

References 98

Appendices 99
Appendix A: Literature review 99
Appendix B: Consultation questionnaire responses 109
Appendix C: Internal and external consultation list 130
Appendix D: Terminology guide 131
Appendix E: Question log 135
Appendix F: Groundwater flood risk assessment 136
Appendix G: Proceedings of stakeholders' workshop 145
Appendix H: Application of FD2121 guidance 154
Appendix J: Pseudocode 161

1 Introduction

Halcrow has been commissioned by the Environment Agency to develop a prototype tool for mapping flooding from all sources. The purpose of the tool is to combine information on flooding from different sources (e.g. river, coastal, surface water etc.) into a single map communicating the probability of flooding. The main tasks of this project are split into two phases:

Phase 1:

1. Review of data, knowledge, methods and models, and initial stakeholder consultation.
2. Summarise input datasets identified under the first task, recording key assumptions, recommendations and intellectual property rights (IPR).
3. Outline challenges of developing a consistent probabilistic method.
4. Conceptualise modular methodology/framework suitable for combining and analysing the probability of flooding from a range of sources.
5. Develop functional design and mock-up demonstrations of the conceptual framework.
6. Develop specification for a prototype tool to assess and map probabilities of flooding from all sources.

Phase 2:

1. Finalise functional design, develop and alpha test prototype software tool. Write user guide.
2. Pilot testing and user acceptance testing.
3. Update prototype following testing.
4. Develop implementation plan.

This report covers Phase 1 only and is aimed at the following stakeholders:

- The Environment Agency project board for this Science project.
- Technical experts and other key stakeholders who need to understand the proposed method.

The report is structured as follows:

- requirements of the project (Chapter 2);
- approach adopted by the project team in delivering the project (Chapter 3);
- review of relevant previous projects and literature (Chapter 4);
- initial consultation process feedback (Chapter 5);
- list of flooding sources to be included in the project scope (Chapter 6);
- review of available datasets (Chapter 7);
- proposed calculation method (Chapter 8);
- high level design illustrated by initial pilot testing (Chapter 9);
- requirements specification for the prototype software tool (Chapter 10);

- consideration of benefits realisation (Chapter 11);
- conclusions (Chapter 12).

Appendix D provides a guide to the following terminology used in this report: natural variability, knowledge uncertainty, deterministic, probabilistic, risk-based, systems-based, hierarchical framework and RASP.

2 Requirements

2.1 Project drivers

The main drivers for this project are:

- The Environment Agency's need to respond to the challenges of the Government's **'Making Space for Water'** and **'Future Water'** strategies which seek a more integrated approach to addressing all sources of flooding and managing a wider range of flooding systems.
- **PPS25** is encouraging Local Planning Authorities (LPAs) to consider all sources of flooding including flooding from rivers, sea, overland, sewers, groundwater and flooding from reservoirs, canals and other artificial sources. (In Wales, **TAN 15** encourages a full consideration of the likely sources of flooding.)
- The European Directive on the Assessment and Management of Flood Risks (the EU **Floods Directive)** indicates that European member states should map flooding from all significant sources of risk by 22 December 2013. The Directive defines 'flood' to mean 'the temporary covering by water of land not normally covered by water'; this terminology is inclusive of all sources of risk, and the Directive only specifically provides for the potential exclusion of floods from sewerage systems.
- The **Pitt Review** recommends that the range of flood risks which are mapped is increased and stresses the need to forecast, model and warn against all sources of flooding.
- The draft **Floods and Water Management Bill**, which aims to create a more comprehensive and risk-based regime for managing the risk of flooding embracing all sources of flooding.

The requirements of the Floods Directive and the Pitt Review are seen as the most important project drivers. The Environment Agency's key strategic role is to provide an overview of risk and communicate the risk effectively. The communication of risk is the chief outcome for the flood mapping from all sources tool.

There are an increasing number of sources of flooding being mapped, most recently with the commission of maps of areas susceptible to surface water flooding, and with reservoir inundation mapping under way. It is a challenge to effectively communicate, to a range of stakeholders, the flood risk from all these sources – combining the data is one approach to meeting this challenge. The drivers all point to taking a holistic view of flooding, where different sources are integrated, rather than segregated due to their mechanism. This is better aligned with the needs of many stakeholders, who are not directly concerned with the source of flooding itself, but the risk of flooding from any source.

2.2 Ultimate outcomes facilitated by the project

The project will contribute to the overall goal of communicating flood information (such as likelihood, extent, depth, velocities) from different sources in an integrated and consistent fashion. This will be of benefit to a broad range of flood risk professionals, decision makers and the public. An integrated map will benefit a number of elements in flood risk management programmes from national to local level (Table 2.1). (Note that information on which sources of flooding contribute at specific locations will still be available).

Table 2.1 Benefits of mapping flooding from all sources.

Flood risk management activity	Benefit
National-scale flood risk assessment	An integrated map will promote a move towards an assessment of total flood risk, rather than risk from just fluvial and coastal sources.
Broad-scale flood risk assessment, e.g. CFMPs and SMPs	Provide better understanding of flood risk at the catchment scale.
Flood forecasting and warning	An integrated map will help direct forecasting and warning programmes towards the most significant sources and help them to convey information about combined risks in simple terms.
Flood incident response	An understanding of the effect of different sources at specific locations will allow improved emergency response from government agencies, utilities and the public.
Communicating flood risk to the public	An integrated map will improve the public's understanding of flood risk and allow the risk from many sources to be communicated concisely and effectively. It could also help people to respond appropriately to reduce the risk of flooding. It may facilitate new methods of communicating risk based on a common understanding of the risks from different sources.
Planning and development control	An integrated map will give planners a holistic overview of flood risk in an area (and thus could contribute to improved spatial decision making).
Developing responses to flood risk	A view of flooding from all sources will allow integrated responses to be developed, and conflicting responses to be reconciled.
Performance management	Performance of flood risk management organisations is dependent on addressing flooding from all sources, and this can be better assessed if total risk is known.

The output of this project will contribute to the ultimate outcomes described above by developing and prototyping a method of integrating the currently disparate sources of flood information. While the tool contributes to the flood probability component of risk rather than the consequences, in combination with receptor data it will give an improved understanding of risk.

2.3 Requirements

In order to realise the benefits described above, a number of requirements of the tool can be defined. The essential high level requirements are:

- The tool should be capable of integrating data of different formats that arise when different sources of flooding are modelled.
- The method should be generic and hence flexible enough to be used on datasets that may arise during the lifetime of the project (i.e. Phases 1 and 2 running through 2009–2010). The method should also allow for the easy

future upgrading of the flood mapping data for specific sources used to generate combined maps. This is also relevant to updating flood information estimates as the climate, land use and flood risk management (FRM) responses change.

- Dependency between sources should be considered and represented. Where joint probability information is available, the method should be able to use it to improve the estimates of flood information in the combined map. An understanding of dependency will be required for sources linked by a strong causal relationship (e.g. storm surges and rainfall events), and for those with a weaker relationship through seasonality (e.g. fluvial flows and groundwater).
- The prototype tool should be acceptable to its end users as software, having been tested and trialled for usability and compatibility.
- The output of the tool must be in such a format as to allow the end users to realise the benefits listed above. The output should simplify and promote, rather than complicate, the assessment of risk from all sources.
- The method must include ways of dealing with our incomplete knowledge of flood probabilities presented by the input data, and generate some form of uncertain probabilistic output or indicator of uncertainty. Quantifying uncertainty will allow the tool to communicate confidence in the outputs.

These requirements are regarded as essential by the project specification.

Optional features are:

- The tool may allow some local, specialised data to be included (e.g. local depth–probability grids). Although the tool would not restrict subsequent use of the mapping (e.g. at an individual property scale), the scale at which the outputs will be suitable for use will depend on the quality of *all* input data for that geographical location.
- The output of the tool may provide an indication of the proportional contribution of different sources to total flood hazard.
- The tool may provide explicit information on flood risk response to climate change and facilitate the inclusion of land use changes and FRM responses, allowing users to view current and future risk.

Chapter 10 provides a full specification of the essential and optional features proposed for Phase 2 of the project.

This project focuses primarily on nationally available datasets. In some cases, this means a single dataset covering the whole of England and Wales (e.g. the map of areas susceptible to surface water flooding). For other sources, data may be spread across several datasets, but still available in a common format and containing standard information.

2.4 Limitations

Having outlined the high level requirements, it is useful to highlight what this project will not deliver:

- The project will not produce any new flood mapping products, relying on previously generated data.
- The project will not develop any new information on joint probabilities.

- The project will not deliver a tool specifically intended for use at the scale of individual properties. Although the method itself is not restricted to use at specific scales, there may be resolution and other constraints that impede the tool's use at small scales.
- The project will not deliver a finished software product, but a prototype that will allow the Environment Agency to perform a preliminary evaluation of the tool and to act as a proof of concept (these aspects will be assessed in more detail in Phase 2).

3 Approach to the project

The project approach is to undertake a targeted consultation process, and to use this to inform the development of a methodology for combining flood sources and maps. Along with a functional design encapsulating the needs of end users, this methodology will be implemented as a prototype tool.

The project tasks (split into Phases 1 and 2) are now described in more detail.

Task 1.1: **Review of data, knowledge, methods and models, and initial stakeholder consultation.**

The review and initial consultation included:

- Comprehensive review of relevant literature/reports, with results presented in Chapter 4 of this report.
- Consultation with experts on sources of flood data available now and in the near/medium-term future, challenges in developing the method, and desirable outputs from the tool. This consultation was in the form of a questionnaire followed up by face-to-face meetings and teleconferences where necessary. The responses to this questionnaire are included in Appendix B.

The aim of this consultation was to define better the scope of the project, to assess likely methods, and to develop a list of sources to be included in the project. The list of sources to be included in the project is included in Chapter 6.

The source list should not be viewed as limiting the data that can be included by the tool, as data from other sources can also be included. If these new data sources do not share a common format with one of the sources already included, modification of the tool may be required (which lies beyond the scope of this project).

A further round of consultation took place at a stakeholder workshop in July 2009, once the draft methodology had been developed (see Section 5.8). Leading academics, consultants, and Environment Agency staff and other operating authority representatives attended.

Task 1.2: **Summarise input datasets identified under Task 1.1, recording key assumptions, recommendations and IPR.**

Datasets representing the sources defined in Task 1.1 are summarised in a record of their properties, IPR status and fitness for purpose (Chapter 7). This allowed the differences and similarities between the datasets to be identified (important in developing the tool's core methods), and any gaps to be discovered.

Task 1.3: **Outline challenges of developing a consistent probabilistic method.**

The findings of Tasks 1.1 and 1.2 were reviewed and drawn together into a list of key challenges (Section 5.5). These fed into task 1.4 to identify the best method.

Task 1.4: **Conceptualise modular methodology/framework suitable for combining and analysing the probability of flooding from a range of sources.**

In order to deliver on this requirement we scoped a range of methods which range from a fully integrated approach through to development of a 'map combination' approach. The two end members of this range are:

Fully integrated approach: The fully integrated approach refers to a method in which sources of potential flood water are input into a simulation engine which routes potential combinations of sources through pathways, calculating the resulting flood maps within, for example, a Monte Carlo framework. This method allows the interaction between water from different sources to be explicitly represented hydraulically, and requires an understanding of event duration, coincidence and dependency.

Map combination approach: The map combination approach will use Bayesian methods to combine probabilities from different sources and information on their dependencies, where available. When no dependency information is available, bounding assumptions will be made, and the assumptions flagged in metadata associated with the combined probability output map. This method will not represent interaction between water from different sources, but combines probabilities.

The essential properties of these two methods are listed in Table 3.1. The chief difference is how pathways are modelled. In the fully integrated approach the sources are combined, and then routed along pathways to the risk receptors. In the map combination approach, the water from different sources is routed separately, and the resulting probabilistic flood maps combined. The optimum solution may lie in some combination of these methods.

As an example of applying the two approaches, consider mapping flooding from fluvial and surface water sources:

The **fully integrated approach** would first specify a number of design events, corresponding to different magnitudes of fluvial flow and local rainfall. These would be assigned a probability through analysis of dependency information. These events would then be used as inputs to a single hydraulic model, for example a linked 1D–2D model representing the channel–floodplain system, with rainfall added directly as a source to the 2D model. The model would then route the combined water across the floodplain to receptors.

The **map combination approach** instead would take the outputs from separate models of fluvial flooding and of surface water, which give separate estimates of flood probability for each floodplain cell. These probabilities would then be combined to give a single estimate of probability of depth exceeding some threshold, from either or both of the sources.

Table 3.1 Typical properties of fully integrated and map combination approaches.

Fully integrated approach	Map combination approach
Combines water at source	Combines probabilities at receptors
Represents some hydraulics of interaction between water from different sources	Interaction represented heuristically
Same treatment of routing of water from different sources	Different routing methods used for different sources
Monte Carlo approach to combining probabilities	Analytical Bayesian approach to combining probabilities
Sources treated at same (spatial, temporal) scale	Sources treated at different (spatial, temporal) scales
Routing calculation performed after sources combined	Flood maps combined directly

The final probabilistic map of multiple sources of flooding may need to be communicated to the end user in new ways. Further information (assigning proportions of the risk to different sources, most significant source, velocity etc.) may also need to be communicated in the same or supplementary maps. The need for clear communication of the results may affect the method chosen.

Chapter 8 describes the proposed method.

Task 1.5: **Develop functional design and mock-up demonstrations of the conceptual framework.**

The aim of the functional design element of this task is to develop the methodology described in preceding tasks into a high level design for a tool, focusing on how it will be used as well as how it will work. The mock-up will be used to demonstrate the 'look and feel' of the tool. Chapter 9 describes the high level design and initial piloting.

Task 1.6: **Develop specification for a prototype tool to assess and map probabilities of flooding from all sources**.

This task takes the functional design from the preceding task and uses it to develop a specification of the software tool to be developed in Phase 2 – see Chapter 10 for details.

Phase 2 (not covered by this report) will consist of the following tasks:

- Task 2.1: Finalise functional design, develop and alpha test prototype software tool. Write user guide.
- Task 2.2: Pilot testing and user acceptance testing.
- Task 2.3: Update prototype following testing.
- Task 2.4: Develop implementation plan.

4 Review of relevant projects and reports

Relevant project reports and other literature have been reviewed by project team experts in a broad range of disciplines. The results and their relevance for this project are summarised in Appendix A.

The key points from this literature survey are:

- Uncertainty is an emerging theme in academic and applied research. Estimation and communication of uncertainty in mapping of flooding from multiple sources may be required in the prototype tool.
- Climate change drivers mean that risk now and in the future is being assessed by a number of projects. The prototype tool should be able to ‘integrate’ flood information from multiple sources for consistent climate change scenarios (e.g. mapping data predicted for a consistent medium–high climate change scenario at the year 2080). Less detailed climate sensitivity inputs may need to be incorporated, for example by taking current scenarios and adjusting associated probabilities to represent increased likelihood of extreme events in the future.
- Techniques exist for estimating and using the coherence/dependence between flood sources. These methods are not in widespread use in flood mapping projects.
- Information on the relative magnitudes of risk from different sources is scarce.
- Some techniques (e.g. from TE2100 study, NaFRA) have been developed for combining probabilistic flood maps, and the ‘Spatial coherence of flood risk’ project has looked into spatial aspects of probabilistic flood mapping.
- Flooding due to combinations of factors (e.g. sewer capacity reduced due to high river levels) may have made a significant contribution to damages in the summer 2007 floods, but no figures exist to quantify this contribution.
- Blockages of sewers and point infrastructure may be an important contributor to flood risk, but the increased risk due to such asset failures has not been quantified.
- Some existing datasets (e.g. NaFRA) are well suited as part of a combined probabilistic flood map; others (high level vulnerability methods for groundwater, maps of areas susceptible to surface water flooding) will be more difficult to interpret and integrate.
- Further information on surface water flooding and inundation from reservoir failure will be forthcoming in 2009–2010.

5 Initial consultation

The initial consultation with experts involved sending a questionnaire with follow-up discussions where necessary. The list of initial consultees is shown in Appendix C which also notes where completed questionnaires have been returned. Appendix B contains the questionnaires received.

The consultation has raised a number of issues, guided by the themes in the questionnaire. These are now described according to those themes. Information from the literature survey has also been included.

5.1 Sources

As a list of sources identified for potential inclusion in the prototyping method was one of the key aspects of the consultation, the finds are included in Chapter 6.

5.2 Methodological approach

The 'end members' of the methodological approaches identified were (i) the fully integrated approach, with water volumes from different sources being combined and then routed over the floodplain, and (ii) a higher level method, where flood maps and other model outputs are combined probabilistically.

The consultees have generally considered the higher level method more appropriate to this project, for the following reasons:

- Information from the most appropriate model (whether part of the RASP family or not) can be combined. Different pathways may be appropriate to different sources (e.g. the pathways for surface water and reservoir flooding may be very different).
- Historical data can be included.
- It may be possible to include modelled velocities.
- The fully integrated approach is better for looking at management options (e.g. the effectiveness of defences) where a fuller description of flood probability may be required. This may be overcomplicated for an overall holistic view of flood hazard.
- Use of the fully integrated approach will complicate flood map delivery, development of the prototype and ongoing development of other RASP tools.
- The method is more feasible for short-term delivery in this project.

However, there are some advantages of the fully integrated approach:

- Pathway data can be reused for some sources (e.g. surface water and sewer flooding).
- The fully integrated approach may be the only way of dealing with non-coastal/fluvial flooding that is strongly influenced by artificial pathways (e.g. sewer flooding).

- The fully integrated approach will give more accurate and credible results where water from different sources interacts.

The responses indicate therefore that, while the fully integrated approach may be more suitable/necessary in special circumstances, for the scale and accuracy required of the tool the higher level method is more appropriate. The fully integrated approach would represent a considerable cost in computational and data requirements, which would not result in significantly improved modelling in the majority of locations. The method we propose combines the fully integrated and higher level approaches, and the rationale for this method is discussed in further detail in Chapter 8.

5.3 Current data sources

Data in existence now are described below, classified as either historical or modelled.

Historical data can in some respects be considered as more reliable than modelled data, as it represents a ‘real’ flood event, but only for the events that happen to have occurred and been observed. Historical data must also be associated with some probability to be of use in assessing flood risk, and for extreme events this may be difficult. It also only tells us about examples of flood hazard in the past, and nothing about future scenarios. Potential sources include:

- Point levels vs. probability/return period, e.g. fluvial, coastal water levels. These typically rely on some degree of extrapolation to long return periods to describe extreme events.
- Point levels of historical floods, e.g. from wrack marks.
- Point flooding incidents where depth exceeded a small threshold, but with no associated levels.
- Historical flood extents, e.g. from aerial photography.

Modelling is the primary source of flood information. Potential sources are:

- The Environment Agency Flood Zones for fluvial and coastal sources. These are in the form of depths or wet/dry state for probabilities 1%/0.5% (for fluvial and coastal sources) and 0.1%. There are also more detailed model outputs available (e.g. from mapping studies and strategies) that can include depths and velocities and other probabilities.
- Reservoir inundation maps commissioned by owners. These are likely to be in the form of flood depths, and possibly velocity and other hazard indicators, conditional on the dam failure. Some sensitivity analysis for different breach hydrographs, floodplain roughnesses and other parameters may also be included.
- Sewer flooding data.
- NaFRA probabilities of flooding (and depth–probability data).
- The Environment Agency’s map of areas susceptible to surface water flooding. For the ‘first generation’ outputs this is in the form of a map with three bands (labelled ‘less’ to ‘more’) – for the ‘second generation’ depths and different probabilities may be provided.
- Groundwater flooding susceptibility maps.
- Canal breaching data.

These sources may be subject to a number of restrictions. Sewer exceedance hazard maps commissioned by water companies may exist, but may be difficult or impossible to access for a publicly available flood map (without legal intervention). Reservoir inundation maps are sensitive due to the potential of dams as terrorist and military targets. Although it is expected that dissemination of the raw reservoir inundation data will be restricted, it is likely that the restriction will be removed once the data have been combined with other sources of flooding and/or presented at a broader scale.

Further analysis of current datasets is provided in Chapter 7.

5.4 Future data sources

The data sources listed above are expected to be augmented in the near to mid-term (within 5 years) by:

- Further information on climate change and updated future scenarios. These may occur incrementally as recent sea level data are integrated with past records, or as step changes in our understanding of future climate impacts (e.g. UKCP09).
- Further flood event records being compiled now or for flood events in the future.
- Improved topography data, e.g. improved accuracy and resolution of Environment Agency LiDAR. This may need to be combined with modelling outputs produced using previous versions of the topographic datasets.
- Outputs from future strategy studies and flood mapping exercises. The new Flood Mapping, Modelling and Data strategies describe the plans for new products including hazard mapping for significant risk areas to meet Floods Directive requirements.
- Surface Water Management Plans in many towns (six local authorities in the first phase), which may include improved sewer data.
- Water Resources Act 2003 obliges owners of 'large raised reservoirs' to produce flood plans. The Environment Agency has also commissioned a first phase of reservoir inundation maps for delivery in 2009.
- Coastal flooding information may be updated as coastal morphology changes.
- Drain London Scoping Study (MWH for Greater London Assembly).
- The Atlantis initiative, which aims to provide integrated geographic and environmental datasets to better support water management in flooding and water quality for the 21st century.
- The INSPIRE Directive, which will establish an infrastructure for spatial data management.
- Outputs from application of the methods being developed in Environment Agency Science projects such as 'Method for Local Probabilistic Flood Risk Assessments' (SC090008/WP2) and 'Developing the Next Generation of Surface Water Flood Risk Assessment' (SC070059).
- Canals will be treated as ordinary watercourses or main rivers, with a responsibility for mapping risk expected to lie with British Waterways.

These changes may take a considerable time to ‘filter through’ to updated flood information. New estimates of climate change impacts, for example, may require further modelling studies/updating of existing models before updated flood maps can be produced.

5.5 Key challenges

A broad range of challenges have been identified by the consultees:

- Managing the large volumes of data will stretch storage and processing capacity. The prototype tool should address potential run times/storage space issues.
- There is a potential conflation of ‘sources’ as discussed in the proposal and tender documents, and ‘natural drivers’. When talking about combining different sources (e.g. surface water and fluvial flooding), these may be a response to the same driver (local convective rainfall). Treating them separately and then combining may miss important dependencies.
- Gaining access to the types of data described above, owned by a number of bodies, national and local, public and private etc. There may be political, data protection, intellectual property, commercial and security issues associated with these datasets.
- Understanding the provenance of different data sources due to different (or non-existent) metadata standards.
- Combining flood maps may not represent the physical processes that occur when flood waters from different sources interact within pathways and at receptors. Combining sources and combining probabilities are not the same, and there may be locations where the combination of sources affects flood risk significantly.
- Representing velocities and hazards to people.
- Representing joint probabilities and spatial coherence.
- Producing a credible overall assessment of flood risk. Will combining information from different flood sources degrade that information to such an extent that the final product is worse than the sum of its parts? How do we communicate this so that the users have an appropriate confidence in the maps?
- How do we test the efficacy of combined maps? Are there datasets that can be used to validate all sources mapping?
- Do we need or want to include flood duration effects? Timescales may be very different (e.g. for groundwater and surface water flooding), so how can we model both? Can this be represented by the same approach as for flood dependency?
- Subsurface flow processes associated with surface and sewer flooding are difficult to represent because of a lack of data on asset performance.
- Combination of sources with information at different spatial scales, and determining an appropriate spatial scale for the output.
- Combination of sources where our knowledge and modelling skill differ widely (e.g. groundwater modelling is less advanced than fluvial modelling).

- Can we simplify the approach by only including *significant* sources of flooding, rather than *all* sources? How do we assess what is significant?
- How do we match our method of communicating of the maps to a wide variety of users? This is outside the scope of this project, but the method should produce outputs that are capable of being communicated clearly.
- Do we need to estimate and communicate the uncertainty in a combined map?
- Do we need to include future flood hazard?
- What is the best (efficient, flexible, transparent) computational architecture for the tool?

Challenges that may be addressed by further Science projects are highlighted in Chapter 12.

5.6 Communication

The consultees have suggested the following methods of communicating a map of combined flood sources:

- Use Defra 'hazards to people' guidance to inform depth thresholds, and estimate the probability for these.
- Probability of flood depths >0 m.
- Type of flooding may need to be communicated.
- Uncertainty – e.g. combining high resolution/accuracy fluvial mapping with groundwater flood susceptibility method.

In anticipation of the questions that may be asked of the project and its outputs, a list of expected questions is being developed. The draft form of this is shown in Appendix E and will be reviewed during Phase 2.

5.7 Relative risks from all sources

There is currently no national picture of the relative magnitudes of risk from different flood sources. However, using some very crude assumptions, we can make some estimates of the relative magnitudes for some sources.

Fluvial and coastal risk: The NaFRA 2008 analyses estimate the expected annual damages from rivers and the sea to be £1.2 billion in England and Wales, with 2.6 million properties with an annual probability of flooding greater than 0.1%. This gives an annual expected damage per property at risk of £500. The expected number of properties flooded per year is 20,000, equivalent to a damage to each of £60,000. These figures represent the risk with defences in place.

Reservoir failure: The Environment Agency is responsible as enforcement authority for some 2,000 reservoirs. Of these, 700 are category A, where more than 10 people are affected. If we assume that each category A reservoir puts, on average, 20 properties at risk, 14,000 properties are at risk. If we estimate reservoir failure probability at 0.01% (i.e. reservoirs are designed for a flood with a 1 in 10,000 chance of occurring in any given year), this produces an annual expected number of properties flooded from reservoir failure at approximately 1. Expected annual damages are therefore £60,000 nationally.

This estimate understates the case, since a reservoir failure inundation will produce more damage (due to hydrodynamic forces and greater water depths) than coastal and fluvial flooding. There is also greater risk to life from a reservoir failure. The 1 in 10,000 probability of failure is an extremely crude estimate based on the flood carrying capacity of the reservoir. This does not include all failure mechanisms (e.g. sunny day failures), and an overtopping event may not cause the dam to fail. Even so, we can estimate risk from reservoir failure as being about four orders of magnitude less than coastal and fluvial flooding.

Surface water: A substantial proportion of the £3 billion damages from the 2007 summer floods have been attributed to surface water. While drawing general conclusions from a single (perhaps exceptional) year, it does show that widespread damage from surface water flooding is possible, as well as the more expected isolated surface water events. This is supported by the figure of £270 million annual average cost of surface water damage in England and Wales quoted in the Foresight Future Flooding report.

Other sources: According to *Flooding from other sources* (*Making Space for Water* – MSfW – report HA4a, JBA 2006), for the autumn 2000 floods 14% of properties were flooded 'due to non-river causes' (which includes surface water), and 40% of the damages occurred outside the indicative coastal/fluvial floodplain. If these flood events are representative, expected annual damages from these other sources are £200–700 million, with 2,000–5,000 properties affected annually. MSfW HA5 estimated that 1.6 million properties in England and Wales may be at risk from groundwater flooding.

This estimate may be underestimating the risk from non-river/coastal sources, since it is limited to events that occurred at roughly the same time as major river flood events. The isolated events mentioned above will also contribute. The evidence does point to surface/other sources contributing significantly (some tens of per cent) to flood risk in England and Wales.

5.8 Project workshop

Subsequent to the initial consultation, a project workshop was help on 14 July 2009. The purpose of the workshop was to:

- raise awareness of the project among specialists and potential end users; and
- facilitate the exchange of ideas on mapping flooding from all sources.

A summary of the workshop is provided in Appendix G.

6 Sources of flooding

6.1 Review of sources

As introduced in Chapter 2, a main objective of this project is to facilitate mapping of flooding from 'all sources' of flooding. Through the consultation and literature review process a list of all sources of flooding has been developed (Tables 6.1 and 6.2). In Table 6.1 the different sources have been assigned to a **priority** and **difficulty** category. Priority indicates whether the inclusion of the flooding source is essential (1), should be included (2), or can be excluded (3) from the scope of this project. The 'difficulty' is a measure of the potential problems/feasibility of including that source (E = easy, M = medium, D = difficult). Note that the 'D' difficultly category does not necessarily imply that it would be difficult to generate the particular data source, rather that suitable data are not expected to be readily available in the near future.

Table 6.1 Main categories of flooding sources.

Source		Priority*	Difficulty*	Issues
Fluvial	Main rivers	**1**	**E**	Different approaches available for same area (local flood model, NaFRA/MDSF2 modelling).
Coastal	Tide/surge	**1**	**E**	
	Wave overtopping	**1**	**M**	May be less accurate than still water levels. Included in some modelling.
Estuarine		**1**	**M**	Joint probabilities.
Surface water	Local rainfall before interception by sewers	**1**	**M**	Map of areas susceptible to surface water flooding not associated with a specific probability. Pluvial flood maps may also pick out fluvial flooding – potential for 'double counting'.
Sewer	Capacity exceedance	**2**	**D**	Capacity poorly understood for different urban catchments.
	Blockage/ failure	**3**	**D**	Probability and consequence poorly understood. The Floods Directive allows this source to be excluded.
Small watercourse	Not covered by fluvial flood maps	**2**	**M**	Catchments too small for flood maps. May be picked up in areas susceptible to surface water flooding and Surface Water Management Plans. Some ordinary watercourses are covered by Environment Agency Flood Zone mapping.
Tsunami	Tectonic	**3**	**D**	Some impacts known, event probability poorly understood.
	Landslide	**3**	**D**	
Groundwater	Rising groundwater levels in Chalk aquifers	**2**	**E**	Several datasets are available for England and Wales. They present indicative outlines, rather than predict depths, return frequencies etc.
	Rising groundwater levels in non-Chalk aquifers	**3**	**M**	BGS have published mapping. Difficulties with comparisons between aquifers.
	Pump failure	**3**	**D**	Affects infrastructure protected by pumping (tunnels, deep basements). Site-specific and depends on future groundwater abstraction scenarios.
	Alluvial aquifers	**2**	**M**	Underground flow paths from fluvial flood events. BGS have published mapping. Difficulties with comparisons between aquifers.
Canal	Breach	**2**	**D**	Consequences not mapped, failure probability unknown (may be available shortly) – approximately 1 breach per year.
	Overtopping	**3**	**D**	
Reservoir breach	>25,000 m^3	**2**	**M**	Covered by Reservoirs Act 1975. / Impact is being modelled in RIM study, but failure probability hard to estimate.
	<25,000 m^3	**2**	**M**	Note proposed introduction of new 10,000 m^3 threshold.
Water supply failures (e.g. burst main)		**3**	**D**	Lack of integrated infrastructure data, no maps of risk available.
FRM infra-structure failure	Breaching	**1**	**M**	Failure probability through generic fragility curves. Significant risk for many areas. NaFRA focus.
	Pumps, barriers etc.	**1**	**D**	Failure probability may not be well understood (even if consequences are). Significant risk for some areas (e.g. Hull).

*Key: **Priority**: 1=must be included, 2=could be included, 3=exclude. **Difficulty** (measure of potential feasibility of including in scope): D=difficult, M=medium difficultly, E=easy.

Table 6.2 Definition of main categories of flooding sources.

Fluvial	Flooding from water emerging from a surface watercourse not subject to tidal influence. Responsibilities for mapping fluvial flooding from main watercourses and ordinary watercourses falls on more than one organisation. The category includes flooding from failure of passive and active flood defences (walls, sluices etc.) and blockages.
Coastal	Flooding from the sea, either as a result of high astronomical tide, meteorological surge or wave action, not subject to the influence of fluvial flows. Includes hazard from failure of passive flood defences (walls, levees etc.).
Estuarine	Flooding from surface watercourses influences by both tide/surge and fluvial flows.
Tsunami	Ocean wave produced by underwater seismic or volcanic activity, or landslide. Distinct from tide/surge of astronomical/meteorological origin.
Groundwater	Flooding caused by water emerging locally from subsurface permeable strata.
Surface water	A surface water flood event results from rainfall generated overland flow before the runoff enters any watercourse or sewer. Usually associated with high intensity rainfall (typically >30 mm/hour) resulting in overland flow and ponding in depressions in the topography, but can also occur with lower intensity rainfall or melting snow where the ground is saturated, frozen, developed or otherwise has low permeability. Urban underground sewerage/drainage systems and surface watercourses may be completely overwhelmed, preventing drainage. Surface water flooding does not include sewer surcharge in isolation. (JBA, 2006)
Sewer	Water exiting from an overloaded subsurface, passive drainage system due to design capacity being exceeded, or due to failure (blockage or collapse). Complementary to surface water flooding, which includes water failing to enter the drainage system due to over capacity.
Canal	Canal breach flooding is caused by embankment failure on artificial navigable waterways. Canals may act as temporary flow paths during storm events and hence be a source of flooding when no embankment failure occurs (termed canal overtopping herein). Canals will be considered as either main river or ordinary watercourses and may thus fall under the 'fluvial' category above.
Reservoir breach	Flooding due to catastrophic failure of dams. Overtopping/spilling events that leave the dam intact will be covered by fluvial flooding.
FRM infrastructure failure	Flooding caused by failure of linear defences (walls and embankments) or active (point) flood defence elements such as pumps (e.g. Hull) and the Thames Barrier. (May also be included in categories above.)

6.2 Priority sources

Following review of Tables 6.1 and 6.2 and discussion with the Environment Agency project board, it was decided to use the following categorisation of sources in this project.

Table 6.3 Prioritisation of sources.

Include and considered feasible	Include but significant issues exist	Exclude (but may still be able to be used and included in the future)
Fluvial (main river) national	Reservoir breach national	Sewer (blockage/failure)
Coastal national	Fluvial (small watercourses) local	Tsunami
Estuarine national	Groundwater local	Canal overtopping
Surface water local	Canal breach*(subject to review)	Water supply failures
Breaching (fluvial/coastal)	Sewer (capacity exceedance)	
	FRM point infrastructure failure (pumps, barriers...)	

Notes: The *national* and *local* descriptions of sources are based on the draft Flood and Water Management Bill and this is subject to change. National-scale mapping is taken as being the Environment Agency's responsibility; local is taken as the responsibility of local authorities. The Environment Agency will still take a strategic overview role for some sources mapped by local authorities (e.g. surface water).

The significant issue with including reservoir inundation is defining the probability of the event – the Environment Agency is undertaking research on this issue and the proposed method will be able to integrate the reservoir inundation mapping once probabilities are assigned.

*Canal breaching and overtopping are likely to be classed as a 'Fluvial' source.

The sources of flooding in the first column of Table 6.3 must be able to be used with the method developed during the project. The method should also be developed to use the sources in the second column; however, it is acknowledged that there are issues with the data and/or physical attributes of these flooding phenomena which may make their inclusion difficult (e.g. their inclusion may degrade the utility of the results). The sources listed in the third column will be excluded from consideration during the project (but with availability of suitable data it may subsequently be possible to use these sources in the developed method).

It should be noted that the sources in column 1 of Table 6.3 correspond to the list of sources for which the Environment Agency is expected to have responsibility for **national-scale flood mapping**. This does not necessarily mean that responsibility for managing the risk from these sources lies with the Environment Agency, as for example surface water is the responsibility of local authorities.

The significant issue with including reservoir inundation is defining the probability of the event – the Environment Agency is undertaking research on this issue and the method proposed in Chapter 8 will be able to integrate the reservoir inundation mapping once probabilities are assigned.

Surface water flooding could be categorised into either column 1 or 2 of Table 6.3; again the issue is assignment of appropriate probabilities to the inundation mapping. The first generation maps (defined as 'areas susceptible to surface water flooding') are not associated with a specific return period, although the storm probability used (0.5% AEP) could represent an approximate probability with large uncertainty due to the lack of representation of drainage and infiltration, and the single storm duration used to generate the maps. The second generation maps (due for delivery in early 2010) will have a better representation of the probability of flooding rather than a semi-quantitative estimate of susceptibility, and though they will be an improvement over the first generation there may still be problems associated with extrapolation to other depths generating a large uncertainty in output maps (see Chapter 8). Surface water maps compatible with an all sources mapping methodology may have to wait until the outputs of the ‘Next Generation of Surface Water Flood Risk Assessment’ project are available.

7 Datasets

7.1 Dataset details

This chapter provides details of the key data sources identified during consultation and literature review. The details of the data sources are provided in Tables 7.1 to 7.7. Section 7.2 provides a summary of the suitability of current and planned future datasets for use within the proposed ‘all sources’ method.

Table 7.1 Flood maps from fluvial, estuarial and coastal sources (key Environment Agency products).

	NaFRA (main output)	**MDSF2 outputs**	**Flood Zone modelling**
Dataset name	NaFRA 2008	MDSF2 outputs	Flood Zones
Flood sources included	• Fluvial • Estuarine, and coastal (includes consideration of joint probability of tide and surge and waves) • Flood risk from overtopping and/or breaching of linear assets or in undefended areas	• Fluvial • Estuarine, and coastal (includes consideration of joint probability of tide and surge and waves) • Flood risk from overtopping and/or breaching of linear assets **and** some point assets or in undefended areas	• Fluvial • Estuarine • Coastal • Flood risk from overtopping only or in undefended areas
Output	Probability of flooding at centre of each cell, subsequently classified into **low** (≤0.5% year), **moderate** (1.3 to 0.5% year), **significant** (>1.3% year) (internal product also has very low and very significant classifications). Secondary output includes probability of flooding above a specified depth threshold (these outputs are the preferred data for use in the prototype tool)	To be defined	Flood extent polygons at selected probabilities (1% fluvial, 0.5% coastal and 0.1% both fluvial and tidal)
Scale and coverage	50 m square cells	Expected to be variable but similar to NaFRA 2008	Variable – dependent on source model
	Bounded by Flood Zones 2	Bounded by Flood Zones 2	Defines Flood Zone 2 extent (0.1%)
	England and Wales coverage	Local in England and Wales – expected by catchment/subcatchment	England and Wales coverage
Data structure	SQL-server relational database system and ESRI grid/polygon	Oracle-based relational database system and ESRI grid/polygon	NFCDD polygon and attribute data
IPR	Results owned by Environment Agency	Results owned by Environment Agency	Results owned by Environment Agency
Accuracy and resolution	Broad-scale	Intermediate	Variable – dependent on source model
Ease of use in new method	Readily usable	Expected to be readily usable	Readily usable
Other comments	'Broad-brush (or high level) assessment of the likelihood of flooding at a national scale' applying the RASP methodology	Development of the RASP methodology as applied in NaFRA 2008 in a desk-top application	Uses national generalised modelling (JFLOW) and locally updated using local JFLOW re-runs and detailed 1D–2D hydraulic hydrodynamic modelling.

Table 7.2 Flood maps from fluvial, estuarial and coastal sources (studies).

	SMP and CFMP studies	Strategy/PAR studies
Dataset name	Varies	Range of modelling methods used to generate flood data, e.g. TUFLOW, ISIS, MIKE 21
Flood sources included	• Fluvial, coastal and estuarine dependent on study • Can including surge, tide and waves • Flood risk from overtopping, breach or erosion (SMP)	• Fluvial • Estuarine • Coastal including (surge, tide and waves) • Flood risk from overtopping, breach or erosion dependent on study
Output	Varies Some studies did not undertake new flood modelling (especially SMPs)	Water levels at defences to derive overtopping discharge; flood extent polygons at selected return period events and in selected epochs; erosion lines developed from high level analysis (i.e. Brunn Rule or specific erosion modelling)
Scale and coverage	As in source data/model	Varies, can be at 10–50 m square cells
	High level England and Wales	As and where available both in England and Wales
Data structure	Varies, grid, polygon	Varies, can be grid or polygon
IPR	Environment Agency	Environment Agency
Accuracy and resolution	As source data	Varies
Ease of use in new method	As in source data/model	Varies
Other comments	High level assessment based on Environment Agency flood data. Produces broad economic assessment	More detailed assessment than at SMP/CFMP level

Table 7.3 Flood maps from surface water sources.

	Surface water flooding
Dataset name	(i) At the national scale, a map showing areas susceptible to surface water flooding has recently been produced and will be improved through a subsequent project (second generation surface water flood risk maps). (ii) In some local areas there are maps of surface water flooding and it is expected that the Surface Water Management Plan (SWMP) programme of work will contribute further mapping.
Flood sources included	Surface water flooding. Note that some surface water flood mapping may also include flooding from sewers and from minor watercourses.
Output	National surface water flood mapping is presented in three bands of likelihood (more, intermediate, less). Although it was generated using a 200 year rainfall, there is no probability attached to the flood map. At the local scale, the SWMP maps are likely to vary in format/content. The current SWMP guidance recommends a range of return periods are used and that hazard indicators will be produced.
Scale and coverage	At the national scale the coverage is England and Wales and the expected viewing scale is 1:50,000.
	At the local scale the coverage will be intermittent and scales will vary.
Data structure	National – polygons. Local – various.
IPR	Results from the SWMP mapping will probably be the IPR of the local authorities. The first generation national surface water map is licensed to the Environment Agency for specific uses only. The second generation map will be owned by the Environment Agency.
Accuracy and resolution	Will vary.
Ease of use in new method	National surface water flood mapping – first generation maps have no probability and thus are hard to use, second generation maps are likely to be easier to use. Local surface water flood mapping – ease of use will vary.
Other comments	The Environment Agency Science project 'Developing the Next Generation Surface Water Flood Risk Assessment' is expected to lead to improved and more consistent surface water flood maps.

Table 7.4 Flooding from surface water drainage networks.

	Drainage networks/sewers
Dataset name	No specific dataset has yet emerged at the national scale.
Flood sources included	Surface water flooding caused by exceedance of the capacity or collapse/blockage of sewers and related infrastructure such as permeable pavements, filter drains, filter strips, soakaways and swales. Foul flooding can occur due to runoff entering the foul network or due to blockage or collapse of the sewers. Foul flooding as a result of insufficient capacity is already monitored by water companies.
Output	Other than records of historical flooding incidents, maps of flooding in this category are rare and therefore there is insufficient information to complete subsequent rows.
Scale and coverage	No information.
Data structure	No information.
IPR	Historical sewer flooding maps will have water company IPR. Water Companies would not generally release the data, but have done so to allow completion of Water Cycle Strategies.
Accuracy and resolution	Not known.
Ease of use in new method	Difficult.
Other comments	Surface Water Management Plans may assist in the generation of suitable data.

Table 7.5 Flooding from groundwater sources.

	Groundwater
Dataset name	Information on flooding incidents may be available in the following systems: 1. Environment Agency 'Historic surface and groundwater flood records' project (ongoing due to finish March 2010) will collate data from a variety of sources nationally. 2. FRIS: Flood Reconnaissance and Information System, some regions, e.g. SW. 3. NFCDD: National Flood and Coastal Defence Database, contains flood event outlines including groundwater flooding incidents. 4. WISKI: for field monitoring data, available at a national scale. 5. Collation of Fire Service Callout Database records on groundwater flooding by Mills (2004), MSc thesis at Birmingham University. Indicative groundwater flood mapping is available from the following sources: a) Groundwater Emergence Maps (GEMs) produced for Strategy for Flood and Coastal Erosion Risk Management: Groundwater Flooding Scoping Study LDS23. b) Groundwater flooding maps produced originally for MSfW HA4, now held by JBA, for indicative 100 year floods in Chalk aquifers. c) BGS susceptibility to groundwater flooding maps, for all aquifers.
Flood sources included	1. Groundwater flood due to rising groundwater level (GWL) – Chalk aquifers. 2. Groundwater flood due to rising GWL – alluvial aquifers. 3. Groundwater flood due to rising GWL – non-Chalk aquifers. Flooding may occur at springs or due to rising water table and therefore pumping operations have a direct bearing.
Output	Not yet known.
Scale and coverage	Such data are regionalised and within the responsibility of the Environment Agency, although utility companies are stakeholders for abstracting groundwater by pumping. Jacobs, JBA and BGS have indicative mapping for England and Wales at around 1:50,000 scale.
	Following the recommendation by Jacobs, the Environment Agency has collected such data since 2006, but these are in terms of flooding incidents, which may help to better establish flood frequency analysis related to groundwater.
Data structure	Various. Data collected for Environment Agency 'Historic surface and groundwater flood records' project to one consistent format; other collected incident data are based on the NFCDD and/or FRIS and data may include occurrence maps and current hazard and risk maps. GIS/geodatabase files.
IPR	Environment Agency holds incident data. Jacobs/Defra hold GEM mapping. JBA and BGS/NERC hold their respective datasets.
Accuracy and resolution	Incident data are expected to be related to discrete incidents. Groundwater floods are sometimes obscured by greater volumes of fluvial flooding so may be under-represented in the incident databases. Indicative mapping uses simplified assumptions of groundwater movement. Areas of groundwater emergence are indicated well in Chalk aquifers, but no account is taken of surface flow (i.e. the flooding). Less certain in non-Chalk aquifers.
Ease of use in new method	Not yet known.
Other comments	There are anticipated problems for integrating groundwater and fluvial/coastal modelling capabilities for the following reasons: 1. Groundwater flow operates at a different temporal scale than other sources. 2. Groundwater flooding problems are temporally more persistent but the amount of excess water is not often as large as fluvial flows. 3. Groundwater modelling grapples with more uncertainty due to heterogeneity of aquifers than open channel flows. Therefore best practice is yet to emerge on probabilistic flood frequency analysis due to groundwater flooding problems. There are also methodological problems in definition of the probability of exceedance associated with the flooding from groundwater sources. 4. Antecedent conditions that may cause groundwater flooding can also lead to increased surface water flooding.

Table 7.6 Flood risk posed by reservoir breaches.

	Large raised reservoirs: breaching (reservoir capacity > 25,000 m^3)	Non-statutory reservoirs: breaching (reservoir capacity between 10,000 and 25,000 m^3)
Dataset name	No information is available at a national scale yet but is in preparation through the Environment Agency's National Reservoir Inundation Mapping project for England and Wales, which was due for completion end 2009. Consequence information in terms of risk to life, in bands, will be available. Detailed maps for some reservoirs exist, prepared by the owners. New legislation for Scotland enables SEPA to request or prepare flood plans.	Work is in hand by the Environment Agency to identify small reservoirs and to assign them to consequence categories. A pilot study to develop a methodology to estimate risk from limited information is proposed by the Environment Agency.
Flood sources included	Flooding due to reservoir dam breach – capacity above 25,000 m^3.	Flooding due to reservoir dam breach – capacity in the range of 25,000–10,000 m^3.
Output	Reservoir flood plans will be available by the end of 2009 for local resilience forum (LRF) use.	A database may be developed. Under new proposed legislation, flood maps will be required for 'high risk' reservoirs of over 10,000 m^3 capacity.
Scale and coverage	Scale expected to be as per model scale, which are usually 2D models.	Scale expected to be as per model scale, which are usually 2D models.
	Coverage: all 2,100 reservoir catchments in England and Wales; also plans for Scotland under new Scottish legislation.	Coverage: in Scotland, flood plans will not be required for reservoirs of less than 25,000 m^3 capacity.
Data structure	Not known yet.	Not known yet.
IPR	The IPR for the new flood plans in preparation lie with the Environment Agency. The IPR for plans in Scotland will lie either with SEPA or the dam owners. There are also security issues with the information concerning flood risk maps from this source.	Not known at present. There are security issues with the information concerning flood risk maps from this source.
Accuracy and resolution	Depends on the breach formula and on the 2D model.	Depends on the breach formula and on the 2D model.
Ease of use in new method	Difficult – probability will not be defined and there may be issues with dissemination of results.	Difficult – probability will not be defined and there may be issues with dissemination of results.
Other comments	Driven by the Water Act 2003 and the Pitt Review. The methodology consists of empirical breach hydrographs fed into 2D models to give extreme flood levels, hazard estimates and travel times.	A recent project has estimated at there are at least 4,700 reservoirs falling in this band of reservoir capacity. Separate studies indicate that a breach at about a third of these reservoirs might pose a risk to life.
	Assigning probability of exceedance to flooding from reservoir breaches remains an issue and this is unlikely to be defined by the ongoing RIM project. However, pragmatic rule of thumb methods may be devised based on the median probability of dambreak failure.	

Table 7.7 Incidental flooding risk stemming from lack of maintenance or similar problems.

	Water company asset failure (e.g. burst mains)	**Environment Agency or others infrastructure point or local failure**
Dataset name	There are calls to collect data on incidents leading to flooding but a national register has not yet emerged.	Not known.
Flood sources included	Sources can be diverse, including: 1. Pipe bursts in water supply systems. 2. Surcharged sewer outfalls or problems associated with valves and non-return valves. 3. Pump failure resulting in flooding at the pumping station and the vicinity. 4. Pipe bursts in raw water or sewerage rising mains. 5. Sewer blockage or collapse (see other table). 6. Aqueduct failure.	FRM assets are large and each prone to their particular risks: 1. Rivers may be heavily weeded or suffer from debris and fallen trees. 2. Bridges may block by debris and due to siltation. 3. Sluices and gates may be blockage or their moving gears fail. 4. Temporary/demountable flood. prevention measures may fail to operate 5. Bunds may breach.
Output	All Water Companies have data and report this annually in the June returns to Ofwat. Data is available from Ofwat	Varies. May be included in NaFRA, MDSF2, CFMP, SMP, strategy models etc.
Scale and coverage	Such data are expected to be related to discrete incidents held by all utility companies.	Flooding from asset failures are expected to be discrete.
	Water companies have internal and external registers, e.g. DG5 register.	
Data structure	No common structure.	Not known.
IPR	Existing databases are thought to be the properties of their collecting utility companies.	Likely to be Environment Agency.
Accuracy and resolution	There may be consistency problems across different utility companies.	Not known.
Ease of use in new method	Likely to be difficult.	Likely to be difficult.
Other comments	Resilience of critical assets emerged as an issue in the summer 2007 floods and Ofwat is driving a strategy to assess them and make them more resilient. This therefore can give rise to new datasets. There will be other initiatives for mapping flooding related to critical infrastructure.	

Table 7.8 Canal overtopping and breaching.

	Canal overtopping	Canal breaches
Dataset name	Overtopping database records details of flood overtopping incidents that occur each month, available as a GIS database, but no inundation maps included. Some SFRAs and local FRAs (e.g. Crossrail) have produced inundation modelling. British Waterways (BW) holds a record of studies on which it has been consulted (but not in database format). Longer term, a high level screening project is expected to identify canal pounds for which BW may want to carry out inundation mapping, but this will be a very long-term programme, which will probably never achieve national coverage.	Research programme under way to identify and record historical breach locations, dating back to 1770. Database/GIS layer currently has 260 records, but no inundation mapping. Potential failure of assets (e.g. embankments or aqueducts) assessed through 1-year and 10-year inspections, with condition grade (A–E) and consequence grade (1–5) assigned. Consequence grade is by visual inspection of nearby receptors, not through any modelling. Longer term, asset inspection could be used to inform high level screening, which would identify pounds with highest potential consequence and then recommend them for inundation modelling.
Flood sources included	Flood water from other sources, principally local feeder channels.	Water from within the canal system or flood water from other sources.
Output	Principally in SFRA's and FRA's at present. Likely to be high level, possibly for one event probability only.	Principally in SFRA's and FRA's at present. Likely to be high level, possibly for one event probability only.
Scale and coverage	Variable scale (SFRA/FRA).	Variable scale (SFRA/FRA).
	No consistent, national coverage. No plans to produce a national dataset.	No consistent, national coverage. No plans to produce a national dataset.
Data structure	n/a	n/a
IPR	BW owns IPR to overtopping database. Local authorities will own SFRA inundation maps. Developers will own FRA inundation maps.	BW owns IPR to breaching database. Local authorities will own SFRA inundation maps. Developers will own FRA inundation maps.
Accuracy	Unknown – likely to vary. BW seeks to guide on modelling when consulted.	Unknown – likely to vary. BW seeks to guide on modelling when consulted.
Ease of use in new method	Unknown – depends on availability of local data and format.	Unknown – depends on availability of local data and format.
Other comments	Phase 2 of 'all sources' project could seek to identify and obtain canal overtopping/breaching data from one or two SFRAs for testing against the framework.	

7.2 Gaps

The following knowledge gaps covering data for mapping flood risk from all sources have been identified:

- Assigning probability of exceedance to flood risk from groundwater sources and from reservoir breaches remains an issue. Work is currently being undertaken by the Environment Agency to estimate probabilities of reservoir failure.
- There are initiatives on mapping flood risk by IDBs, SFRA and historical data but the relevant information is not readily available.
- The contribution of swell waves (long-wavelength ocean waves from non-local wind action) to coastal flood risk is poorly understood at present.
- Best practice procedure is yet to be developed for mapping floods from some of the sources, e.g. incidental flood risk from lack of maintenance.
- Different resolutions are associated with flood risk data from the various sources and therefore some form of resampling may be required to bring all the information to the same basis.

7.3 Summary of suitability

Table 7.9 provides a summary of the suitability of the datasets described in Tables 7.1 to 7.8 for use within the proposed 'all sources' method. The suitability is assessed under three headings:

- 'Current data' – the current version of the dataset is considered suitable for use within the method, particularly at the national scale.
- 'Datasets expected before 2011' – suitable datasets are expected to be available before January 2011.
- 'Future planned datasets' – there are plans to provide datasets which are likely to be suitable but the date the data will be available is either not known or is after December 2010.

The criteria used to determine suitability are:

- Contents of the dataset represent a source of flooding that should be included in the 'all sources' map.
- Coverage is 'England and Wales' for national datasets, or 'extensive' for local datasets.
- IPR or secrecy issues are unlikely to cause problems with use of the data, although there is a security issue associated with flood maps from reservoir breaches.
- Whether the dataset assigns a quantified probability to the flood likelihood, or vice versa.
- Formats are suitable for use with the method.
- Accuracy is such that the dataset is expected to be useful for the 'all sources' requirements.

Other considerations are the availability of uncertainty information and dependency/joint probability information.

Table 7.9 Summary of suitability of datasets.

Dataset or source of flooding	Flooding sources	Suitability for use in proposed 'all sources' method			Comments
		Current data	Datasets expected before 2011	Future planned datasets	
NaFRA	Fluvial, estuarine, coastal	Yes	Yes	Move towards local probabilistic models, facilitated through MDSF2. Outputs will be aggregated into national picture, similar to approach used for current Flood Zones.	Uncertainty information available through sensitivity analyses (SC050064 RASP HLM+ Sensitivity Analysis) and validation studies (forthcoming FDI(09)29 Improved model validation/calibration).
Flood Zones	Fluvial, estuarine, coastal	Yes	Yes	Taken as a scenario from our probabilistic modelling, i.e. single event loading with all defences failed to represent no defences.	Uncertainty likely to be high, not well understood.
SMP studies	Coastal	No	No	Will be taken from one source rivers and sea probabilistic outputs, i.e. NaFRA update.	
CFMP studies	Fluvial	Probably difficult	Probably difficult	Will be taken from one source rivers and sea probabilistic outputs, i.e. NaFRA update.	Uncertainty likely to be high, not well understood.
Strategy studies	Fluvial, estuarine, coastal	Yes	Increasing coverage	Will be taken from one source rivers and sea probabilistic outputs, i.e. NaFRA update.	Uncertainty in water depths reasonably well understood.
Surface water flooding	Surface runoff and sewer capacity exceedance	No	Second generation maps	Third generation approach, run by owners of local flood risk, as part of SWMP. We hope/anticipate they will adopt our standards to allow for easy integration into a national picture.	Likely to move towards probabilistic modelling for next generations.Second generation to include single depth/probability. Sewer capacity exceedance covers water before entry into sewer system, rather than water emerging from it.
Surface water drainage networks	Sewer capacity exceedance	No	No	Third generation approach, run by owners of local flood risk, as part of SWMP, where the probabilistic approach will promote inclusion of surface water risk in 'all sources' mapping.	Sewer models are likely to include capacity exceedance for water both entering and emerging from the system. Local models may use very simplistic surface models.
Groundwater	Range of flooding mechanisms	No	No	May still not have a probability attached. Will also be looking towards probability of duration of inundation.	Concentrated on collection of incident data; these may be translated into risk maps. Data may be qualitative. Some qualitative local data available.
Reservoir breach	Breach of reservoir >10,000 m^3	No	RIM outputs	Non-statutory reservoir information.	If lacking breach probability data may need to classify as qualitative, e.g. as reservoir risk zones.
Pipe bursts etc.	Clean water supply infrastructure failure	No	No	Yes.	Assessment of critical assets may provide qualitative information.
FRM point asset failure	Failure of gates, barriers etc.	No	No	Move towards local probabilistic models incorporated into national picture, similar to approach used for current Flood Zones.	Some local models and NaFRA/MDSF2 will include these. May be only as scenarios with no probability data.
Canal overtopping		?	?	Move towards local probabilistic models incorporated into national picture, similar to approach used for current Flood Zones.	
Canal breach		?	?	No national programme, but further development of inundation maps may result from further development of SFRAs (e.g. Level 2), FRAs and BW's own local (risk-based) inundation mapping.	

7.4 Suggestions for overcoming data issues

Table 7.9 summarises the status of potential datasets to facilitate flood risk mapping from all sources. There are a series of potential issues associated with use of the data for mapping all sources – this section summarises these issues and suggests potential ways forward.

- Mapping data from different sources have different resolutions. While their integration in terms of wet cells and dry cells is feasible through some form of resampling, the introduction of new errors is likely. To minimise errors, facilities for some sensitivity tests would be appropriate.
- Although SWMPs are produced in response to flooding from storms of designated probabilities of exceedance, this does not necessarily mean that floods from surface water sources will have the same probability of exceedance. However, equivalence of probability could be a first order assumption, until improved relationships exist between the probability of exceedance of storms and resultant flooding from surface waters.
- There are methodological problems associated with assigning probability of exceedance to flooding from some of the sources (e.g. flooding from groundwater and reservoir and canal breaches). It is possible that 'rules of thumb' can be developed to assign probability of exceedance. This needs to be progressed separately from this project.
- Different levels of uncertainty are associated with each flood mapping dataset product. In reality, a methodology needs to be established for the propagation of uncertainty through the combined sources, and using this to communicate confidence in the results to the user.
- There is currently no information readily available on the nature of datasets for flooding from some of the sources. For the development of prototype mapping software, assumptions can be made until more appropriate data becomes available.
- There are a range of sources of flooding for which there is no suitable data available (e.g. tsunamis and swell waves). The Environment Agency should maintain a 'watching brief' for developments in these areas.
- Best practice procedures are also evolving for each source and the methodology proposed by this project will need to be reviewed periodically to ensure it remains appropriate.

8 Proposed method

8.1 Introduction

In this chapter we describe the proposed methodology for mapping flooding from all sources, combining the map combination and fully integrated approaches as defined in Chapter 3. The descriptions in this chapter assume the reader has a good working knowledge of the probability of extreme events and some familiarity with joint probability methods as described in Hawkes (2005), FD2308, *Use of Joint Probability Methods in Flood Management: A Guide to Best Practice*.

A flow chart for the methodology is shown in Figure 8.1, and a typical application of the methodology is described as follows:

Stage 1 – Data collection	Collect existing flood information maps at national and possibly local scales.
	Collect information on uncertainty and dependencies, if available.
	Collect information on different scenarios (management options, data for future climates) if required.
Stage 2 – Use map combination approach to generate combined flood map	Use the software tool to generate an 'initial' combined map.
	Use tool outputs to assess locations where dependency may generate interactions, and decide whether these are in high consequence areas.
	Use tool outputs to assess areas where uncertainty is high, and decide whether these are in high consequence areas.
Stage 3 – Local modelling	Where necessary, use local models to map flood risk from combined sources and feed back into stage 2.
	Where necessary, use local models to map risk more accurately in high consequence areas and feed back into stage 2.

The two components at the core of this methodology are:

- **Method for combining existing flood maps** producing a single measure of combined hazard for each spatial element (such as floodplain cells or polygon elements). The spatial extent of these cells is determined by the input data.
- **Method for deciding where more accurate local modelling is required**, either due to dependency between sources, or uncertainty in national/catchment-scale maps, where this coincides with high consequence areas.

This methodology represents a mixture of the map combination (combining existing flood maps) and fully integrated (external modelling) approaches. These two approaches are related to the source–pathway–receptor system based model of flooding (Figure 8.2).

The fully integrated approach can be interpreted as combining at source level, with the combined contributions routed along pathways through to receptors. Water from different sources can thus interact at both pathway and receptor level, allowing, for example, hydraulic effects to be represented.

The map combination method represents combination at receptor level. Water from different sources is routed along pathways independently, and then combined in floodplain cells (e.g. model grid squares) to give a single hazard measure (e.g. water depth).

The methodology proposed here, a hybrid of the map combination and fully integrated approaches, has been selected with the following rationale based on the consultation responses, literature survey and workshop:

- Within the time and resource constraints of this project, the delivery of the map combination approach is more feasible and still delivers the majority of the benefits realisable through a fully integrated approach.
- There are considerable technical difficulties in delivering a generalised fully integrated approach. The representation of sources, pathways and to some extent receptors depends on the source of flooding being modelled, and as such is better left to specific models, rather than the 'one size fits all' approach that the fully integrated method would require.
- The method makes the best use of existing modelling and modelling methods, rather than replicating work already undertaken just to assess combined flood probability. This also ensures that the map of all sources is consistent with the estimates of flooding from its components.
- The method also means it will be straightforward to include future datasets as they arise, either for sources not previously mapped, or as better datasets replace existing ones. The approach of bringing in external data where possible means the method should remain appropriate in the future.
- Until the tool has been developed and piloted, it is not clear how much influence interaction between sources has on overall risk. The methodology proposed here allows us to assess which areas are at potentially increased risk due to interaction and dependency between sources, and hence make an informed decision about whether it is worthwhile modelling these interactions.
- The approach also allows probabilistic occurrences such as blockages, defence failures and reservoir breach to be treated in the same framework. Flood probability arising from blockages in culverts or bridges, for example, can be treated as a separate source. This contribution of these types of asset failure sources to the total flood probability may be useful information, for example to inform asset maintenance.

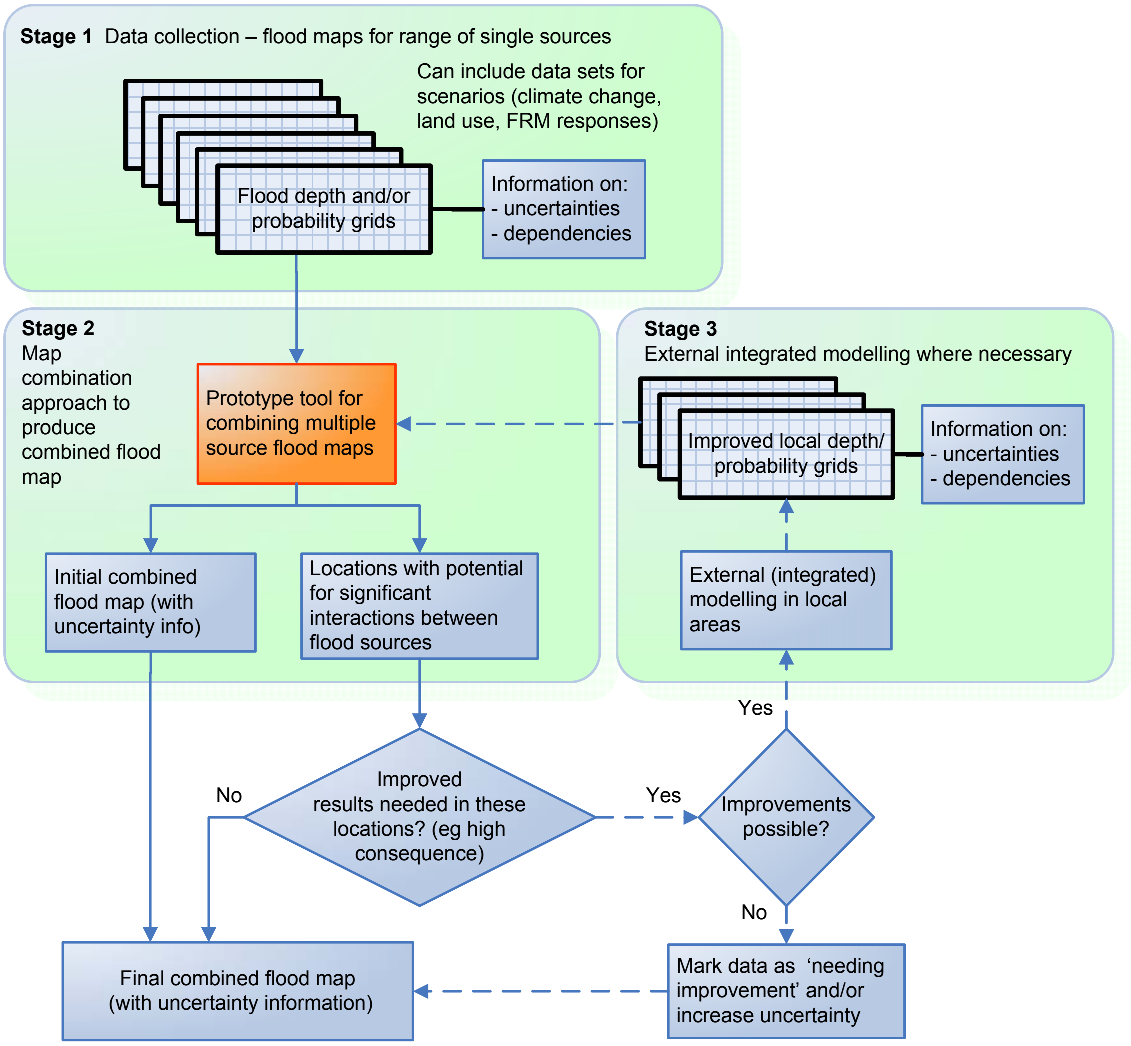


Figure 8.1 Flow chart for combined mapping of flooding from all sources.

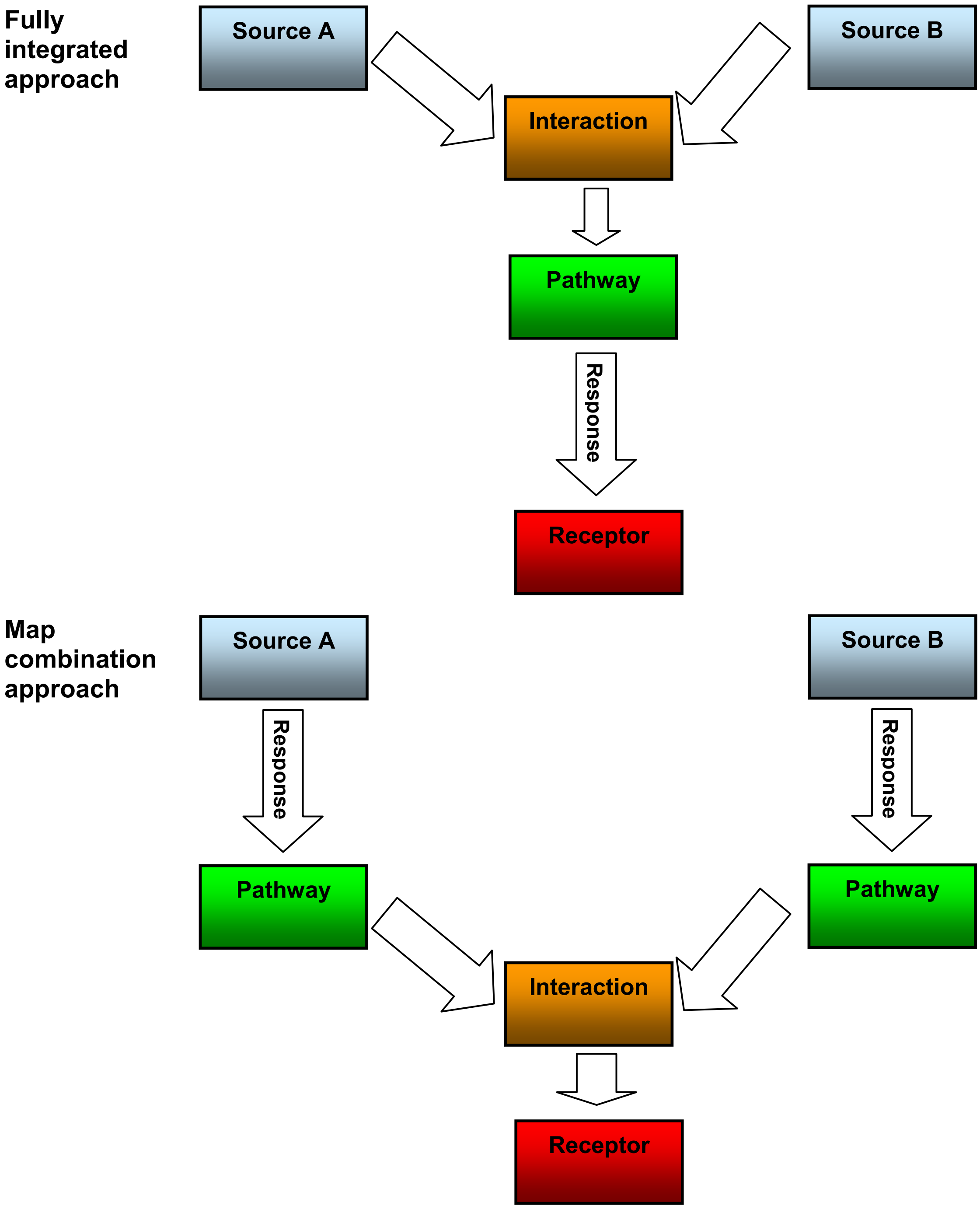


Figure 8.2 Source–pathway–receptor interpretation of the two approaches.

8.2 Proposed method

The method for combining maps needs to take a probabilistic approach to both deterministic and probabilistic model outputs, as both of these types include information on flood hazard probabilities. Development of the method starts with analysis of the case for independent sources, which is a useful simplification.

8.2.1 Theory of combining independent sources

Formally, independence between sources means the probability of event A conditional on event B is equal to the unconditional probability of A:

$$P(A|B) = P(A) \tag{8.1}$$

It is not clear whether any sources of flooding are truly independent, as even a weak seasonality will tend to cause flooding to coincide. However, it is a useful limiting behaviour for weakly dependent sources.

The probability of either of two independent events A and B occurring is given by adding the mutually exclusive combinations of A and B:

$$P(A \text{ or } B) = P(A \text{ and } B') + P(A' \text{ and } B) + P(A \text{ and } B) \tag{8.2}$$

A´ here indicates the probability of event A not occurring. In practical terms, this is a probabilistic description of flood risk from either of two sources, with events A and B representing the sources exceeding given thresholds (e.g. river flow exceeding the 1% AEP flow).

Equation 8.2 can be simplified first by neglecting the $P(A \text{ and } B)$ term. Since sources A and B are independent, and we are dealing with rare events, the probability of them occurring at the same time is insignificant.

Taking fluvial and coastal flooding as an example, the probability of two 1% AEP events occurring at the same time is less than 0.01%. Even if a 1% fluvial flood and a 1% coastal flood occur in the same year, the probability of the flood event (lasting typically a few hours to a few days) coinciding with the highest tide of the year is small. The probability of events coinciding will be increased by dependency between the sources, event duration (long events are more likely to coincide) and seasonality (long flood events in a short flood season are more likely to coincide).

Combining (1) and (2) gives the familiar additive form for either of two events occurring:

$$P(A \text{ or } B) = P(A) + P(B) \tag{8.3}$$

This is easily extended to more than two sources:

$$P(A \text{ or } B \text{ or } C \text{ or } ...) = P(A) + P(B) + P(C) + ... \tag{8.4}$$

For any number of sources, the addition of probability can be ‘chained’, by combining A and B, then combining the result with C, and so on. The validity of equations 8.3 and 8.4 decreases as the probabilities of events increase, and as the numbers of sources increase. This is because for more sources and more likely events, they are more likely to coincide even if independent.

This analysis is valid when applied to two or more sources of flooding at source level. The probabilities $P(A)$ etc. should be interpreted as the probability of source variables exceeding given thresholds, but since the possibility of this forcing occurring at the same time from two sources is being ignored, this is equivalent to the probability of floodplain depths (or other hazard measures) exceeding given thresholds. The probabilities for depth exceedances from two sources can therefore be written as:

$$P(\mathrm{X_{Combined}} > \mathrm{x}) = P(\mathrm{X_A} > \mathrm{x}) + P(\mathrm{X_B} > \mathrm{x}) + P(\mathrm{X_C} > \mathrm{x}) + ... \tag{8.5}$$

$P(\mathrm{X_A} > \mathrm{x})$ represents the probability of the hazard X exceeding a threshold x due to the source A. The exceedance probabilities for floodplain hazards (receptor level) can thus be combined in the same way as at source level. For the following descriptions we deal exclusively with depths; extension to other hazard measures is in theory straightforward, although practical considerations (especially resolution) may make application to these measures more difficult.

8.2.2 Application to flood hazard data

While the theory of adding probabilities is simple, in practice the diversity of probabilistic data source types, uncertainty and incomplete information mean that the addition of probabilities must be carried out carefully.

In an ideal case, where the input flood hazard maps give a complete picture of how floodplain depths vary with probability, addition of probabilities requires interpolation between data points on the depth–probability curve (Figure 8.3). This allows the probability for a required output depth, or depth associated with an output probability, to be determined. In practice, the interpolation is carried out using reduced variates, such as the Gumbel (y_G) or logistic (y_L) reduced variates defined through the exceedance probability P:

$$y_G = -\ln(-\ln(1-P)) \tag{8.6}$$

$$y_L = -\ln\left(\frac{P}{1-P}\right) \tag{8.7}$$

For extreme events associated with small exceedance probabilities, both these variates reduce to $-\ln(P)$. Plotting depths against these reduced variates tends to produce straight (or straightish) lines, and so interpolation errors will be smaller using these variates rather than the exceedance probability itself.

Addition of probabilities as in equation 8.5 also allows the contribution to the probability from each source to be determined.

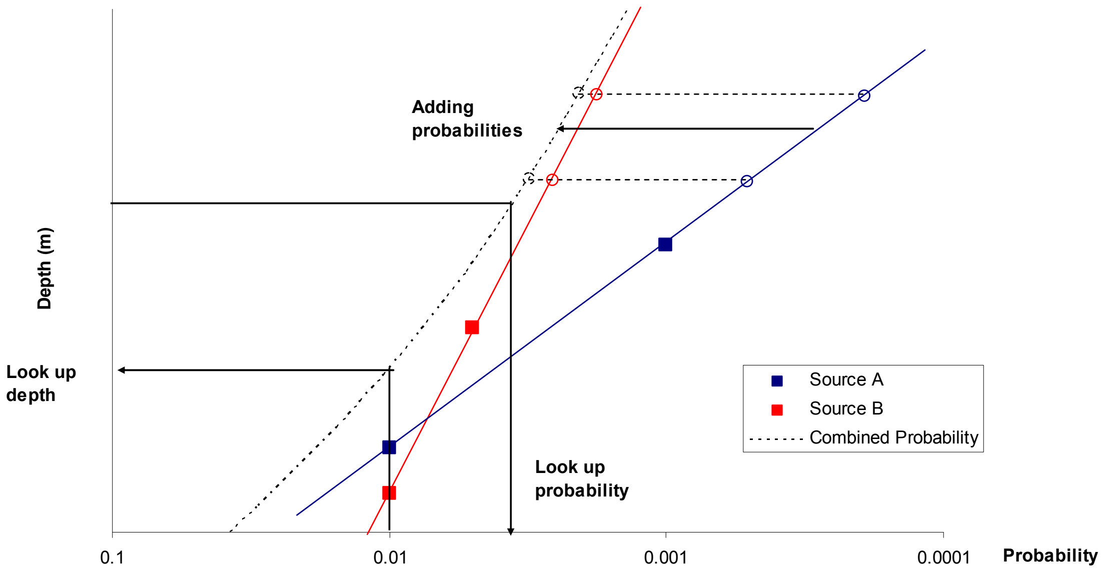


Figure 8.3 Addition of probabilities using interpolation of the depth–probability curve.

Uncertainty, incomplete information and dependency are all dealt with using methods based on interpolation and addition of probability-reduced variate curves. These are described in detail in the following sections.

8.2.3 Representing uncertainty

The probability-depth curves of Figure 8.3 correspond to a complete understanding of the natural variability of floodplain depths. In practice, these estimates will be uncertain, and this uncertainty can be expressed as a range of probabilities at each depth, or a range of depths at each probability. For a more detailed discussion of this issues see Appendix D.

For this project, we propose dealing with uncertainty in a relatively simple way that still captures the interactions between knowledge uncertainty and natural variability.[1] The depth–probability curve is represented as a range of probabilities, the ends of this range representing an upper and lower bound for the probability at each depth. Adding ranges of probabilities is carried out using the following equations:

$$P_{Upper}(\mathrm{X}_{\mathrm{Combined}} > \mathrm{x}) = P_{Upper}(\mathrm{X}_{\mathrm{A}} > \mathrm{x}) + P_{Upper}(\mathrm{X}_{\mathrm{B}} > \mathrm{x}) + P_{Upper}(\mathrm{X}_{\mathrm{C}} > \mathrm{x}) + \ldots \quad (8.8)$$

$$P_{Lower}(\mathrm{X}_{\mathrm{Combined}} > \mathrm{x}) = P_{Lower}(\mathrm{X}_{\mathrm{A}} > \mathrm{x}) + P_{Lower}(\mathrm{X}_{\mathrm{B}} > \mathrm{x}) + P_{Lower}(\mathrm{X}_{\mathrm{C}} > \mathrm{x}) + \ldots \quad (8.9)$$

These estimates of the upper and lower bounds will tend to overestimate the uncertainty (quantified as the difference between the bounds). A more rigorous approach would represent uncertainty as a convolution of the uncertainties in the input parameters, but this would require a knowledge of the probability distributions representing uncertainty. These are unlikely to be available for many types of input data, and so equations 8.8 and 8.9 represent a pragmatic approach to understanding uncertainty based on limited input data.

This addition of probability ranges is illustrated in Figure 8.4 for two sources. These upper and lower bounds can be used to represent any type of uncertainty that can be quantified in terms of the effect on flood depths and/or probabilities. Parameter uncertainty (due to imperfectly defined model inputs) and model error (due to imperfect models) will combine to produce an uncertainty in model outputs, and this can be represented as the upper and lower bounds.

For most data sources the uncertainty is expressed in terms of the depth rather than the probability. This can be transformed into upper/lower probability bounds using the gradient of the depth–probability (or reduced variate) curve.

[1] Knowledge uncertainty and natural variability, along with other terms, are defined in Appendix D.

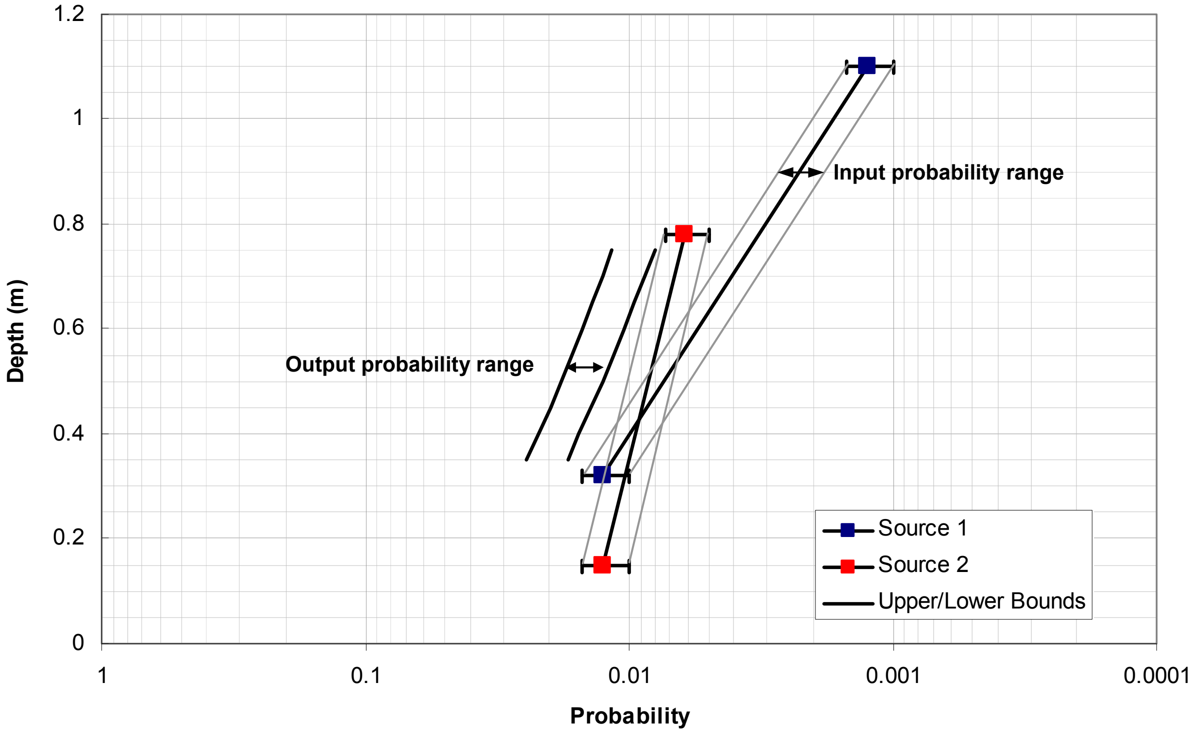


Figure 8.4 Addition of uncertain probabilities.

Displaying both upper and lower bounds spatially will be difficult, as the user will probably only be able to interpret a single colour scale. There are two options for displaying a single probability measure from uncertain probability data:

- Use a combined exceedance probability (CEP) approach (Kirby and Ash 2000), which merges the probability distributions from natural variability and knowledge uncertainty into a single distribution. In this case, assuming a uniform distribution between the upper and lower bounds means that the CEP probability lies halfway between the upper and lower bounds. A drawback of this approach is that extrapolation from incomplete data produces a wide range between the bounds, and the CEP value may not represent our understanding of flood probability. For example, Figure 8.5 shows that for small depth thresholds, the upper and lower bounds span [0.02,1], giving a CEP probability of approximately 0.5. A user may interpret this as an extremely high risk of flooding, rather than a reflection of our lack of model data at small depths.

- Use extrapolation to extend the 'best estimate' values (i.e. the depth–probability grid inputs) to smaller and greater depths. The method described in the next section for dealing with uncertainty will produce widely spaced upper/lower bounds where extrapolation is used, and, if this is communicated to the user, it can be used to place less weight on these areas.

The map combination method will also need to represent qualitative uncertainty information. An example is where two datasets represent the same source, where double counting would lead to an overestimation of flood probability if they were treated as separate sources. For two sources with different uncertainties (e.g. represented by a simple uncertainty score), the one with the lowest uncertainty should be used. When there is a conflict between two datasets, with different results but the same uncertainty scores (i.e. we have no reason to favour one over the other), then taking the maximum depth from the two datasets may be a reasonable assumption that retains data from both but avoids double counting.

This qualitative treatment will also provide a way of including local data brought in to represent high consequence areas or areas where dependency and interaction are important. These local models can be assigned a lower uncertainty score and thus will overwrite the broader scale datasets within the domain of the local model.

8.2.4 Representing incomplete information

The method of combining probabilities must also deal with situations where there is incomplete information about the depth–probability curve:

- **Extrapolation** to depths/probabilities beyond the range of the input variables.
- Data sources associated with a **single data point** on the curve (e.g. depth at a single return period).

The method used to deal with uncertainty described in the previous section can also be used to deal with incomplete information.

Data input to the method is in the form of lists of probabilities and associated depths. Extrapolation is required when the probability or depth to be calculated lies outside the range spanned by this list. An example might be where data input is in the form of depth maps for 1% and 0.1% AEPs, and the depth for the 5% AEP is required. By assuming that the depth–probability curve is monotonic (i.e. depth always increases as exceedance probability decreases), we can make the following statements about probabilities for depths outside the input range:

$$\begin{aligned} 0 \le P(X > x) < P(X > x_{Max}) \quad & \textit{for } x > x_{Max} \\ 1 \ge P(X > x) > P(X > x_{Min}) \quad & \textit{for } x < x_{Min} \end{aligned} \tag{8.10}$$

Using the example above, this means that if at a point on the floodplain the 1% and 0.1% depths are 0.5 and 1 m respectively, then all we can say about the probability of the depth exceeding 0.2 m is that it is greater than 1% and less than or equal to 1. Similarly, the probability of exceeding 1 m is less than 0.1% and greater than or equal to zero. These inequalities are used to calculate upper and lower bounds when data are extrapolated, with an example shown in Figure 8.5.

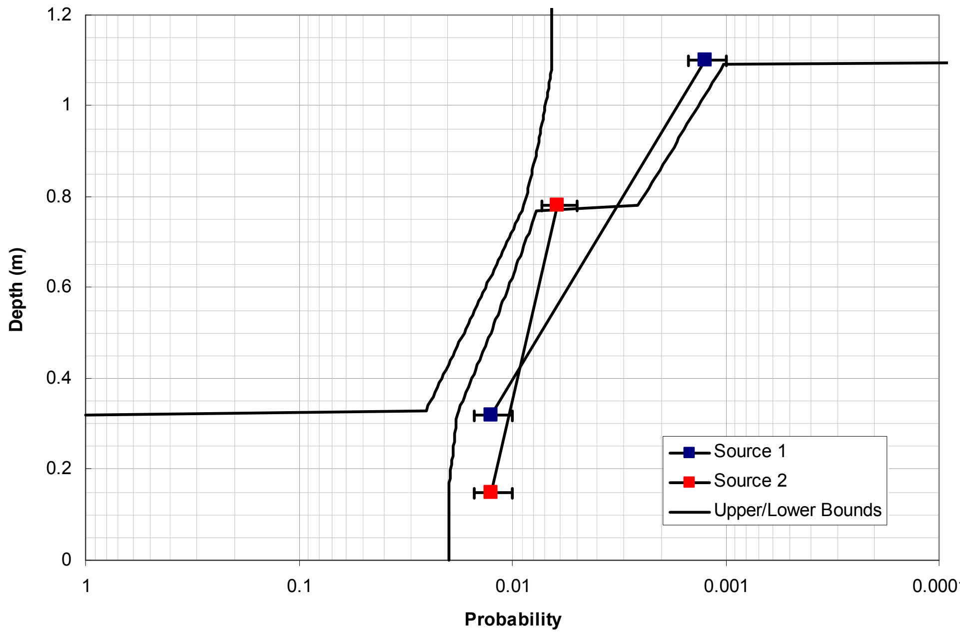


Figure 8.5 Addition of uncertain probabilities with upper and lower bounds used to represent extrapolation.

The same method can be applied to data with only a single data point representing the depth–probability curve (e.g. a 1% flood depth map). This is illustrated in Figure 8.6. Representing uncertainty in the probability associated with the single data point from uncertainty in the depth is difficult since there is no gradient information.

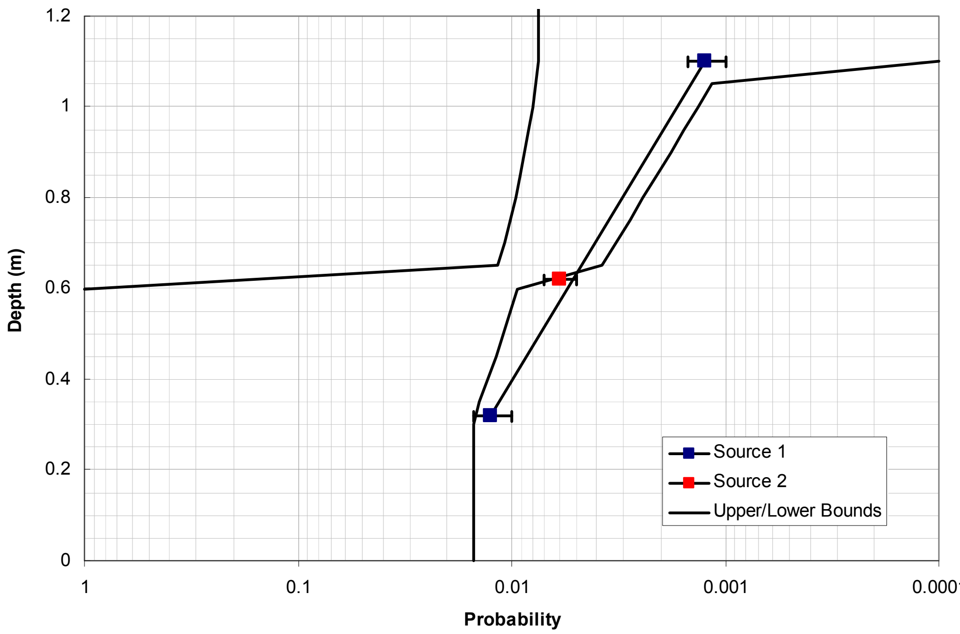


Figure 8.6 Addition of uncertainty probabilities with upper and lower bounds used to represent extrapolation from a single data point.

8.3 Representing source interaction and dependency

The method described in the previous sections assumes both that:

- Sources do not interact, so that the depth resulting from a forcing event (e.g. fluvial flood) is not affected by another forcing event (e.g. surface water event), unless the depth from the second exceeds the first.
- Sources are independent, so that the probability of two forcing events occurring at the same time can be ignored.

These assumptions allow depth grids to be combined without modelling any interaction between water from different sources. Modelling interaction requires detailed local data on joint probabilities and hydraulics, which at present have only been used in small-scale models. These effects may be important for some areas with high consequences, and the methodology described in this report allows the results of these local models to be included. To determine where local modelling may improve flood mapping, a method for determining where interaction and dependency may be significant is required.

The effects of the assumptions of no interaction and no dependency are strongly interlinked. If forcing events are more likely to occur together, their interaction may be more significant in generating flood risk. Nevertheless, interaction may be significant even for independent sources. An example is where it is extremely unlikely for sources alone to produce flooding above a depth threshold (say 2 m), but there is significant consequence to flooding above a higher threshold (say 2.5 m). The probability of depth exceeding 2.5 m will in that case mostly come from a combination of sources, rather than each individually.

Representing the effects of interaction and uncertainty analytically in a computationally feasible way is extremely difficult, perhaps impossible within the limitations of this project. Instead, we have identified the key effects of interaction and dependency, and develop an approximate method to represent these effects (Table 8.1). Within the framework developed here, if a more detailed understanding of source interaction and dependency is required, external modelling results can be brought in to represent this. External modelling should use the more rigorous methods described in Hawkes (2005).

Table 8.1 Effects of interaction and dependency to be represented by the method.

	Effect	Example
1	Interaction and dependency will be most significant for areas where probabilities from two (or more) sources are approximately equal. Where one source dominates, interaction and dependency have little effect.	Dependency and interaction between fluvial and coastal flooding is most significant in the estuarine zone. Up and down stream of that, one or the other dominates.
2	Effects will depend on threshold depths and probabilities.	For a large threshold depth, the only events capable of producing flooding may be from a combination of sources. In this case the interaction and dependency may be major factors in determining risk.
3	Effects vary spatially across England and Wales, and the dependency can be characterised quantitatively using numerical dependency measures.	Tide/surge levels and river flows are more highly correlated for catchments in the west, compared to catchments draining to the east coast.

Effect 1 in the table also includes situations where sources may combine to produce flooding for events that, if they occurred in isolation, would be too small to generate flooding. An example is locking of storm water sewers or small drainage channels by high tide or river levels during a local rainfall event. This will tend to occur in locations where there is a probability of surface water flooding with ground levels close to sea or river levels, and hence where there is also a probability of fluvial or coastal flooding. If this interaction is potentially significant, then external modelling should be brought in to represent the effect.

We make the following assumptions to simplify the method used to represent the interaction and dependency effects:

- The interaction is modelled as being simply additive. If forcing events from two sources occur at the same time, then the depth (or other hazard) is taken as the sum of the depths from the individual sources. This assumption will tend to overestimate depths. For example, increasing volumes of water filling a topographic depression will have less effect as depth increases, so doubling the volume will not double the depth.
- If we have a depth–probability relationship for each source, then the depth–probability relationship for the sum of the depths is taken as the sum of the two source curves (adding depths). This approximates what should strictly be a convolution of the two probability distributions by a single point.
- Interactions and dependencies between three or more sources are assumed to be represented by the interactions and dependencies between each pair. Areas where three or more source interactions are important will thus coincide with interactions between each source pair, and the method described here will identify these.

Using these assumptions, we can simplify the FD2308 method of Hawkes (2005) to give a computationally tractable problem.

The description of joint probability between two variables is simplified by considering the distributions of individual variables (the marginal distributions) and the joint distribution separately. The marginal distributions can be described using standard statistical models (Gumbel, Extreme Value, Generalised Logistic etc.) which capture their behaviour for extreme events, and may be conditional on a threshold being exceeded (as in the Peaks-over-Threshold method). To represent the dependency, variables are transformed to variables with simpler distributions (e.g. uniform or normal), and dependency between these transformed variables expressed. The joint probability information in this transformed space is known as a copula. One useful property of the copula is that it is invariant under transformation of the marginal variables, and is thus independent of the details of the marginal distributions used to describe the probabilities for the individual sources.

FD2308 defines two copulas for use in modelling dependency between flood sources. The revised χ model is defined for two uniformly distributed random variables u and v (obtained by transforming the marginal distributions), and a dependency parameter α:

$$P(U > u \text{ and } V > v) = \exp\left(-\left[(-\log u)^{1/\alpha} + (-\log v)^{1/\alpha}\right]^{\alpha}\right) \quad (8.11)$$

The parameter α can be obtained from the χ parameter through:

$$\alpha = \frac{\log(2-\chi)}{\log 2} \quad (8.12)$$

The bivariate normal model is based on marginal variables transformed to normal distributions, and is itself a bivariate normal distribution with zero mean, unit standard deviation for both variates, and correlation parameter ρ. The parameters χ and ρ are given in map form in FD2308 for a number of pairs of variables, allowing the combined probabilities to be calculated. FD2308 defines the copulas for forcing (e.g. tide/surge height and river flow) rather than at receptor level, but since the copulas are invariant under transformation of the marginal variables they will also be applicable to floodplain depths. The only condition is that the relationship between forcing and water depth is monotonic and increases with flood magnitude (i.e. depth increases for bigger events).

The first step in the FD2308 method is to define the copula using χ or ρ, and to select a probability contour in [u,v] space, for example corresponding to the 1% AEP. Examples of the χ copula are shown in Figure 8.7, and for the bivariate normal copula in Figure 8.8. The copula for χ=0 corresponds to independence between u and v, so that $P(U > u \text{ and } V > v) = P(U > u)P(V > v)$. For χ=1, u and v are totally dependent, and so $P(U > u \text{ and } V > v) = P(U > u)$ for u>v, and $P(V > v)$ for v>u. Figure 8.7 and Figure 8.8 illustrate how dependency, quantified by χ or ρ, makes most difference to probabilities near the diagonal of the plot. So for situations where u>>v, or v>>u, dependency has little effect on the probability, which accords with our intuitive understanding of joint probabilities in the first effect listed in Table 8.1. Dependency has the greatest effect for areas of the [u,v] space near the diagonal, where the probabilities of flooding from the two sources are approximately equal.

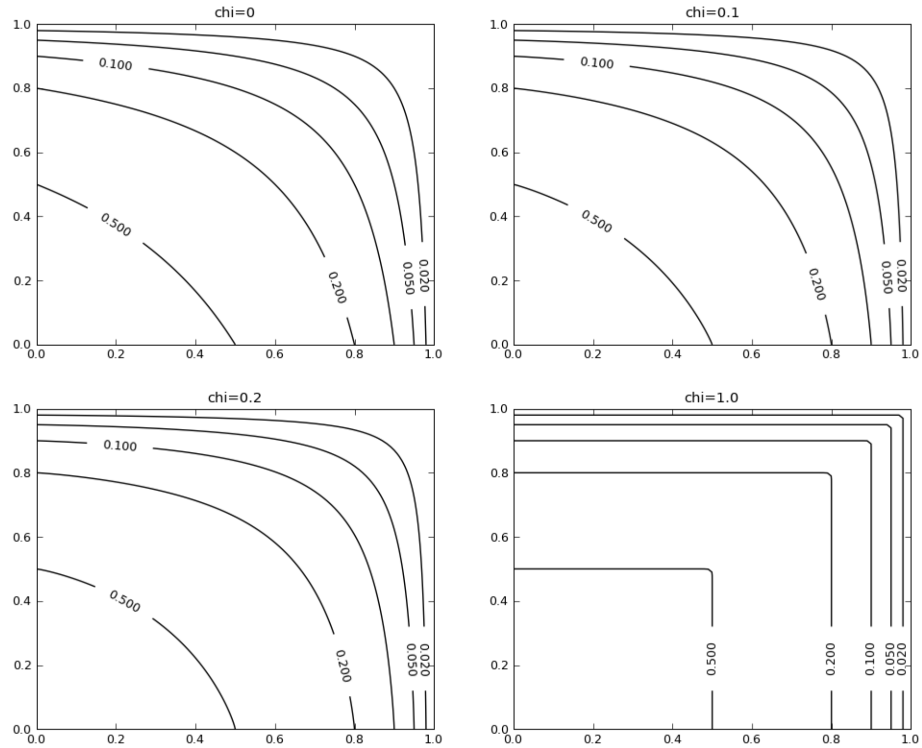


Figure 8.7 χ copulas for four different dependency parameters, drawn as contours of equal probability (labelled, not per cent).

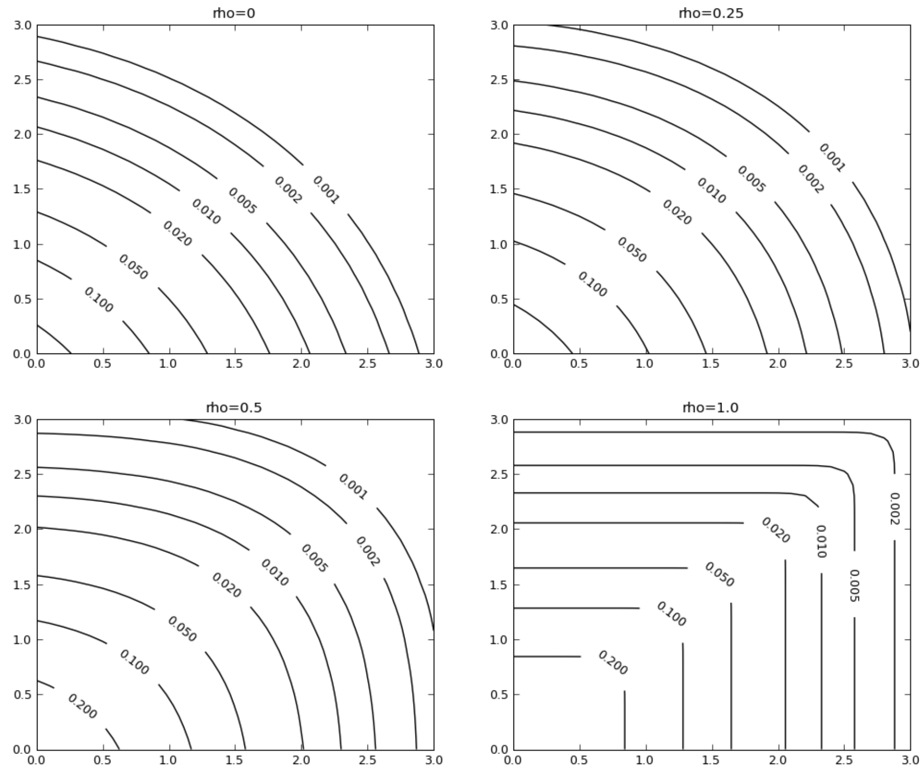


Figure 8.8 Bivariate normal copulas for four different dependency parameters, drawn as contours of equal probability (labelled, not per cent).

The second step of the FD2308 method is to sample a number of forcing event condition sets (e.g. tide/surge levels and river flows) from along the chosen probability contour (1% in our example) in the [u,v] space. These can then be used to drive a hydraulic model, and the maximum depth (or other impact) from these models at the site of interest is taken as the 1% AEP depth.

For this project, we simplify this approach by firstly assuming that the depths are additive, meaning that further model runs are not required. Secondly, we replace searching for the maximum depth along a probability contour with searching for the maximum probability along a given depth contour. Thirdly, we only search three points: one on each axis, and one on the diagonal. The diagonal point is chosen as this is where dependency has the most effect on the probability given by the copula. The relative magnitude of the probability values at these three points gives a measure of the significance of interactions and dependency. This is illustrated in Figure 8.9, along with the equivalent method from FD2308. The algorithm is described further below (with further detail in the pseudocode in Appendix J) for two sources:

1. Select a depth threshold.
2. Calculate probabilities of this threshold being exceeded by each source individually, P_1 and P_2.
3. Create a look-up table of $d_{sum}(u)=d_1(u)+d_2(u)$, where u is the marginal variable transformed into a uniform or normal distribution (according to the copula being used), and d_1, d_2 are the depths from the two sources.
4. Use this look-up table to calculate the value of u for which d_{sum} is equal to the depth threshold.
5. Calculate the joint probability from the copula function, P_{Joint}.
6. Compare the probabilities from steps 5 and 2, and classify according to:

   $\frac{P_{Joint}}{Max(P_1,P_2)} < 0.5$ Weak interaction/dependency

   $1 > \frac{P_{Joint}}{Max(P_1,P_2)} > 0.5$ Potential interaction/dependency

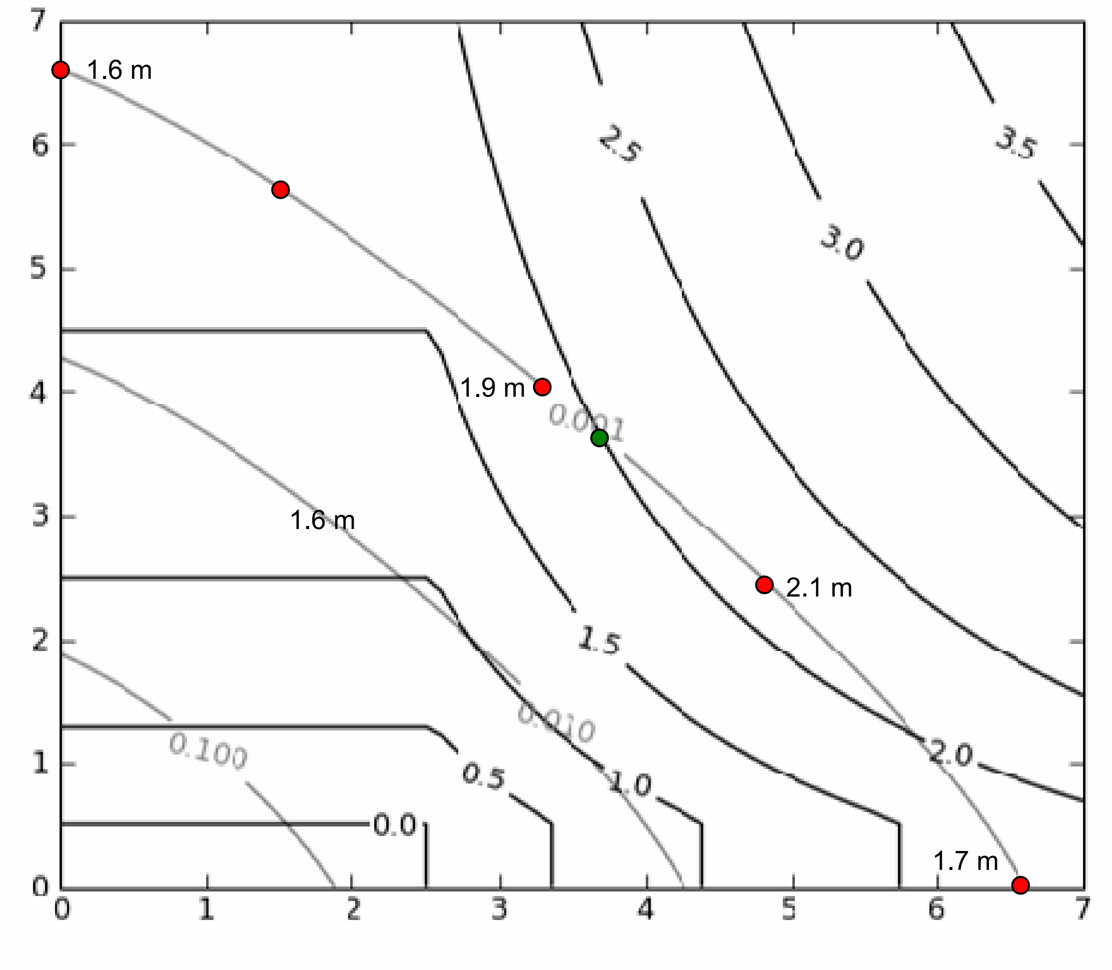
   $\frac{P_{Joint}}{Max(P_1,P_2)} > 1.0$ Likely interaction/dependency

7. When this classification has been calculated for each cell, display spatially.

The classifications in step 6 are essentially arbitrary, but the pilot testing described in Chapter 9 demonstrates that these are reasonable initial values. It will be a functional requirement of the software that the user can change these from the default values if necessary. Validation of the method through an assessment of how it represents the effects listed in Table 8.1 is described in the initial piloting in Chapter 9.

For dependency between three sources, the method described above is applied to each pair of sources. This is possible since the method does not actually affect the output probability, and hence the calculation can be performed independently for a number of source pairs. Where interaction and dependency between three sources may be important, it can be identified as interaction/dependency areas common to both pairs. This is a heuristic argument for multivariate dependencies between flood data, but little information on suitable techniques and their application to England and Wales is available.

The method described here relies on the most common approach to representing dependency between extreme events: factoring out the marginal distributions and then describing the uncertainty between more well-behaved variates (uniform and normal in this case). This means the method should be relatively easy to update as further dependency information becomes available in the future. Currently, we have only a limited understanding of the contribution that dependency makes to flood risk at different scales (national, regional, local). While the method developed for the prototype tool is limited in its treatment of interaction between sources, it will allow us to make a consistent assessment of how widespread the effects of dependency and interaction are likely to be, and hence whether future risk assessments need to include more detailed approaches to mapping these effects.



FD2308 Method

1. Select probability contour, e.g. AEP=0.1%
2. Sample points along this contour (red dots)
3. For each of these points, calculate the equivalent values (e.g. tide/surge, flow) from marginal distributions
4. Use these as boundary conditions in model to calculate depths (shown near red dots)
5. Take maximum depth as representative of 0.1% AEP = 2.1 m

FMAS Method

1. Select depth for which probability is required (e.g. 2 m)
2. Find intersection of diagonal and 2 m contour (green dot)
3. Find joint probability value of this point from copula (0.1%)
4. Compare this value with equivalent marginal distributions – if of similar magnitude, dependency and interaction may be important

Figure 8.9 Method for calculating relative significance of interaction and dependency. The copula function is plotted here in grey, against reduced variates to show extreme values better. The depths for interacting sources is shown in black, representing the result of summing the depths obtained from the marginal distributions.

8.3.1 Issues not addressed by this methodology

This report represents the current state of the methodology, which is ready to be developed into a prototype tool. There remain some issues that need to be resolved in Phase 2 of the project as the prototype tool is developed and piloted:

- The method gives no indication of whether the combined exceedance probability or the extrapolated value is a better (i.e. more useful) estimate of combined probability. This is being left as a choice for the user, but will remain a subjective decision. Care will be required to inform users (who may be accessing outputs 'second hand' rather than using the tool themselves) about the appropriate level of information that can be inferred from the tool's outputs. We propose that the extrapolation method is used as a default, as this has been shown to give better results for the Hull pilot site, but this will need to be reviewed after more piloting in Phase 2. Both the combined exceedance probability and the extrapolated value must be viewed alongside the uncertainty estimates, which will guide the user into viewing the results with an appropriate level of confidence.
- Upscaling data to appropriate level or averaging probability over a larger cell.
- The classification thresholds for potential and likely interaction need to be confirmed as appropriate.

There are further limitations imposed by available data or methods:

- Uncertainty information may not be available for some sources, even for 'probabilistic' model outputs such as NaFRA. Future projects will supply information on uncertainty for some model outputs (such as SC090008/WP1 Validation and Calibration of Probabilistic Flood Models) which can be used with the method developed for this project. Other uncertainties, such as those associated with national surface water mapping, may be more difficult to quantify.
- Dependency data is not available for some pairs of sources, most significantly local rainfall and fluvial flows. This is a complex issue linked to spatial–temporal rainfall behaviour and catchment hydrology across a broad range of scales. SC060088 'Spatial Coherence of Flood Risk' may provide some information on dependency between flows in large and small catchments, which may be useful as a proxy for fluvial flows and local rainfall. It is also investigating the dependency between fluvial and coastal flooding, which may be useful in the future in mapping the effects of dependency.
- Explicitly representing dependency between three or more sources. A heuristic method has been described in this section, but a more rigorous approach may be required for future versions of the tool. The tool as specified here will at least give some indications of areas where three-or-more-way interactions may be significant, and hence whether a more sophisticated approach is required.
- The method cannot deal with interaction/dependency for data sources represented by a single depth–probability grid (e.g. current national surface water mapping). More generally, single depth–probability grids give limited information (as illustrated in Figure 8.6) when combined with other sources. A fuller description of the depth–probability curve is required to make the most of combined flood probability information.

- The method concentrates on uncertainty in depths and probabilities, rather than in terms of spatial uncertainties (i.e. uncertainty in flood extent). For some input data types the method will represent spatial uncertainty. For example, if a grid represents the ‘upper bound’ water depth and this includes a greater flood extent than the ‘best estimate’, the difference between the two extents can be represented spatially as a band of uncertainty.

9 High level design and initial piloting

9.1 User requirements

User requirements have been obtained from four main sources:

- Project workshop (Appendix G)
- Essential requirements (Chapter 2)
- Method description (Chapter 8)
- Input/output data requirements (Chapter 7)

These sources have been used to draw up Table 9.1, which lists user requirements and the function features that fulfil those requirements. These requirements are detailed further in Chapter 10, and will be discussed with Environment Agency CIS representatives during Phase 2 of the project.

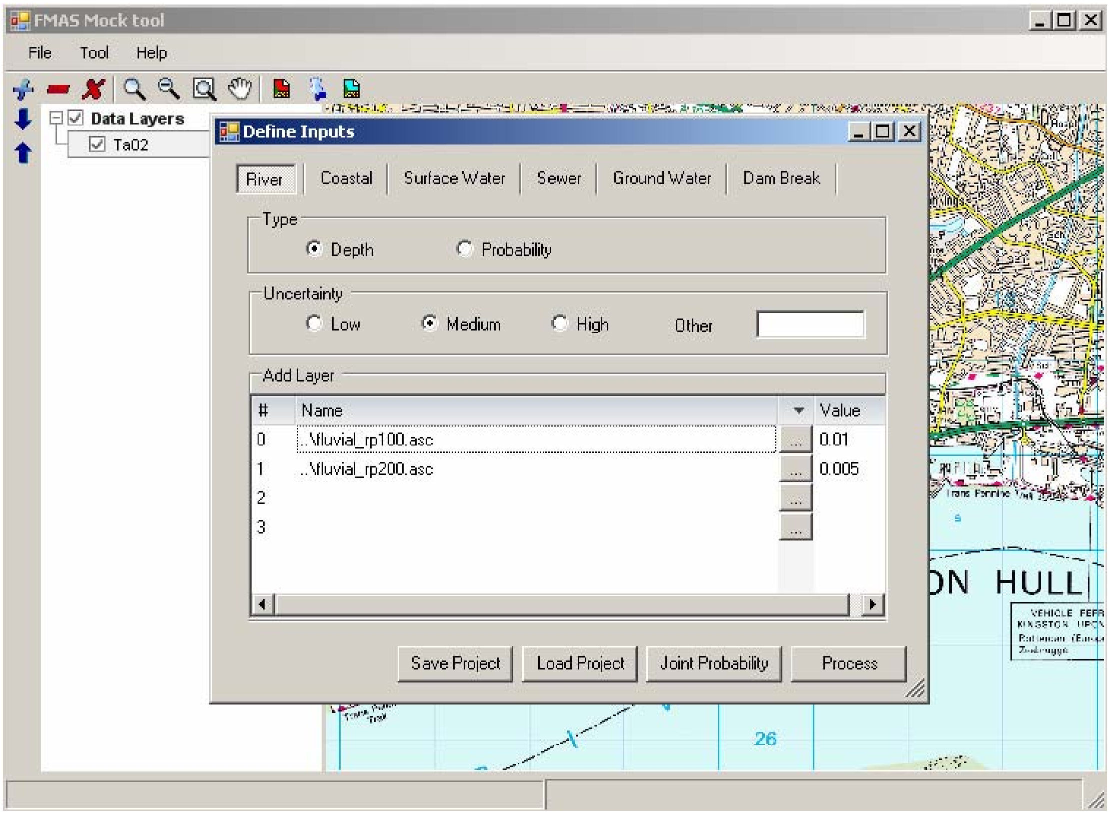
Table 9.1 User requirements and functional features (draft).

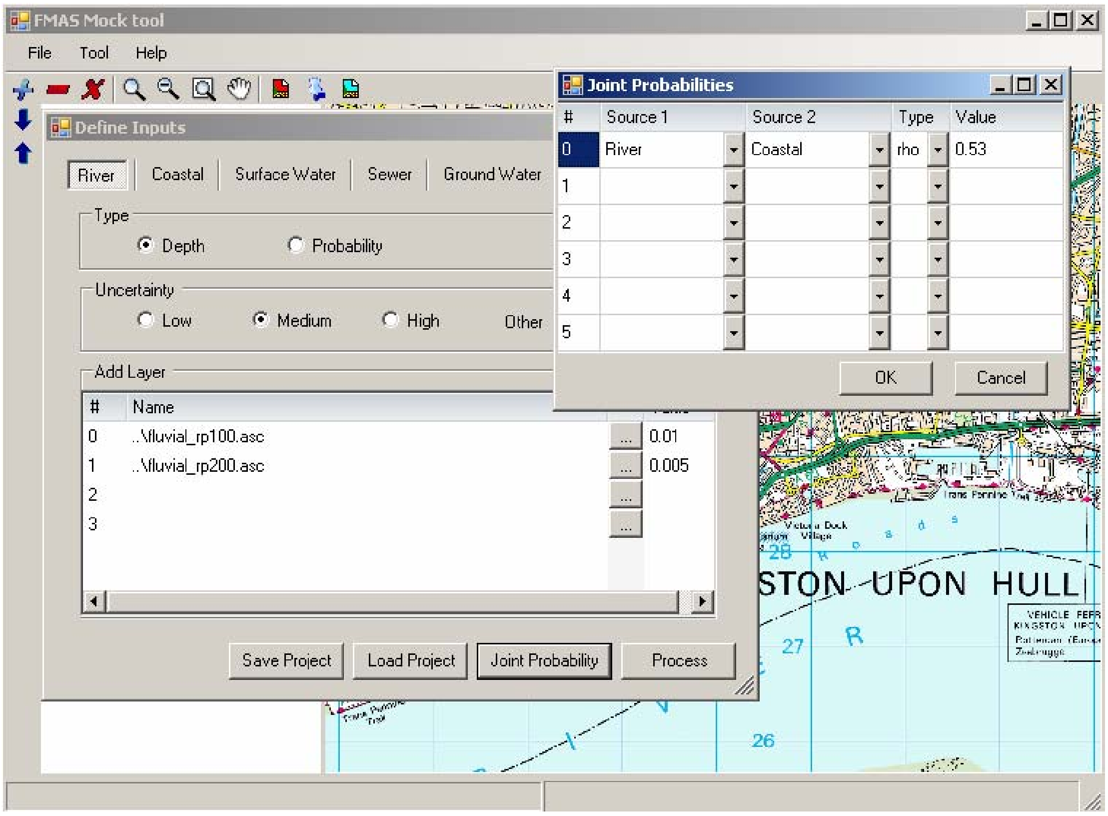
User requirements *What does the user want to do?*	**Functional requirement** *What features of the tool allow the user to do this?*
Compatibility with other software data formats.	Load or pre-process a range of spatial data formats: ascii grids, arc grids, shapefiles, MIF/MID etc. Input formats to be determined from analysis of datasets. Output in standard formats that can be used in other tools.
Produce a common national standard including legends, formats, attributes.	Mostly provided by post-processing systems although the tool should generated appropriate required attributes.
Accommodate differing grid sizes and resolutions.	Input files at different resolutions and extents transparently (as in GIS software).
Capability to produce maps for a diversity of flood variables, e.g. depth, duration, velocity, hazard to people.	Make input/output and processing generic so that different hazards can be processed in the same way (but focus on flood probability).
View spatial patterns of combined flood hazard.	Display a single estimate of combined probability through, for example, colour scale.
View details of combined flood hazard (e.g. uncertainty, contribution from sources).	Display available uncertainty information and contributions at selected points, communicating more information than possible in a single map.
Produce appropriate mapping, which can be used as a basis for appropriate communication of flood risk information to a range of stakeholders to support high level flood risk messages, and more detailed information at a local scale.	Display different levels of information tailored to/by different users, either directly or through output to other systems (e.g. web server). This will affect both scale and content.
The tool should be capable of integrating data of different formats that arise when	Load or pre-process different types of data, e.g. deterministic, probabilistic and qualitative. This refers to the data content, rather than its

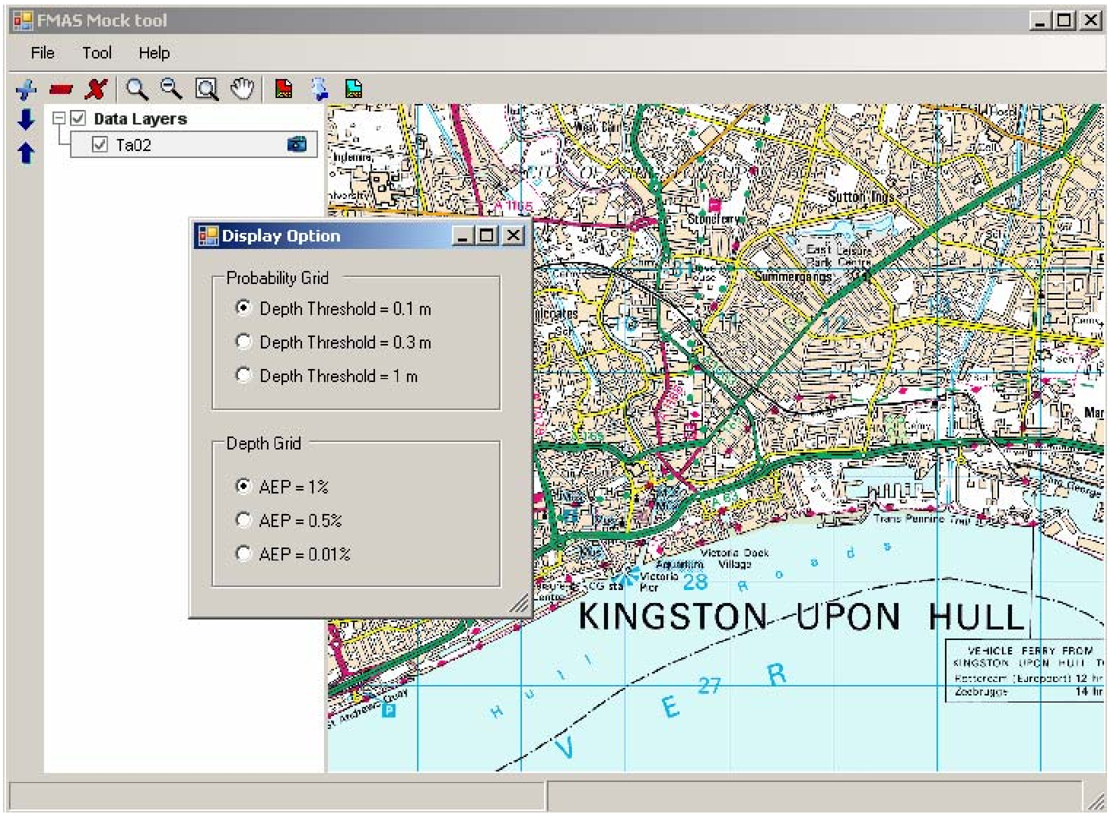
User requirements *What does the user want to do?*	**Functional requirement** *What features of the tool allow the user to do this?*
different sources of flooding are modelled.	format.
Flexibility to include future datasets and formats.	Inclusion of standard formats should cover this. New formats should be convertible to current formats as a temporary 'work around' as new formats are produced.
Updating flood hazard estimates as the climate, land use and flood risk management responses change.	Process data sources in the future in the same way as current datasets.
Dependency between sources should be considered and represented. Where joint probability information is available, the method should be able to use it to improve the estimates of flood hazard in the combined map.	Tool(s) to include features to use current understanding of and information on dependency (e.g. FD2308 measures). This will involve input of dependency measures (uniform and spatially varying) and modified calculation of probabilities.
The prototype tool should be acceptable to its end users as software, having been tested and trialled for usability and compatibility.	Tool to meet a set of well-defined criteria for acceptance. The software will be tested against these criteria in Phase 2.
The output should simplify, not complicate, the assessment of risk from all sources.	Capability of outputting a simple view of flooding from all sources.
The method must include ways of dealing with our incomplete knowledge of flood probabilities presented by the input data, and generate some form of uncertain probabilistic output or indicator of uncertainty.	Allow input of uncertainty information: uniform, spatially varying, related to depth or probability, qualitative. Input uncertainty either integrated with probabilistic calculation or piped through tool to output stage. Represent situations where uncertainty in combined maps is large (i.e. little information content).
The tool must allow some local, specialised data to be included (e.g. TE2100 depth–probability grids). Although the tool would not restrict subsequent use of the mapping (e.g. at an individual property scale), the scale at which the outputs will be suitable for use will depend on the quality of all input data for that geographical location.	Include data with local extent, sampled to appropriate scale. Prioritise use of this data in the calculations where it overlaps with national/catchment scale data.
The output of the tool may provide an indication of the proportional contribution of different sources to total flood hazard.	Communicate proportion of risk from each source either on map (may overcomplicate) or point by point.
Decide where to use local data for improved accuracy.	Include criteria to identify and communicate areas where low confidence and high consequence coincide.
Decide where to use local data to represent dependency.	Include criteria to identify and communicate areas where dependency and high consequence coincide.
Data requirements specific to sources to be included in the tool.	Tool must accept the data types and formats for flooding from fluvial, coastal, estuarine, surface and breaching sources at national scale. Discussed further below.

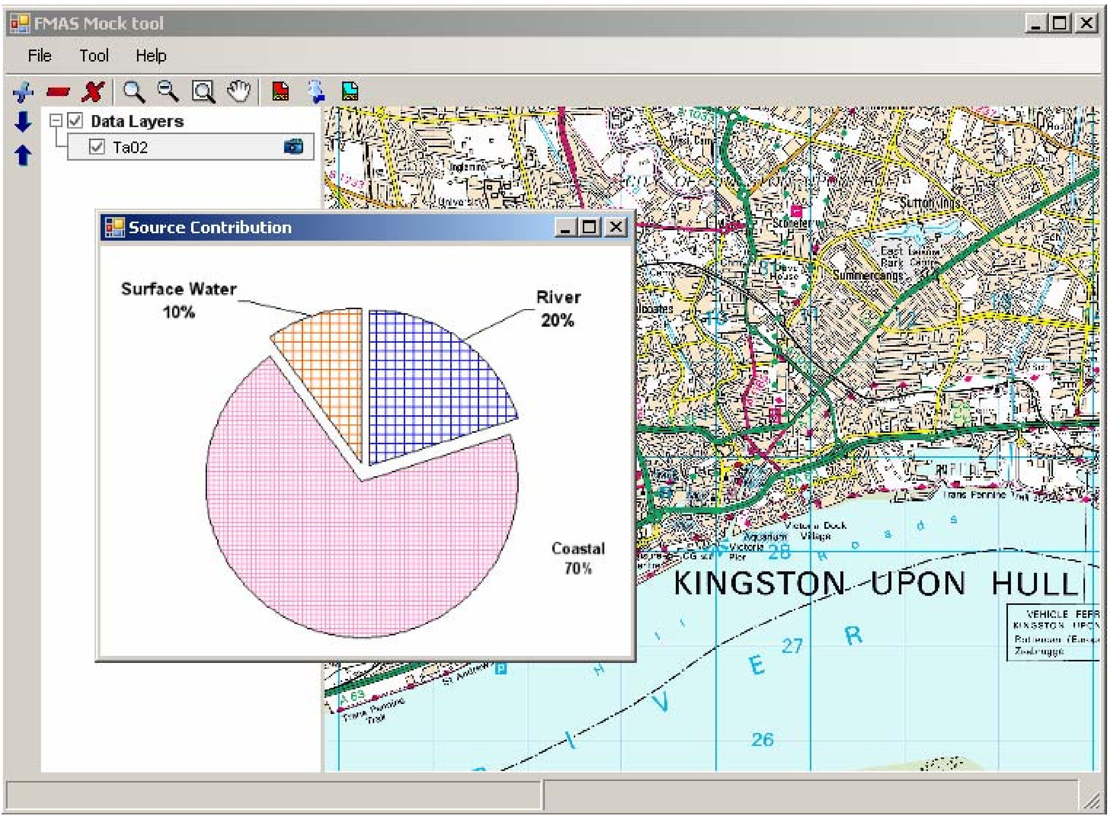
9.2 Mock-up demonstrations

The requirements analysis above has been used to develop a mock-up of the prototype tool. Selected 'screenshots' of the mock-up are provided below. Note that this is *not* the proposed design of the user interface, rather it is a mock-up to facilitate discussion and analysis of what the user interface should look like. The design of the user interface is a Phase 2 task.






FMAS Mock tool
File
Tool
Help
Data Layers
Ta02
Display Option
Probability Grid
Depth Threshold = 0.1 m
Depth Threshold = 0.3 m
Depth Threshold = 1 m
Depth Grid
AEP = 1%
AEP = 0.5%
AEP = 0.01%
KINGSTON UPON HULL


FMAS Mock tool
File
Tool
Help
Data Layers
Ta02
Source Contribution
Surface Water
10%
River
20%
Coastal
70%
KINGSTON UPON HULL

9.3 Initial piloting

The methods described in Chapter 8 have been piloted on data available to the project for Hull. Hull is at risk from coastal, fluvial and surface water flooding, and the effects of reservoir breaching have been modelled using a hypothetical dam location.

The data sources used for the initial piloting are:

Tidal flooding: Outputs from a standard flood risk assessment study mapping Flood Zones 2 (0.1% AEP) and 3 (0.5% AEP). These have been modelled without defences or barriers in place.

Fluvial flooding: Outputs from the River Hull flood risk management strategy study, modelled using ISIS 1D without defences in place, for the 1 and 0.5% AEP events.

Surface water flooding: Flooding events with 1 and 0.5% AEPs have been modelled using direct rainfall onto TUFLOW models. Separate models for urban and rural areas are used, and rainfall depth equivalent to the capacity of the urban drainage system was subtracted from the rainfall inputs for the urban model.

Reservoir dam breaching: A fictitious reservoir was assumed in the north-east of the city. ISIS 2D software was used to simulate the flood propagation resulting from a hypothetical dam breach.

9.3.1 Independent sources: tide/surge and fluvial

The results of combining probabilities for tide/surge and fluvial sources are shown in the maps in Figure 9.1. Figure 9.2 shows detailed depth–probability curves for three test points. All the test points show how the method only predicts narrowly defined upper/lower bounds for a limited range of depths, determined by the values in the input depth grids. The upper bound is especially lacking in information; for small depths it reverts to 1, indicating only that the probability of flooding is less than 1.

The high upper bound reflects the information carried in the depth grids inputs. For much of the study area, depths are greater than 0.3 m for the 1% and below AEP grids input. The only conclusion that can be drawn from the data is that the probability of depth exceeding 0.3 m is greater than 1% AEP, and less than 1. The areas with high upper bounds can be reduced by using more input depth grids or by using probability grids instead, which give a clearer picture of threshold exceedance probabilities.

While more tightly defined bounds would be preferable over a greater range of depths, this cannot be justified from the input data. This highlights the need for input data spanning an appropriate range of depths or probabilities if precise results are required from the method. The uncertainty estimates derived from the method may provide some guidance on the extra data (e.g. from further modelling and mapping) required to reduce the uncertainty, although this will not be straightforward due to spatial variability in the probability and uncertainty estimates.

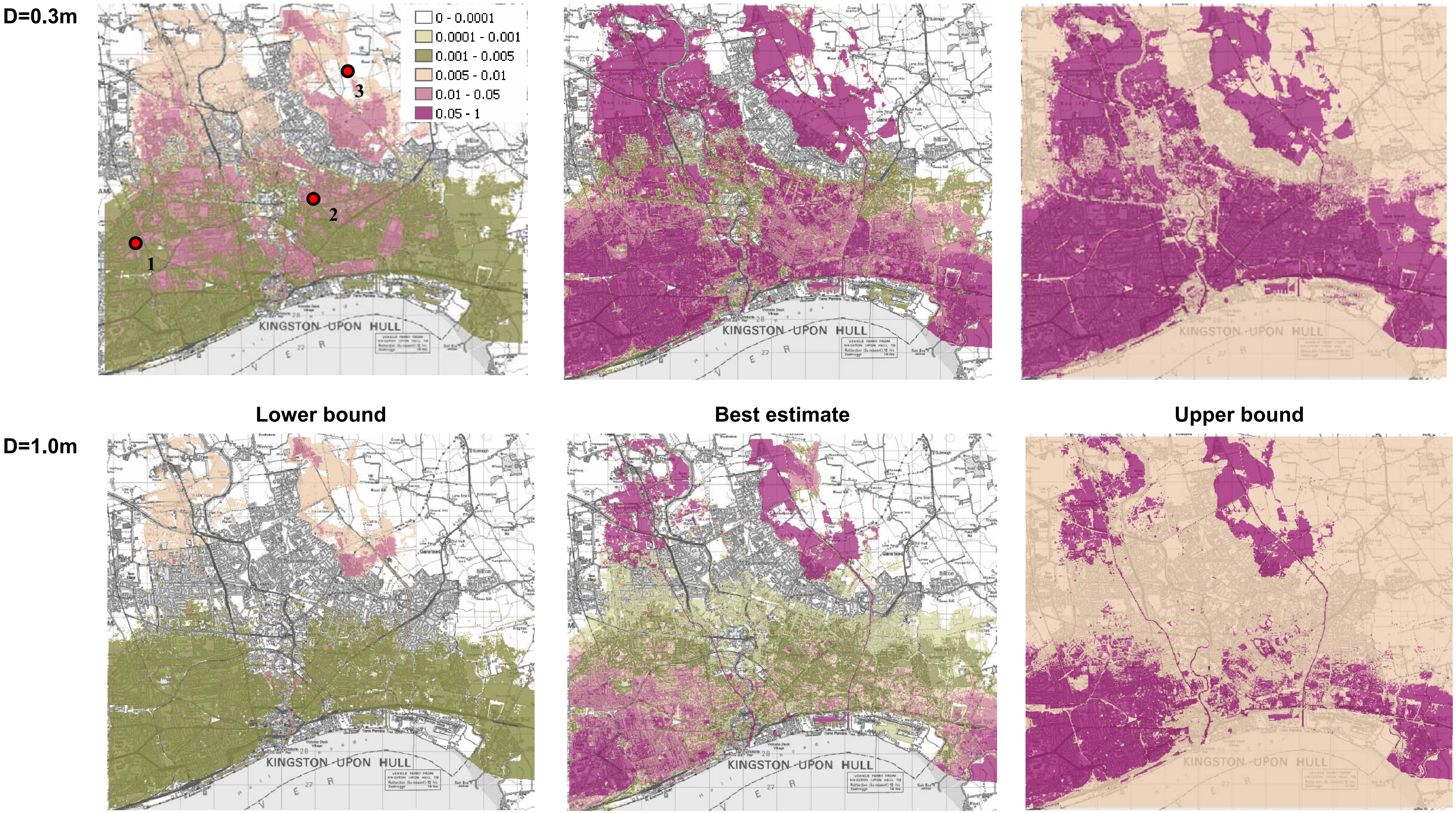
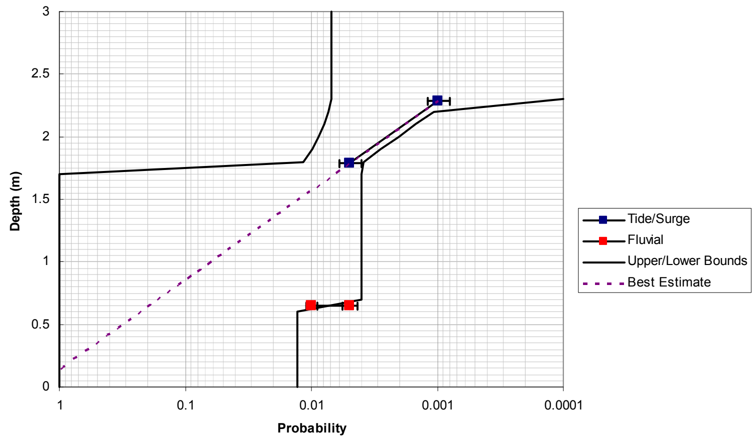
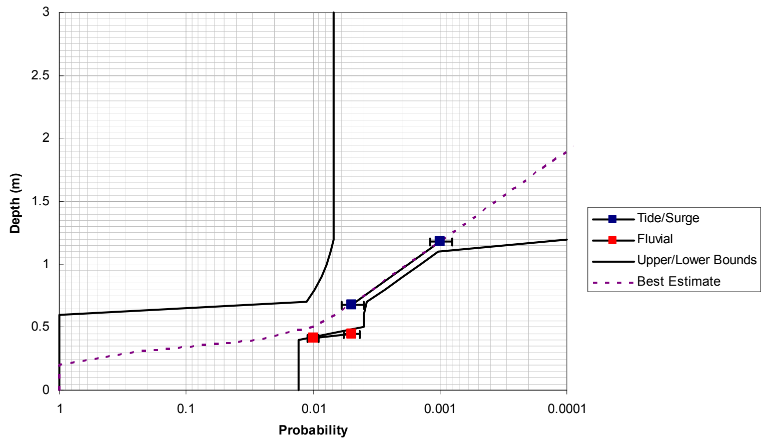


Figure 9.1 Combined probabilities for tide/surge and fluvial sources for depth thresholds 0.3 m and 1.0 m. The 'best estimate' is the extrapolated result (see Section 8.2.3). Locations of test points with depth-probability plots in Figure 9.2 are shown as red dots in the first plot.

Test point 1



Test point 2



Test point 3

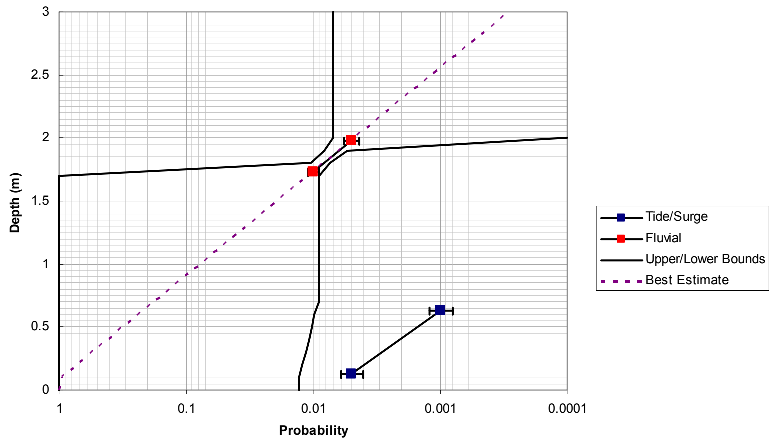


Figure 9.2 Combined probabilities for tide/surge and fluvial sources for test points, with 'best estimate' derived by extrapolation.

9.3.2 Independent sources: tide/surge, fluvial and surface water

The results of combining probabilities for tide/surge, fluvial and surface water sources are shown in the maps in Figure 9.3. Figure 9.4 shows detailed depth–probability curves for three test points. Again the test points show how the method predicts widely spaced upper/lower bounds for depths away from those given in the input data. The inclusion of three sources has further widened the upper/lower bounds for some depths. The spread between upper/lower bounds is expected to increase as more sources are added.

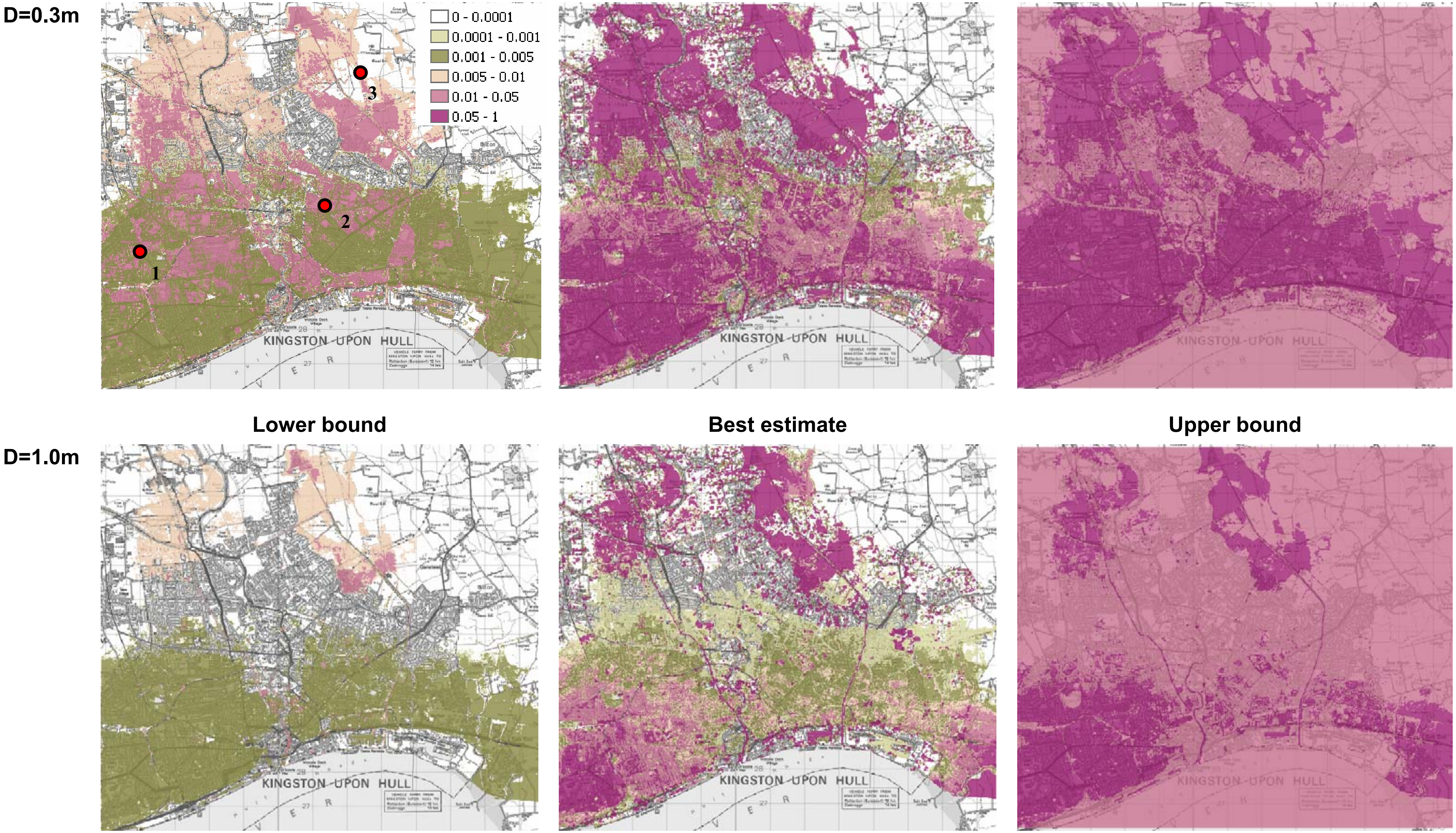


Figure 9.3 Combined probabilities for tide/surge, fluvial and surface water sources for depth thresholds 0.3 m and 1.0 m. The 'best estimate' is the extrapolated result (see Section 8.2.3). Locations of test points with depth–probability plots in Figure 9.4 are shown as red dots in the first plot.

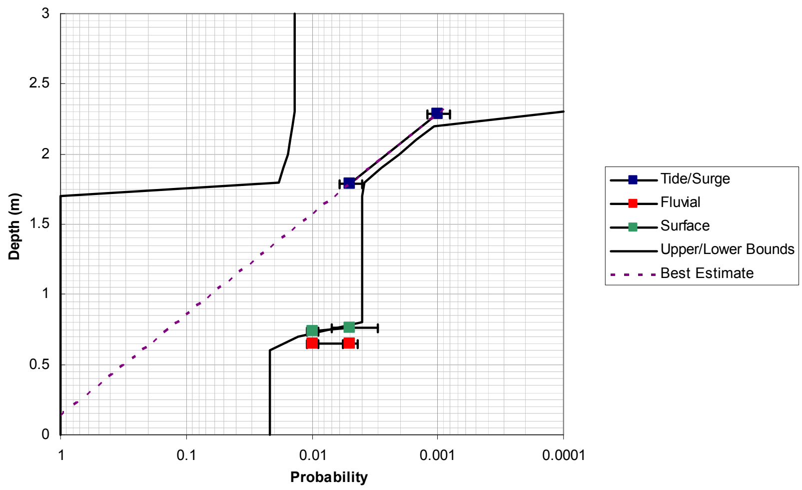
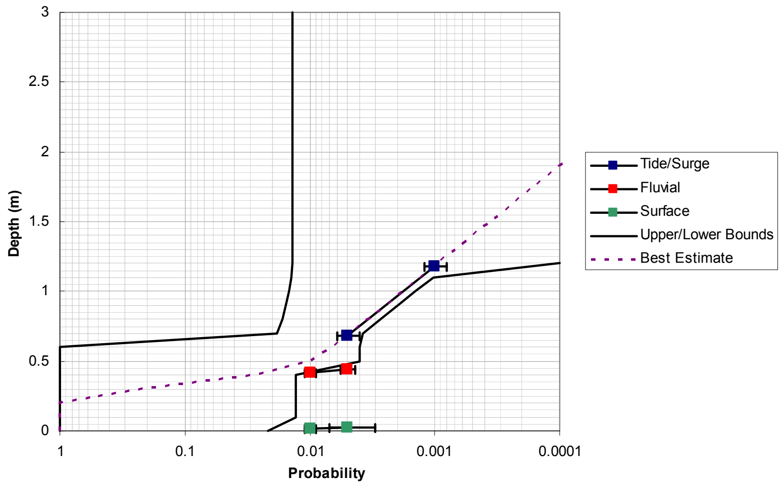
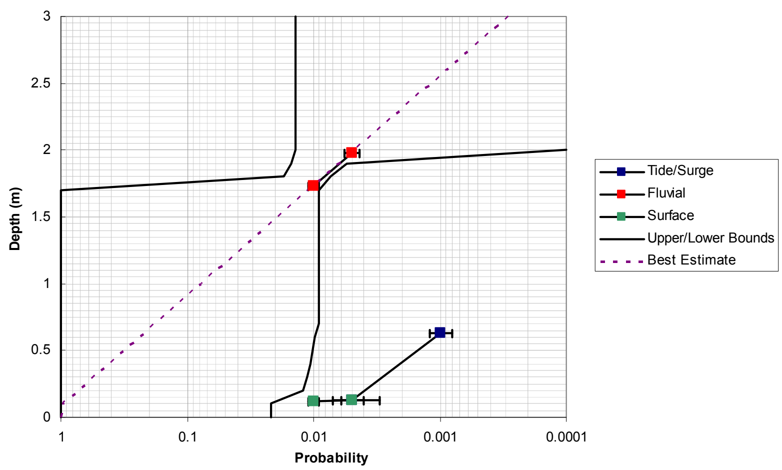
Test point 1	
Test point 2	
Test point 3	

Figure 9.4 Combined probabilities for tide/surge, fluvial and surface water sources for test points, with 'best estimate' derived by extrapolation.

9.3.3 Independent sources: fluvial and reservoir breach

Testing of the method with a reservoir breach source shows that because of the low probability of reservoir breaching (~0.01% AEP), including this source has little effect on probabilities. The depth–probabilities thus reflect the same relative risks as described in Section 5.7.

Figure 9.5 shows upper and lower bound probabilities maps for fluvial and reservoir breach sources.

Figure 9.6 shows the results as depth–probability plots for the three test points shown in Figure 9.5. The single data point on the depth-probability curve available for reservoir breach flooding may be associated with two uncertainties: for the depth and for the probability of failure.

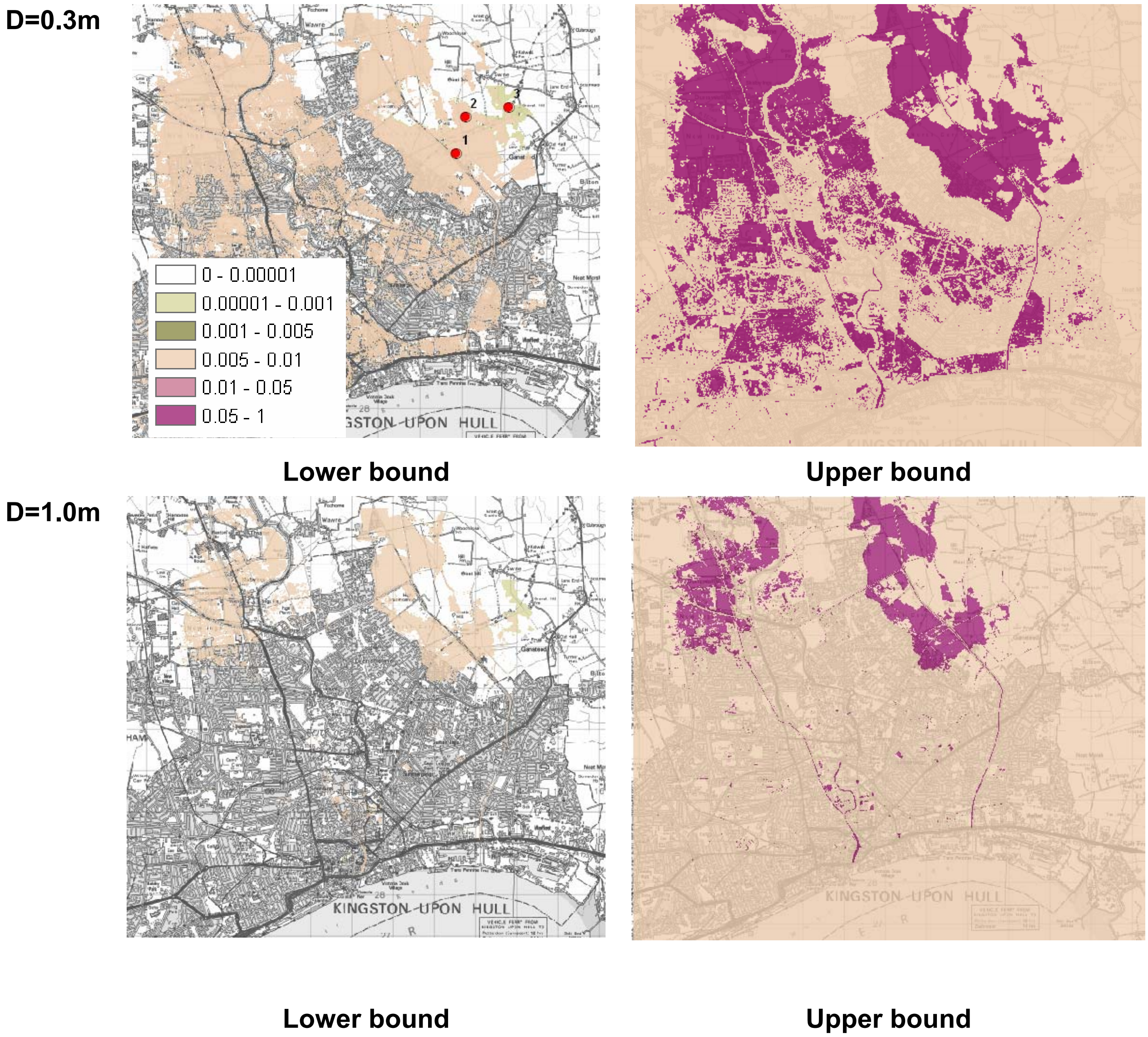


Figure 9.5 Combined probabilities for fluvial and reservoir breach flooding.

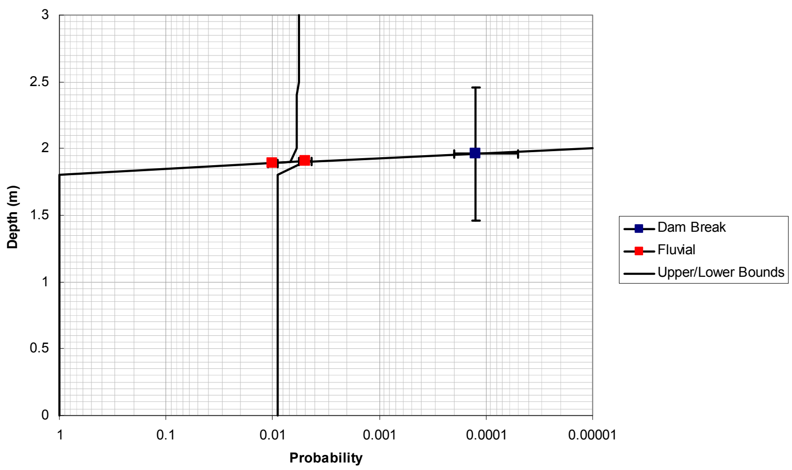
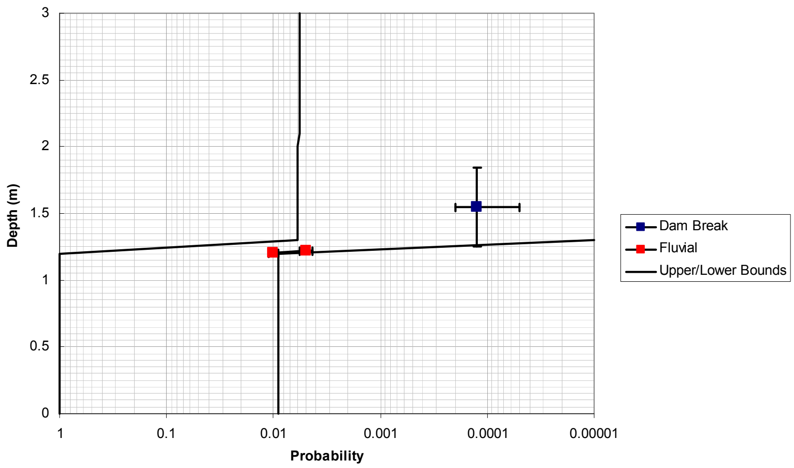
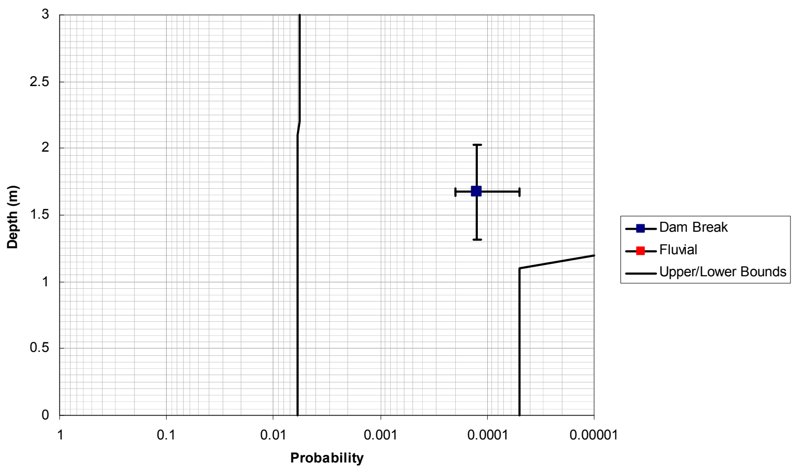
Test point 1	
Test point 2	
Test point 3	

Figure 9.6 Combined probabilities for fluvial and reservoir breach flooding for test points.

9.3.4 Interactions and dependency: tide/surge and fluvial flows

The method for identifying areas affected by source interaction and dependency has also been tested on Hull, using combinations of tidal/fluvial data, and tidal/surface water data. Both of these source combinations have dependency parameters which are derived in FD2308. Hull has large, flat, low-lying areas which are at risk of flooding from the Humber estuary, the River Hull and surface water.

For the case of tidal/fluvial interactions, FD2308 characterises the dependency between flows on the River Hull and tide/surge levels as 'independent', with $\chi<0.010$ (see Figure 9.7). To provide an effective test of the interaction/dependency method, we have instead tested using $\chi=0.125$ ('strongly correlated') and $\chi<0.25$ ('super dependent'). These values are more appropriate for catchments draining to the west coast, but have been used here to show how the method will work for more strongly dependent data than available for Hull.

For the $\chi=0.125$ ('strongly correlated') case (Figure 9.8), there are only small areas with the potential to be affected by interaction and dependency. These areas are different according to the depth threshold applied.

These results have been verified for three test points (locations shown in Figure 9.8) by analysing the marginal distributions and copula function for each point. Figure 9.9 shows the results of this analysis. The cumulative joint probability contours are shown on the left, plotted on reduced variate axes, along with contours showing the sum of the individual source depths at those probabilities. The plots on the right hand side show depth–probability curves for the two sources (blue and green), the joint probability depth curve for the sum of these depths (red), and the values calculated at each depth threshold by the approximate method (blue dots). For test point 1 we would expect interaction/dependency to be insignificant, since tide/surge flooding dominates, and this has been correctly identified by the approximate method.

Test point 2 illustrates a situation where the method has not worked well because of the specific properties of the flood depth grid data at this point. The probability of flooding is high for the fluvial source (blue curve), and this means that when tide/surge flooding occurs it is highly likely to be coincident with fluvial flooding. This situation may be physically unrealistic and due to the way the modelling has been carried out. Areas such as these highlighted by the method are worthy of further investigation.

Test point 3 illustrates a situation where the interaction is potential/likely for 0.3 m and 2.0 m thresholds, but not for 1.0 m. Flooding for low depths is extremely likely (as for test point 2), meaning interaction may be important, and for higher depths, the probability of exceeding the threshold is similar for the two sources, which is generally identified with potential for interaction and dependency to have an effect. This is reflected in the results of the approximate method, identifying potential interaction at 0.3 m and 2.0 m, but not at 1.0 m thresholds.

Figure 9.10 shows how the area around test point 3 is classified as an area of 'likely interaction/dependency' if the dependency parameter is increased to 0.25 ('super dependent'). This shows how the approximate method will increase areas of potential and likely interaction/dependency as the dependency parameter increases. This gives some confidence that when applied to other test sites (on the west coast for example) in phase 2 of the project, the output will reflect the greater effects of dependency in these locations.

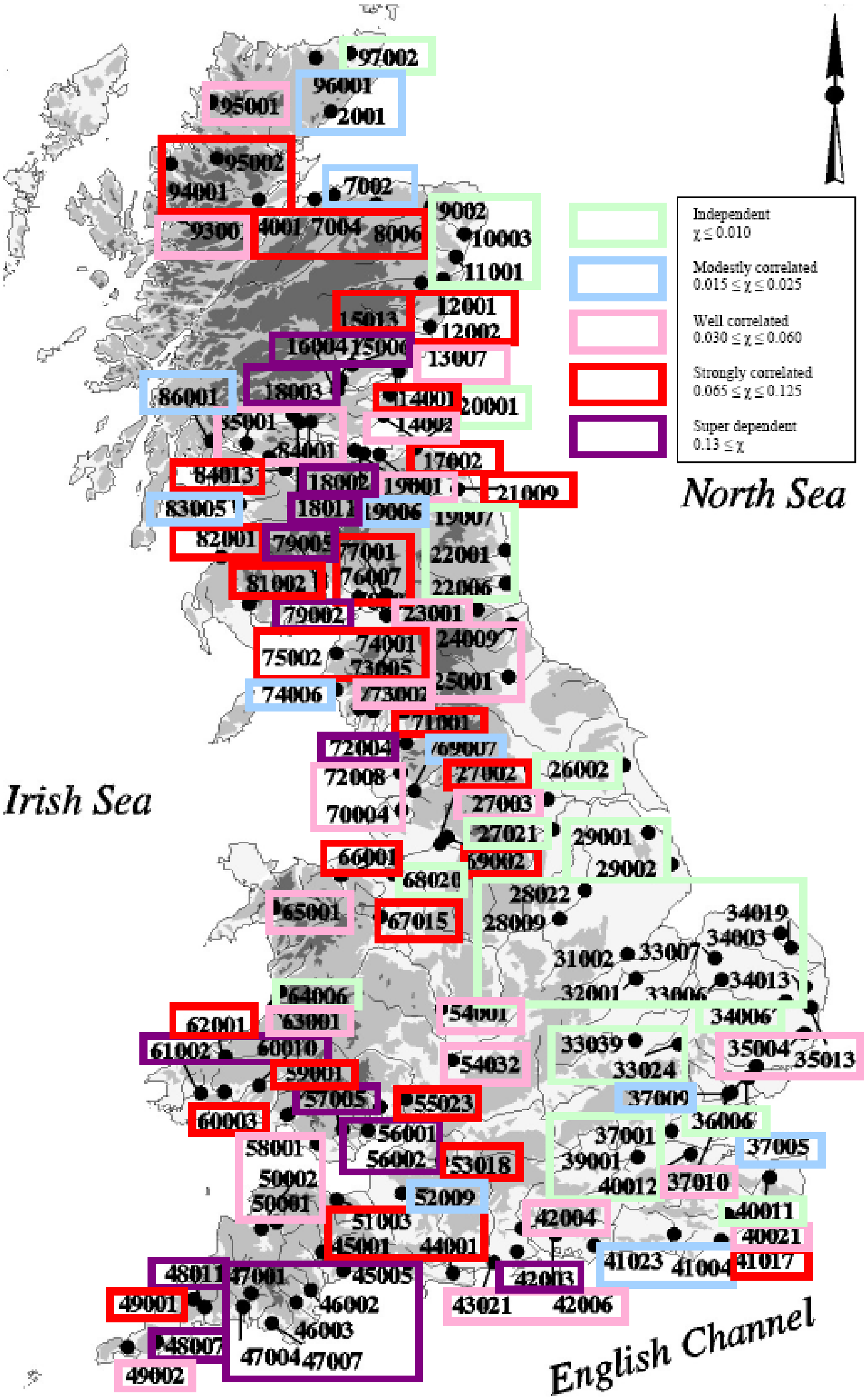
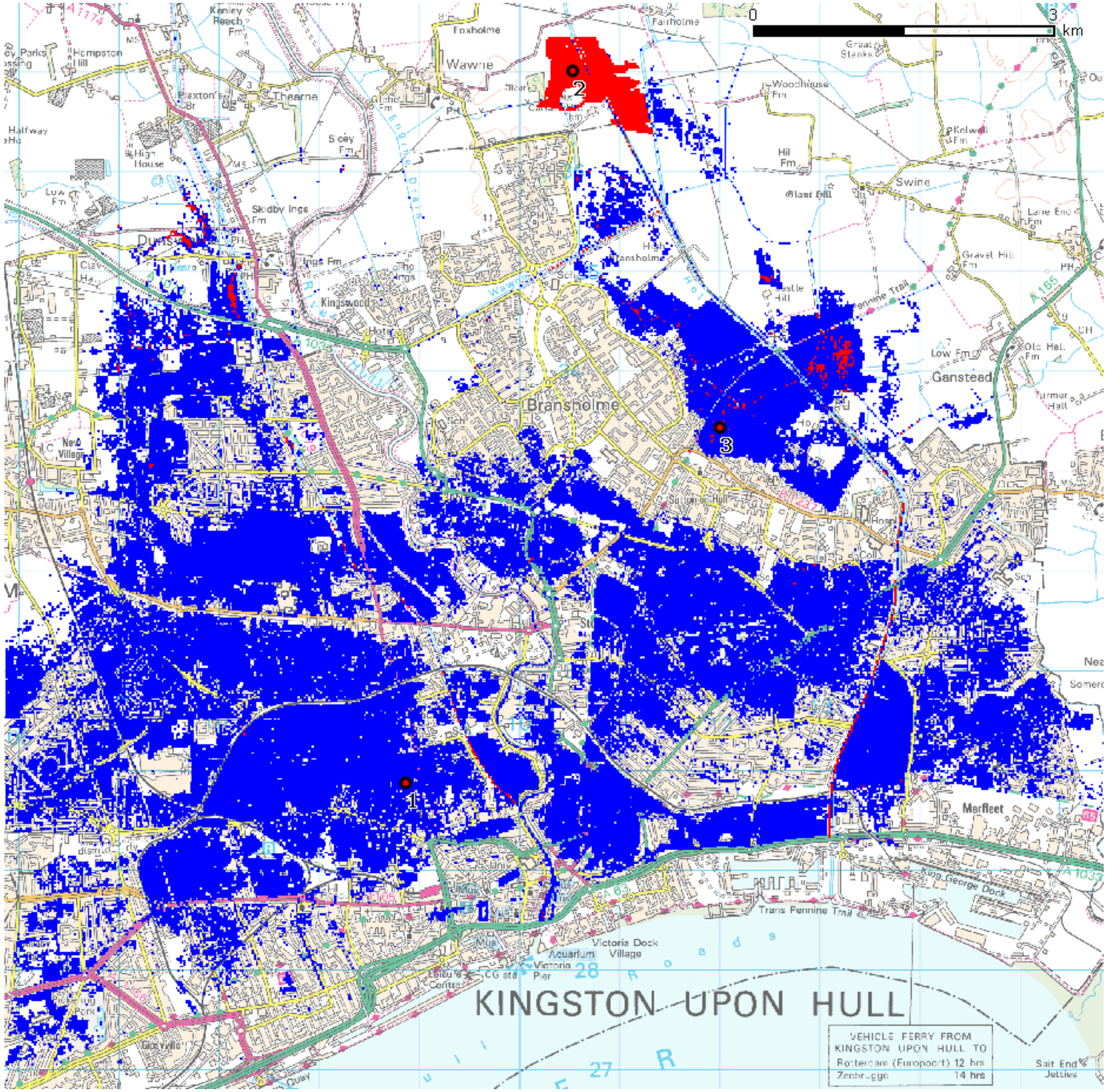


Figure 9.7 Dependency measure χ between tide/surge and fluvial flows, taken from FD2308.

Figure 9.8a

Depth threshold=0.3m



Blue	Weak interaction/dependency
Red	Potential interaction/dependency
Green	Likely interaction/dependency

Figure 9.8b

Depth threshold=1.0m

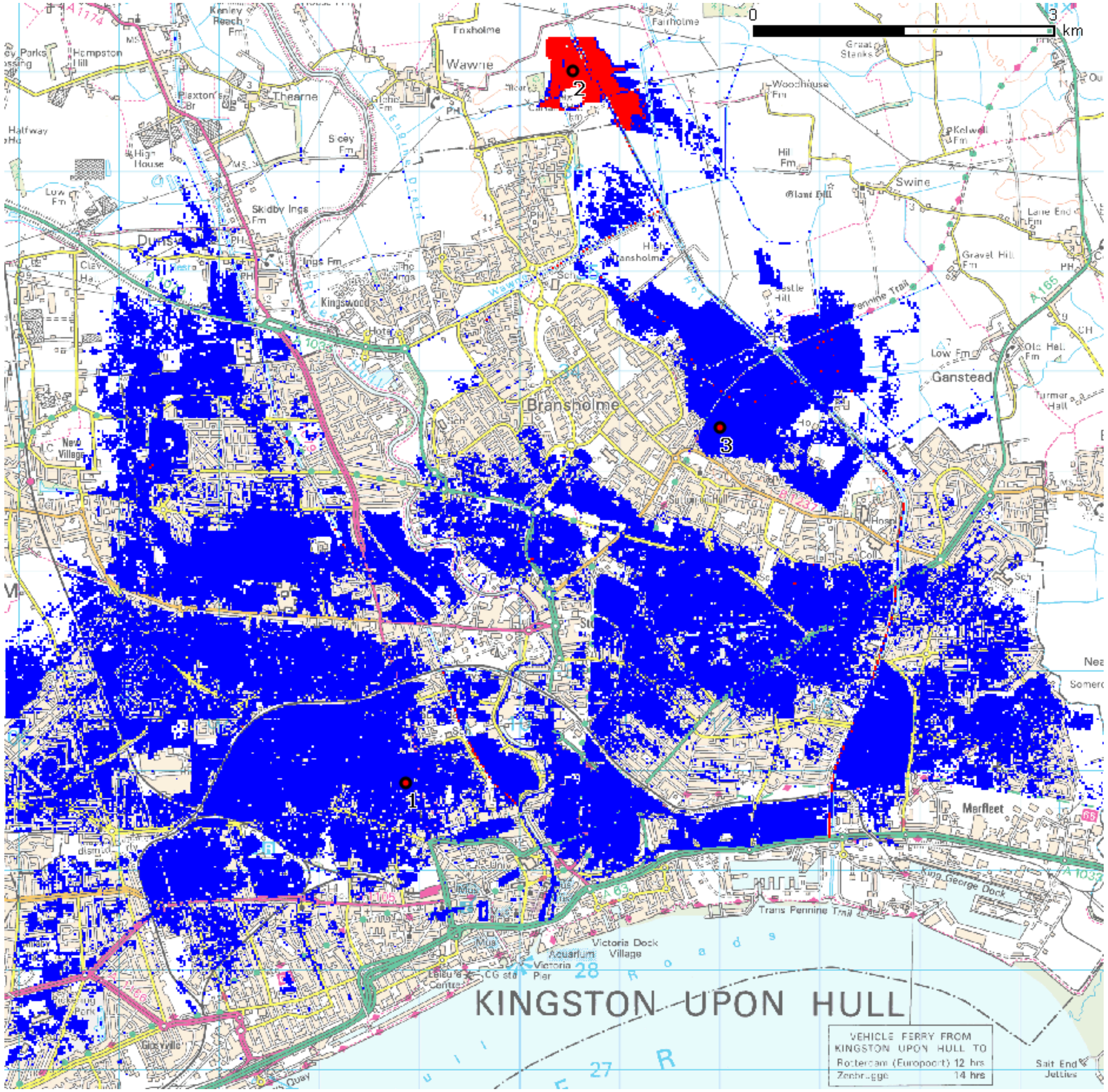


Figure 9.8c

Depth threshold=2.0m

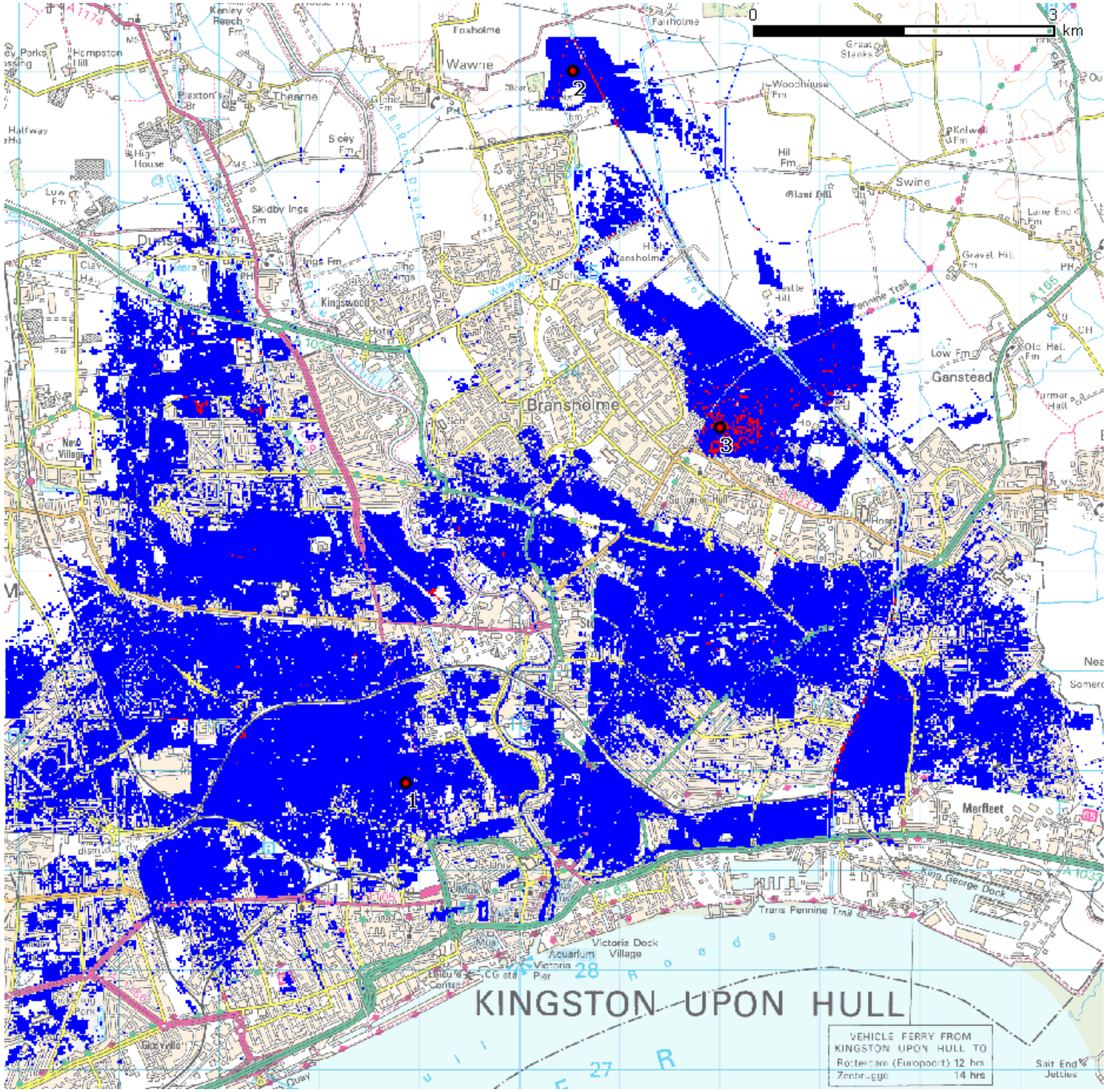


Figure 9.8 Dependency/interaction classifications for tide/surge and fluvial data for Hull, for depth thresholds 0.3, 1.0 and 2.0 m, and dependency parameter χ=0.125.

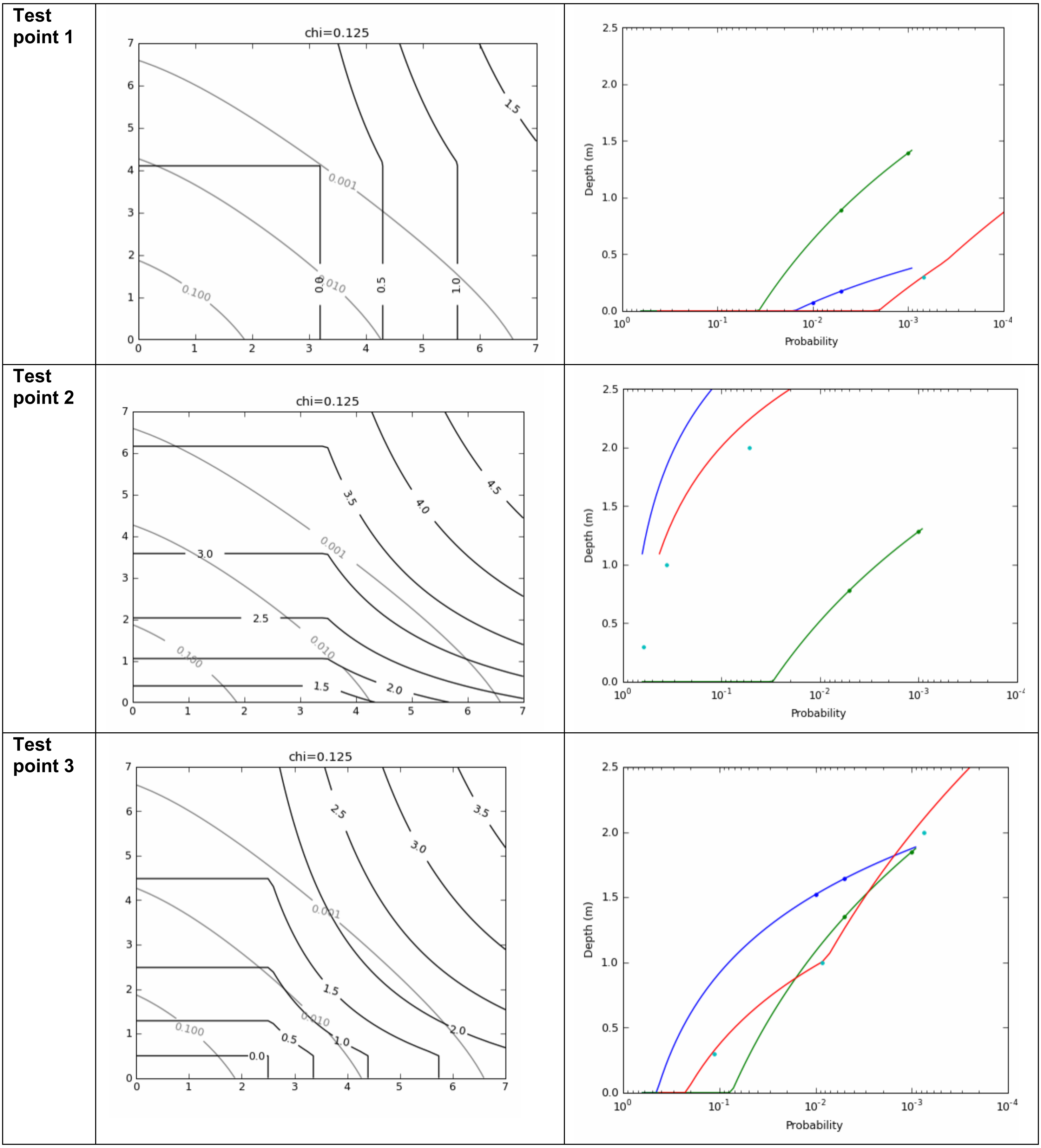


Figure 9.9 Verification of interaction/dependency results for three test points.

Depth threshold=2.0m

Blue	Weak interaction/dependency
Red	Potential interaction/dependency
Green	Likely interaction/dependency

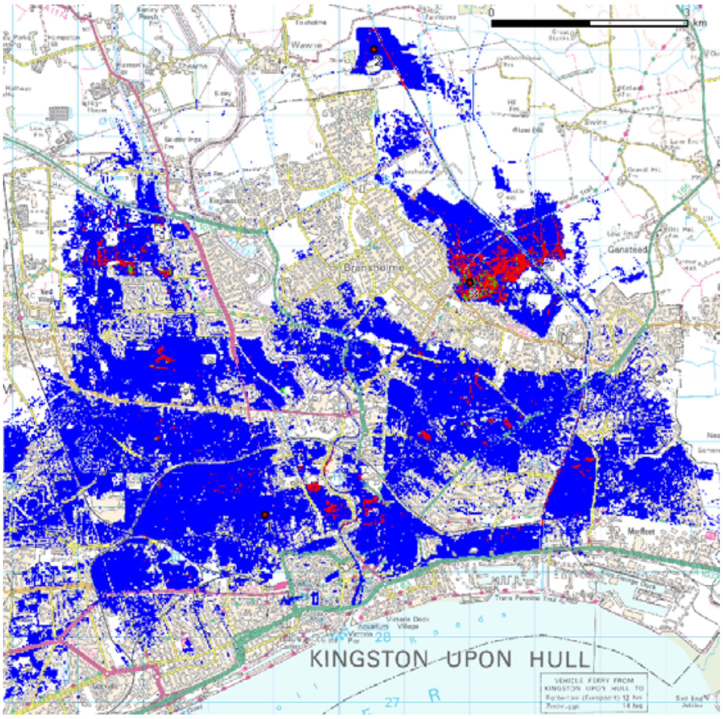


Figure 9.10 Dependency/interaction classifications for tide/surge and fluvial data for Hull, for depth threshold 2.0 m, and dependency parameter χ=0.25.

9.3.5 Interactions and dependency: tide/surge and surface water

Dependency between tide/surge and rainfall are described in FD2308 through a bivariate normal copula with dependency parameter ρ, shown in Figure 9.11. Again for Hull, the dependency is low, so an artificial value of 0.37 ('modest dependency', more appropriate to the west coast) is used here instead.

Figure 9.12 shows that there are only isolated areas of potential or likely interaction between tide/surge and surface water flooding. This is mainly due to the small depths produced by surface water flooding across Hull, when compared to coastal flooding.

Where surface water does produce significant flooding, the topographic control on depths mean that the depth–probability curves are very flat. A topographic depression will tend to fill to a level at which water spills out, and an increase in rainfall beyond that point will not tend to increase flood depths. This is the dominant process in a flat area like Hull, where flow path flooding is not widespread. Figure 9.13 shows that these flat depth–probability curves mean that interaction and dependency will be important for large depth thresholds, where flooding is more likely to arise from a combination of sources rather than each one individually. Again, as for the tide/surge–fluvial interactions, this agrees with our understanding of dependency described in Chapter 8.

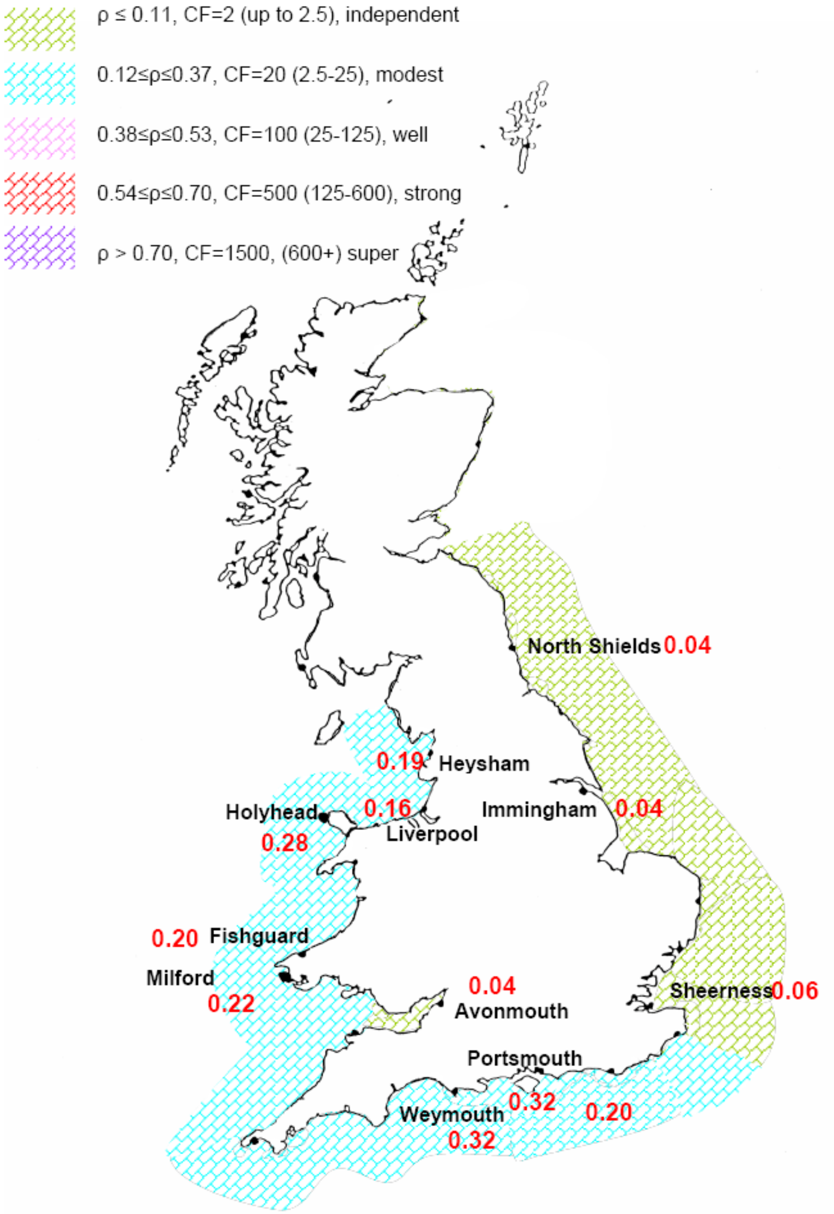


Figure 9.11 Values of bivariate normal dependency parameter for tide/surge and rainfall, taken from FD2308.

Figure 9.12a

Depth threshold=0.3m

(blue)	Weak interaction/dependency
(red)	Potential interaction/dependency
(green)	Likely interaction/dependency

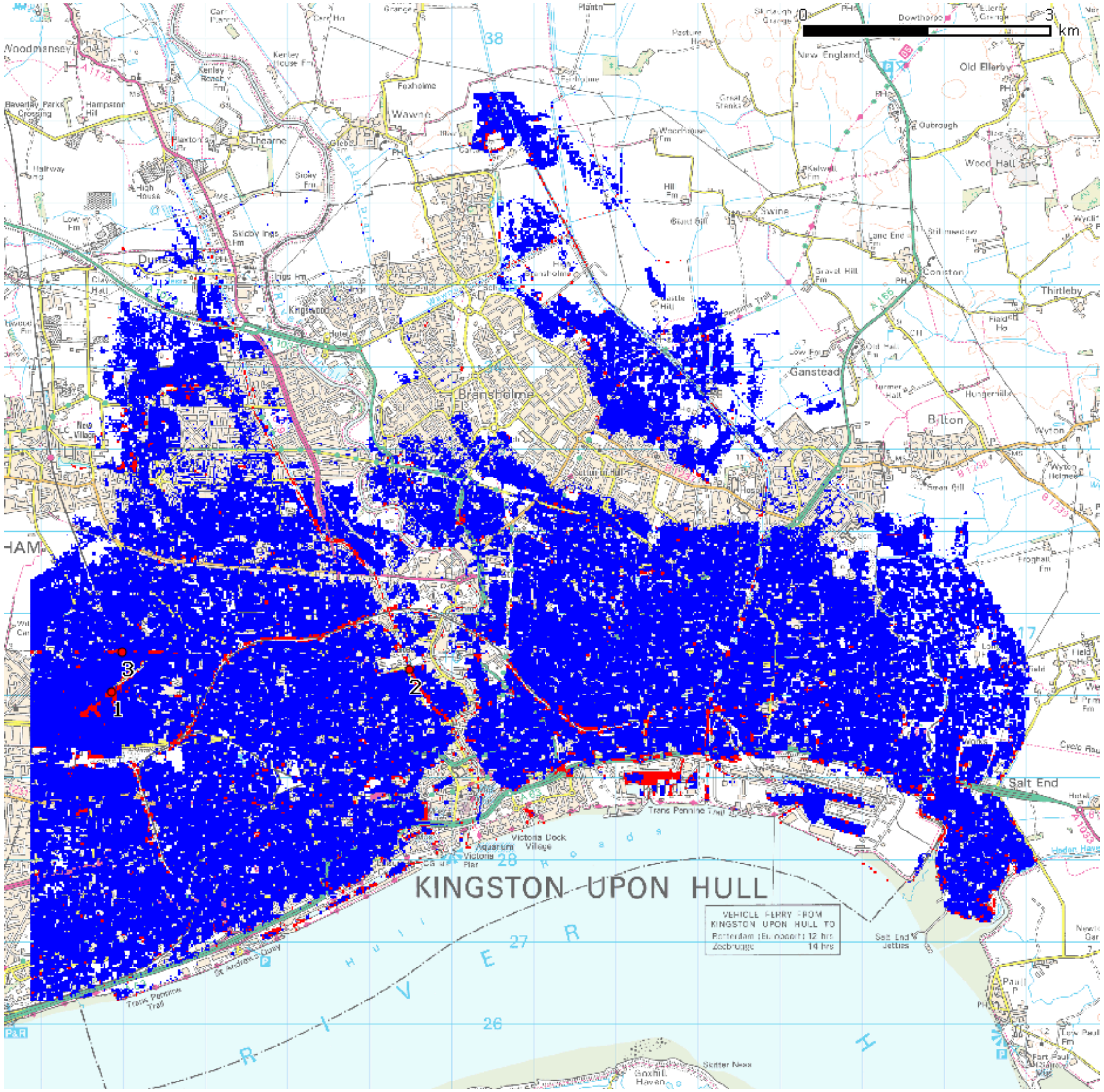


Figure 9.12b

Depth threshold=1.0m

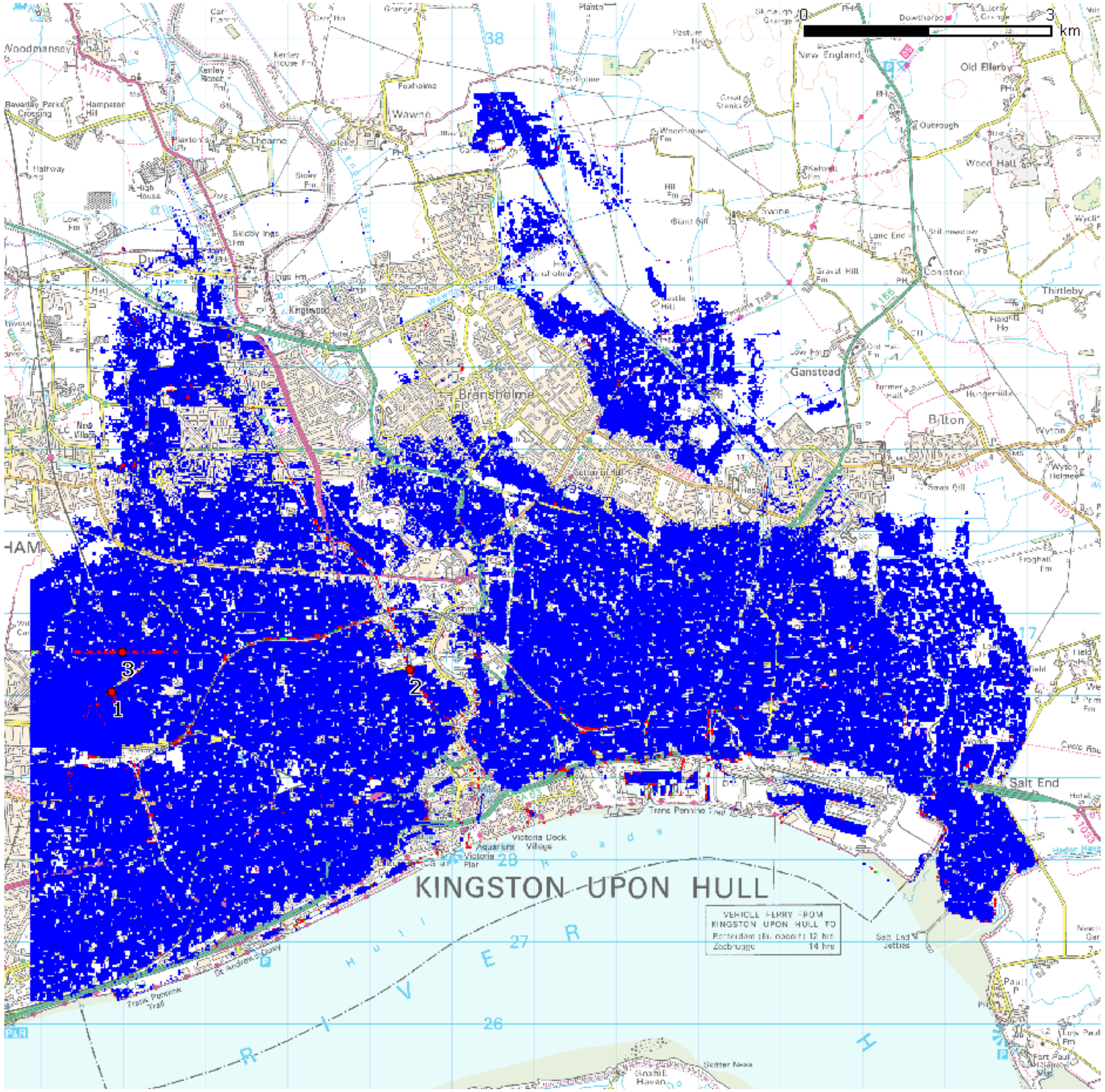


Figure 9.12c

Depth threshold=2.0m

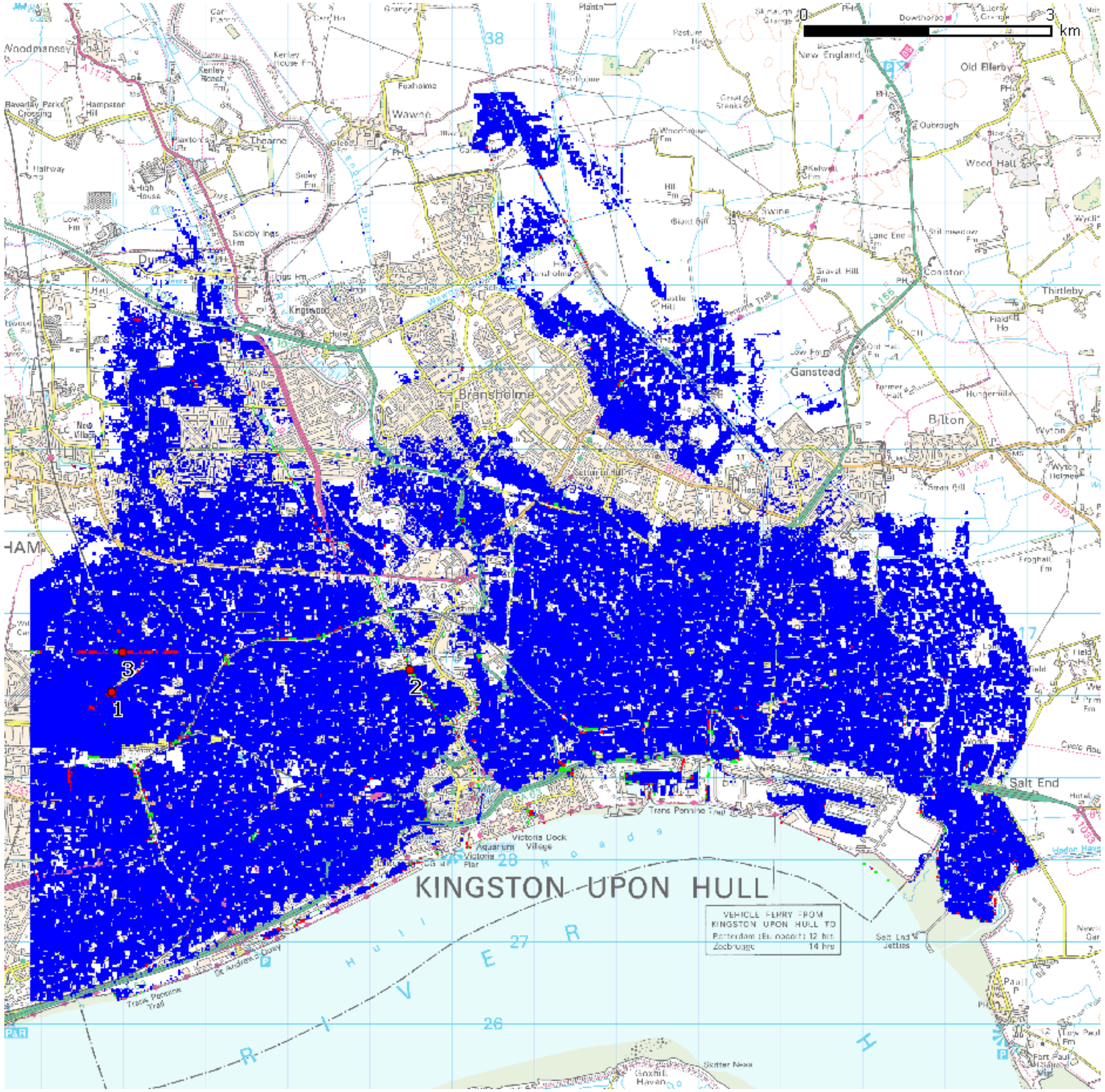


Figure 9.12 Dependency/interaction classifications for tide/surge and surface water data for Hull, for depth thresholds 0.3, 1.0 and 2.0m, and dependency parameter ρ=0.37.

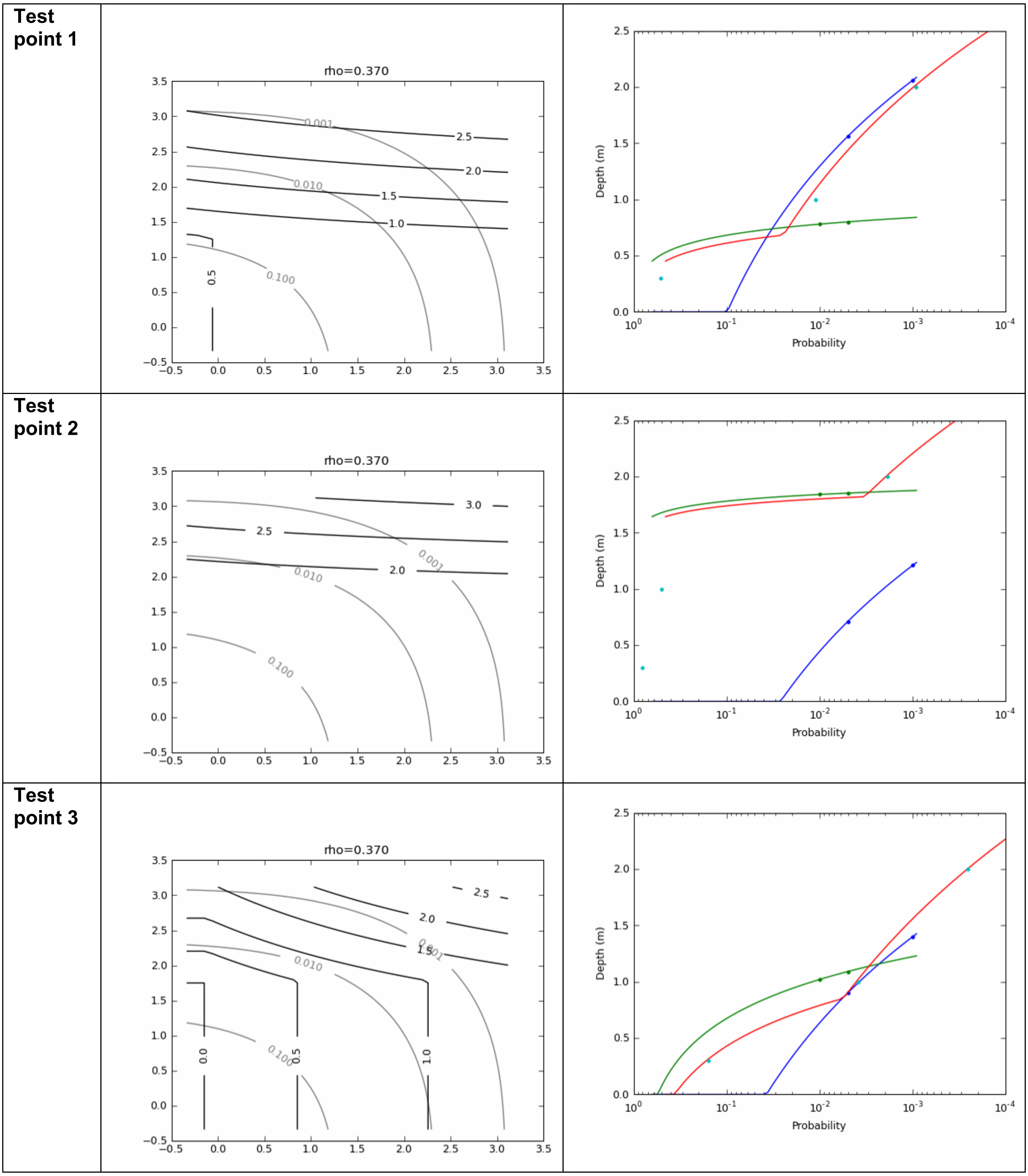


Figure 9.13 Verification of interaction/dependency results for three test points.

10 Specification for prototype software

This section details the non-functional requirements, functional requirements, technical specifications and program specifications for the software.

10.1 Non-functional requirements

Among the most important non-functional requirements are those that arise from the FD2121 project (R&D Software Development Projects – Guidance for Research Contractors, FD2121/TR2, 2007) and Environment Agency CIS needs. The application will be designed from the beginning to be as compliant with the CIS standards as possible to facilitate future use of the tool.

The best course of action is for there to be early discussion with Environment Agency CIS members to reduce risks and ensure all parties agree with the development plan.

Furthermore, we have completed the recommended tables taken from the FD2121 guidance document where possible to give a picture of where the program as currently envisaged sits in terms of compliance (see Appendix H). This will be expanded and enhanced as more details are finalised during Phase 2.

We anticipate some compromises may be required to maintain accessibility to the application by non-Environment Agency users.

Table 10.1 lists the non-functional requirements currently identified for the prototype tool.

Table 10.1 Non-functional requirements identified for the prototype tool.

ID	Priority	Description	Source
NF1	Essential	Efficiency – the prototype tool should address potential run times and storage space issues.	Consultees
NF2	Essential	Usability – the prototype tool should be acceptable to its end users as software having been tested and trialled for compatibility. Must meet a set of well-defined criteria for acceptance and be tested against these criteria in Phase 2.	Consultees
NF3	Essential	Restrict use of proprietary software formats.	CIS Standards/ Consultees
NF4	Essential	The IPR of the tool should remain with the Environment Agency. The Environment Agency will own any code produced.	Environment Agency
NF5	Essential	Maintainability – code written should be well structured and commented allowing for easier bug fixing and future development by the Environment Agency or third party (including potential future release as 'open source' software).	Halcrow/ CIS Standards
NF6	Essential	Interoperability – compatibility with other software formats.	Consultees

ID	Priority	Description	Source
NF7	Essential	Flexibility – write code so that it is easy to include future datasets, data formats and to add new capabilities to the software.	Consultees
NF8	Essential	The tool should not have onerous requirements for installation, such as reliance on expensive third party software.	Halcrow
NF9	Essential	The prototype should run on Environment Agency standard desktop machines.	Halcrow
NF10	Essential	The output produced by the tool should simplify (not complicate) the assessment of risk from all sources.	Consultees
NF11	Essential	Robustness and reliability – the application should be robust (e.g. able to handle invalid user input gracefully) and reliable (program should rarely fail).	Halcrow
NF12	Essential	Testability – it should be possible to verify the results obtained from the tool separately.	Halcrow
NF13	Essential	Usability – where processing occurs, the user should be informed through a mechanism such as an 'hourglass' or 'Please wait' message – where such processing is a matter of seconds. For longer processing tasks a progress bar should be shown to the user informing them of the time remaining for completion of the task. Shorter tasks should be undertaken 'on the fly' to allow maximum flexibility to the user.	Halcrow
NF14	Essential	Provide installer/uninstaller program – should be able to run silently to allow remote installation. Provide a minimal and typical installation where minimal contains only 'essential' components and 'typical' includes additional support material such as demo data.	Halcrow
NF15	Essential	Software to be compliant with relevant clauses of CIS Technical Standards where possible/practical.	Halcrow
NF16	Essential	Facilitate easy access to the software for appropriate Environment Agency staff.	Halcrow
NF17	Essential	Maintain scale-independence of software – data volume primarily limited by hardware rather than software. Smaller areas will support more detail.	Halcrow
NF18	Essential	To operate as a single user desktop tool (no requirement for simultaneous multiple user access to data sources).	Halcrow
NF19	Essential	Make as GIS-system independent as is practical within project constraints. It would be very useful if the prototype tool could be used without the need for a third party GIS system.	Halcrow

10.2 Functional requirements

Table 10.2 lists the functional requirements ('what the program will do') identified for the tool, based on suggestions from consultees and Halcrow.

Table 10.2 Functional requirements identified for the prototype tool.

ID	Priority	Description	Source
F1	Essential	Load or preprocess a range of spatial data formats and data types, including: • Import flood depth grid of specified probability. • Import flood probability grids of specified depth. • Import flood extent polygon (convert to depth grid). Robustly handle errors in input data (e.g. depth increasing with probability, null data values). List of input formats to be determined from analysis of datasets. Output in standard formats that can be used in other tools.	Consultees
F2	Essential	Allow display characteristics to be saved/loaded/distributed as files. Standard display formats can then be used easily.	Consultees
F3	Essential	Input files at different resolutions and extents transparently (as in GIS software).	Consultees
F4	Essential	Make input/output and processing generic so that different hazards can be processed in the same way.	Consultees
F5	Essential	Display varying levels of information for different users, either directly or through output to other systems (e.g. web server). This will affect both scale and content.	Consultees
F6	Essential	Load or pre-process different types of data, e.g. deterministic, probabilistic and qualitative. This refers to the data content, rather than its format.	Consultees
F7	Essential	New formats should be convertible to current formats as a temporary 'work around' as new formats are produced. Allow newer data to replace older data (e.g. through version number or date stamps), using similar method to F12.	Consultees
F8	Essential	Process data sources in the future in the same way as current datasets.	Consultees
F9	Essential	Tool(s) to include features to use current understanding of and information on dependency (e.g. FD2308 measures). This will involve input of dependency measures (uniform and spatially varying) and modified calculation of probabilities.	Consultees
F10	Essential	Capability of outputting a simple view of flooding from all sources.	Consultees
F11	Essential	Allow input of uncertainty information: uniform, spatially varying, related to depth or probability, qualitative. Input uncertainty either integrated with probabilistic calculation or piped through tool to output stage.	Consultees
F12	Essential	Include data with local extent, sampled to appropriate scale. Prioritise use of this data in the calculations where it overlaps with national/catchment scale data. For current tool, effect some limitations to ensure it is applied at appropriate scale.	Consultees
F13	Essential	Provide simple GIS viewing functionality of imported datasets (including panning and zooming).	Halcrow
F14	Essential	Provide simple GIS thematic mapping functionality.	Halcrow
F15	Could	Enable batching where applicable so unattended runs are	Halcrow

ID	Priority	Description	Source
	have	possible.	
F16	Essential	Assign uncertainty to imported flood depth grid.	Halcrow
F17	Essential	Works on standard Environment Agency desktop although possibly with some functionality limitations (dependent on availability of GIS software).	Halcrow
F18	Essential	Assign uncertainty to imported flood probability grid.	Halcrow
F19	Essential	Assign uncertainty to imported flood extent polygon.	Halcrow
F20	Essential	Calculate flood probability grid for different user-specified depth thresholds from multiple flood depth grids.	Halcrow
F21	Essential	Calculate flood depth grid for different user-specified probabilities from multiple flood probability grids.	Halcrow
F22	Essential	Import joint probability relationships between sources.	Halcrow
F23	Essential	Calculate joint probability grid from pairs of individual source flood probability grids for user-specified depth (include nesting for 3+ sources).	Halcrow
F24	Essential	Display background mapping (to include OS Tiles, photogrammetry, DTM, historical flood extents, flood risk management areas – formats to be supported to be confirmed).	Halcrow
F25	Essential	Export joint (co-) probability flood extent maps for user-specified probabilities and depths.	Halcrow
F26	Essential	Metadata (to appropriate standard) to be generated.	Halcrow
F27	Essential	Display source contributions (and associated uncertainty levels) to overall probability for user-selected cells.	Halcrow
F28	Could have	Provide unified data repository for single source flood depth, flood probability grids and flood extent polygons and associated metadata (including information on uncertainties, dependencies).	Halcrow
F29	Could have	Structured case management: climate, management options, receptors.	Halcrow
F30	Could have	Default sets of parameter values (e.g. depth thresholds, probability values) that can be overwritten by the user for specific applications. These defaults can be stored in a XML parameters file to allow them to be changed easily, e.g. to correspond to values used in NaFRA.	Halcrow

10.3 Technical specification

10.3.1 Introduction

This section seeks to define a robust, flexible and maintainable architecture within which to develop the tool. The section recommends a number of software development best practices that should be considered for adoption within the application development (particularly important if the tool is to be developed further after completion of Phase 2).

The recommendations in this section also take into account the FD2121 report recommendations (see FD2121 section, Appendix H).

10.3.2 Platforms – operating system

The vast majority of potential users of the tool are likely to be running a Microsoft-based operating system. As such, the tool will be developed to work on the following operating system platforms:

- Windows 2000
- Windows XP Home
- Windows XP Professional
- Windows Vista

Windows 2000 is specifically included as it is the Environment Agency standard desktop operating system and so must be supported by the application. Note that other legacy operating systems such as Windows 95/98/ME, Windows NT 3.x and Windows NT4 will not be supported.

It is an open issue as to whether to develop and test on Windows 7. A decision will be made during Phase 2 in consultation with the Environment Agency, and with regard to the uptake of Windows 7 and likely future Windows releases. Ideally, the tool should be as operating system independent as possible.

10.3.3 Coding

Suitable coding standards should be adopted by the developers in order to provide a consistent style and easily comprehended code. This will facilitate any future development.

Code should be well commented and it is suggested that a tool such as JavaDoc is used to generate documentation. JavaDoc is a free utility that can be used to generate developer documentation from comments in Java code and is widely used in Java development. There are equivalents available for other languages.

The developers should make use of the language features to develop code in a modular fashion to produce code that can be easily extended and reused where practical.

Coding will need to take account of the potentially large computational burden imposed by processing large grid datasets. The calculation of interaction and dependency parameters is especially intensive, and development must take account of the likely size of datasets that users will expect to process. These datasets could be large for the analysis of broad-scale risk, with large areas of potential interaction for low-lying areas (e.g. the Fens).

10.3.4 File formats

A requirement for three file types (other than the input/output GIS formats) has been identified:

- Project file format, describing input datasets, thresholds, dependency information etc.
- Program settings file format, describing default values, paths etc.
- Display characteristics. Different display characteristics may need to be referred to in the project file format.

All three of these formats will be defined in XML (eXtensible markup language), which is an open standard, text-based file format. This is a common-sense approach and also satisfies the Environment Agency CIS requirement to avoid using proprietary data formats.

10.3.5 Data storage

There will be no centralised data storage (e.g. Oracle database) component to the prototype tool. This would make the application much more complex to install. It would also present a barrier to take-up of the tool and, importantly, is deemed technically unnecessary. The user will simply provide input files (grids etc.), most likely from their local machine (or network), and the application will process these files and output files to local or network disks. While not one of the primary objectives of this project, the development and piloting of the tool may produce recommendations for file formats to be used in exchanging depth, probability, uncertainty etc.

10.3.6 System architecture

In order to be fully compliant with Environment Agency CIS standards, the prototype tool should satisfy the following major conditions (not an exhaustive list, see the FD2121 guidance document in Appendix H for fuller details):

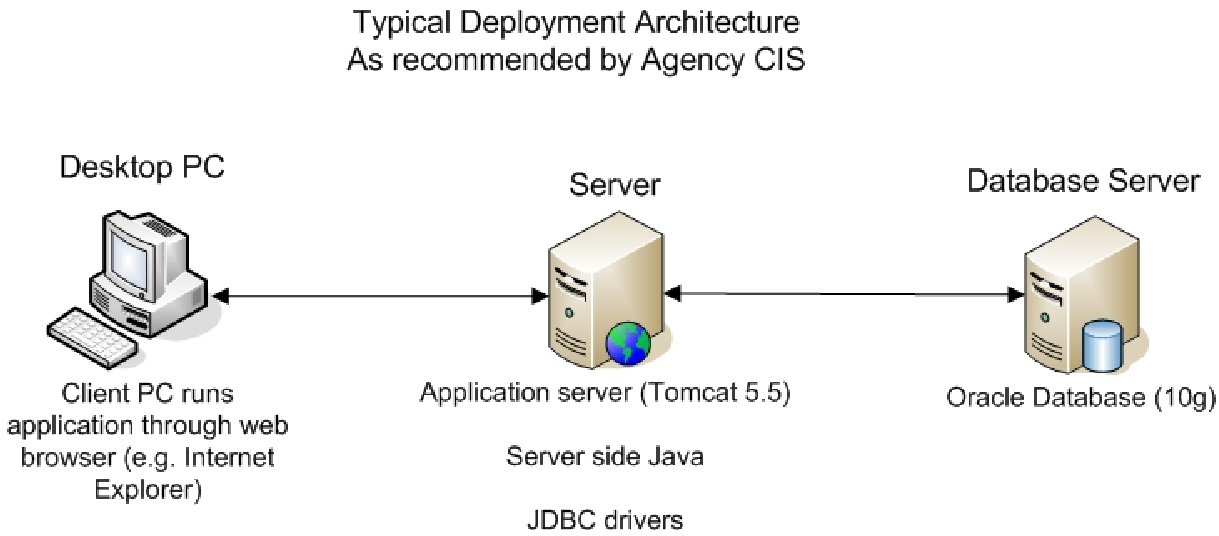
- Development language – developed in Java.
- Deployment architecture – n-tier structure utilising a browser-based thin client.

Language choice should be based on ability to meet the functional and non-functional requirements. Because of the Environment Agency CIS's very strong preference for Java development this must be considered seriously. It is a popular, well-supported language, is object-oriented and would seem well suited to the task (if chosen it is expected that development will be undertaken using version 1.6 of the Java Development Kit (JDK 1.6) – NetBeans is suggested as the integrated development environment to use for development). Table 10.3 shows some of the main things to consider when deciding which language to use for the prototype. We would suggest that the biggest issue identified is that of suitability for GIS development (in particular user interface tasks) and the developer would need to ensure that the language chosen can satisfy the requirements of the tool.

Table 10.3 Considerations on whether to use Java or C# for the prototype tool.

Consideration	Java	C#
Acceptable to Environment Agency CIS	Yes, Java is the preferred development language of the Environment Agency.	No, not without very good justification.
Object oriented	Yes	Yes
Well supported	Yes, Java is a very popular language and there is a large user community.	Yes, C# is also a very popular language with a large user community.
Performance	Java computational performance is good – similar to C#.	C# computational performance is also good – similar to Java.
Suitability for 'rich' user interface development	The Java SWING library provides many controls that can be used in user interface creation. Fewer third party controls available than for C#, but should be able to satisfy the requirements of the prototype tool – and those that are available are more likely to be open source.	C# has many controls available and is better supported by third party controls. They tend to 'look better' than their Java counterparts. These are less likely to be open source though.
Suitability for GIS development	There are GIS toolkits available for Java and some of these are free.	GIS toolkits are available for C#. They tend not to be free. However, it is generally the case that GIS development is better supported than for Java.
Deployment issues	Requires the Java Runtime Engine to be installed (this would most likely already be installed on Environment Agency machines).	Requires the .NET framework to be installed. Depending on the version used this may already be installed on Environment Agency machines.
Can run on non-Windows platforms	Yes, but not considered a requirement for the tool.	No, but may be adaptable through C#/mono implementation on Unix-type platforms.

In order to develop the tool in strict compliance with Environment Agency CIS guidelines the developer would need to develop using an n-tier architecture. Figure 10.1 illustrates this type of architecture, along with one suggestion for a possible better solution for this tool.



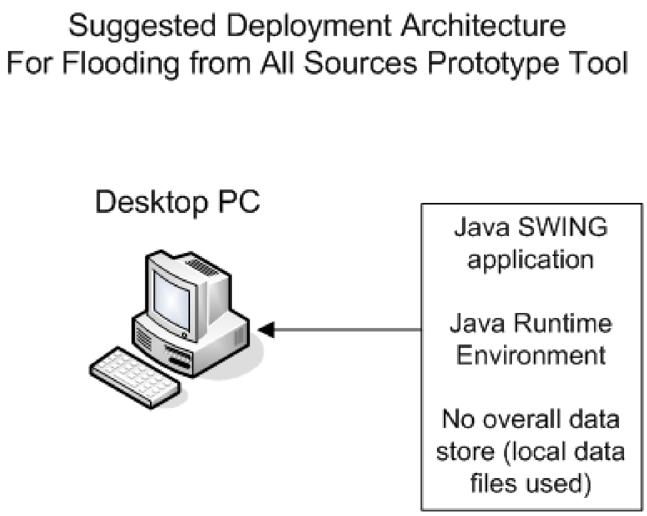


Figure 10.1 Deployment architecture.

With an n-tier architecture, components are separated into logical groups, typically spread over different computers (though not necessarily so). This would normally consist of a number of layers. The layers are:

- A 'thin client' user interface – this implements the user interface part of the application but performs very little processing. In the case of Environment Agency CIS recommendations, the user interface is presented via a web browser application.
- A 'business logic' layer, which implements the program logic and is where the main processing occurs – this would be where the calculations would take place in the case of the prototype tool.
- A 'database access layer' – this layer deals with reading and writing information to the database.

This is a good, logical approach in many cases, but in the case of the prototype tool it would seem not to be the most appropriate for the following main reasons:

- There is no database access layer to the application.
- GIS functionality in web browser interfaces is typically less responsive, giving a poorer user experience, as well as being more difficult to develop.
- Installation issues – following Environment Agency CIS guidelines would involve users needing to install a web/application server, either onto servers in their organisation or their own desktop PCs. This is a barrier to use of the program, particularly for users external to the Environment Agency.

One solution (also shown in Figure 10.1) is to develop the application to run as a standalone tool, so that it is run on a single self-contained PC, in order to overcome the above issues. It is worth noting that requirements (such as the Java Runtime Engine) of the application would be the same as for the MDSF2 user interface currently in development.

The developer of the tool should code it in such a way that the user interface is decoupled from the logic/processing code of the application. This approach would allow the tool to be moved to an n-tier structure in the future, or allow the engine to be integrated with another front-end delivering the GIS functionality.

10.4 Specification for the prototype software tool

10.4.1 Overview

A user will use the program in order to generate combined flood maps or datasets from a variety of data sources. They will do this by adding layers (e.g. flood maps and grids) within the program, setting up joint probabilities between various sources and defining depth thresholds or output probabilities. The user will then choose to process this data, which will result in combined datasets and/or flood maps being produced. The output can then be used to provide information and input to other software tools.

10.4.2 User scenarios

Some scenarios for use of the program and outputs of the program are described below:

Scenario 1 – An Environment Agency employee wishes to combine information on probability of flooding from single sources to provide an assessment of likelihood of flooding from all sources for public information.

Scenario 2 – An Environment Agency employee wishes to identify high risk areas with significant multiple source interaction and/or high levels of uncertainty to assist project commissions.

Scenario 3 – A consultant undertakes the above tasks as part of work for the Environment Agency (i.e. Scenario 1 or 2).

There are two additional potential user categories:

- Local authority staff undertaking Scenario 1 or 2 – this is not an essential requirement of the *prototype* tool.

- Other stakeholders (e.g. public) – these are excluded from the requirements.

Scenarios such as these will be developed further at the start of Phase 2 to improve the functional design of the tool.

10.4.3 Non-goals

There will be no centralised data storage aspect to the program.

The prototype will not need to interface/integrate directly with tools such as ESRI's ArcGIS, but the tool will support loading and exporting to common proprietary file types.

10.4.4 Prototype

The specification applies to the development of a prototype.

10.4.5 Details

A number of software features/modules have been identified from the user functional and non-functional requirements. A mock-up of how these features might look is given in Chapter 9, and Table 10.4 lists the features.

Table 10.4 Features identified from user requirements.

Layer management
The user can click a button or icon to display the layers (e.g. flood extents, grids) currently loaded into the application, or these will be on view in a permanent panel. From here the user can: **Add a layer group** – This will bring up the 'Add layers' screen which will allow the user to set up a layer group. Here the user can set the following values: **Name**: Textual identifier for the group, e.g. 'Coastal NaFRA Output'. **Depth/Probability**: The user can choose between 'Depth' or 'Probability'. **Uncertainty**: The user chooses 'Low', 'Medium', 'High' or 'Other' (other allows them to enter a numerical value manually). **Layers**: The layers associated with this group are shown (initially empty) in a grid. The grid displays the layer name and the associated depth or probability for the layer. The user can add layers via a file open dialog or remove layers from the group. The user may also need to set a depth or probability uncertainty value or values (e.g. upper and lower probability bounds) for each layer. **Type:** This will be used to classify the layer sets, allowing the tool to decide which results can be overwritten (e.g. local results over national results). Several types may need to be selected, e.g. for a local estuarine model replacing fluvial and coastal results. **Numerical/thematic:** Allows the user to define input as being a hazard classification rather than depth/probability. This will then be overlayed on

the output rather than used in the combined probability calculation. **Edit a layer group** – This will bring up the 'Edit layers' screen for a currently selected layer group. This screen will be the same as the 'Add layers' screen and will work in the same way. **Remove a layer group** – This will remove a currently selected layer group (after user confirmation) from the project. If this layer group has been included in one or more 'joint probabilities' relationships (see *Joint probabilities)* these will need to be removed from the project also (the user will be warned if this is the case).
Project management
A 'project' consists of the layer groups, the joint probability relationships set up between them, and the threshold depths and probabilities used to generate the outputs. The user will be able to save and load projects to the interface, and the project file should allow results to be reproduced with no extra information from the user.
Scenario management
The user will be able to group numbers of layer sets, into future climate scenarios and management options. The user can then quickly switch between viewing different scenarios.
Joint probabilities
The user will access an interface to define joint probability relationships between pairs of sources, either using chi or bivariate normal copulas. The user will have the option to define different potential/likely interaction thresholds. Default values of chi/rho will be provided appropriate to the dataset types being combined. The dependency parameter can be entered as a single value or as a spatially varying value.
Settings management
The user will be able to control: **Colour schemes**, e.g. greyscale, colour ramps, user defined, from files **Layer transparency**, e.g. to make background mapping visible **Data scaling**, e.g. linear, reduced variate, histogram equalised **Spatial data displayed**, e.g. combine exceedance probability or 'best estimate' **Options for representing uncertainty**, e.g. transparency, saturation, fill style **Select scenarios** to display, e.g. climate change epoch, management scenario
Display maps/layers
The user will be able to manipulate the tool's view of the output data: **Pan, zoom** in an efficient way **Add** background mapping, aerial photography etc. **Add GIS layers**, e.g. receptors, surveyed flood extents **Switch** between input and output layer views **View scale** through scalebar or gridlines
Probability calculation
The user will actively start the calculation once the input data have been defined, and will be prompted to recalculate if input data are updated. A progress bar and

cancel/pause option will be available.
Query outputs
The user will be able to query a point on the map to obtain information on: **Input depth/probability grids** including uncertainty **Combined output** depth/probability and uncertainty. For a point, this will be recalculated using a larger number of depths/probabilities than for the gridded data **Contribution** from each source to total risk **Scenario to be displayed**, allowing the user to quickly switch between them and compare results
Export results
Export output grids Export point queried data Export screen shot

10.5 Reuse of MDSF2 code

There is little scope for reuse of MDSF2 code within the development of the prototype. However, due to the probabilistic nature of the method used, the tool should be compatible with MDSF2 as a pre- or post-processing stage. For example, MDSF2 outputs representing fluvial and coastal flood probability can be combined with other sources using the tool. The probabilistic nature of MDSF2 outputs will integrate well with the tool, avoiding the large uncertainties that may occur with depth grid data.

10.6 IPR needs

All IPR of the application will be transferred to the Environment Agency.

11 Benefits realisation

11.1 Introduction

Chapter 2 introduced the ultimate outcomes that the project is intended to facilitate. However, these outcomes are some way down the line and it will be important to understand what actions are required, and plan for them, in order to achieve the full benefits of the project. As a reminder, the objective of Phase 1 of this project is to develop a method suitable for mapping flooding from all sources, and in Phase 2 the method will be embedded in prototype software. This chapter introduces the subsequent activities that will be required to realise the benefits of the Science project and also discusses linkages with other Science projects.

11.2 Alignment with high level Science outcomes

Benefits realisation management is recognised in the Environment Agency's Science department as important. A set of strategic outcomes have been identified and projects are aligned to these outcomes and assigned to 'streams' of work designed to facilitate these outcomes. This project contributes to the Science Stream 'S5 – Rural and urban catchment science', within the outcome group 'O-2.1 – Increased capability to define and model rural and urban catchment processes and systems'. It leads towards Strategic Outcome 'SO-2.0 Increased or would increase capability to model FCRM processes and systems'.

The project is also part of the Initiative Groups 'I2 – Developing applications and guidance for Strategy, Planning and Development Control functions' and 'I3 – Developing applications and guidance for Data, Modelling and Mapping functions'. These feed into outcome groups O-3.1 and O-3.2 thus leading to Strategic Outcome 'SO-3.0 Created or would create business applications (models/tools) and best practice guidance for FCRM'.

11.3 Actions to realise the benefits

While the high level Science outcomes are useful in ensuring Science projects are aligned with business needs, they are not necessarily aligned with the evolving FCRM Mapping, Modelling and Data (MMD) policy team desired outcomes. It is recommended that the MMD 'benefits roadmap' is reviewed considering the prototype tool for flood mapping. This task is best undertaken during or in parallel with Phase 2. It is also recommended that during Phase 2 a full analysis is undertaken of activities necessary to realise the outcomes introduced in Chapter 2. An outline of these activities is provided in Figure 11.1.

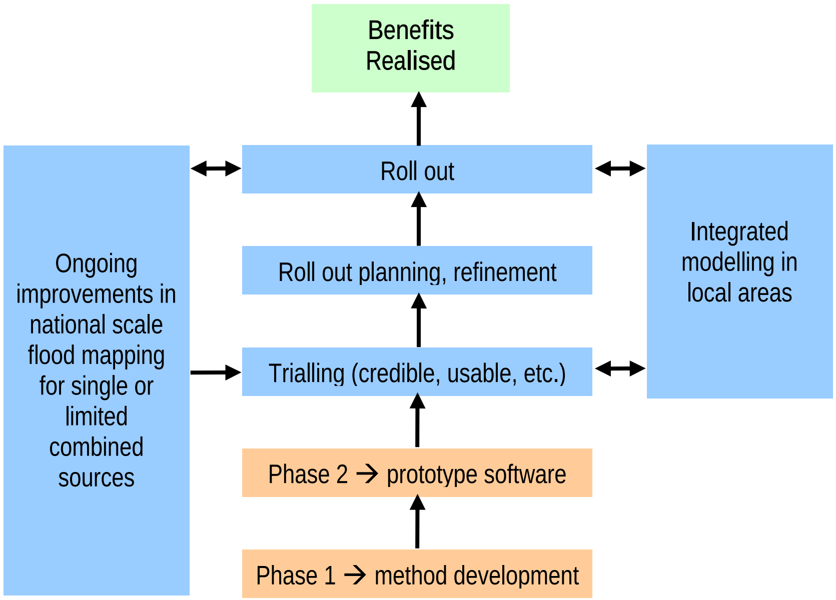


Figure 11.1 Outline road map to benefits realisation.

11.4 Links with other Science projects

Table 11.1 shows the project's links to other Science projects.

Table 11.1 Links between this project and other Science projects.

Science project	Linkage
Scoping the development and implementation of Flood and Coastal RASP Models (SC050065) (complete)	Provides description of RASP framework.
MDSF2 development (SC050051) (ongoing)	Defines the MDSF2 system.
Spatial coherence of flood risk (SC060088) (ongoing)	The main focus of the 'all sources' project is mapping spatial patterns of risk, rather than integrating risk over large areas. Hence issues of spatial coherence (e.g. how likely is widespread flooding compared to isolated incidents?) are of limited relevance. Thus there are no outputs from the spatial coherence work that are directly applicable to the 'all sources' project at this stage. Some methodological advances from the spatial coherence work may be of relevance to future versions of a 'all sources' tool: • Improved joint probability information for fluvial and tide/surge events. • The multivariate statistical regression model used to represent the dependency between multiple sites may be of use to represent joint probabilities between more than two sources, but at present there has been no work done in applying such a model to three or more sources. The 'all sources' project should therefore concentrate on the more established methods for joint probabilities such as those described in

Science project	Linkage
	FD2308.
Joint probability: dependence mapping and best practice (FD2308) (complete)	Primary source of dependence information for many sources, and modelling results are likely to be in formats implied by the use of these methods. Variable pairs with dependencies reported are: • wave height and sea level • river flow and surge • hourly rainfall and sea level • wind-sea and swell • wave height and surge • tide and surge • daily rainfall and surge. Modelling outputs using this approach are likely to be in the form of water levels for a number of joint probability situations (e.g. 100 year flow + 5 year sea level). Maximum water levels are then taken for each location in the model study area.
Broad-Scale Modelling Scoping – a vision for flood modelling and risk science (FD2118) (complete)	Provides general guidance.
Software requirements for Joint FCERM R&D programme modelling outputs and architecture specification for RASP family outputs (FD2121) (complete)	The requirements will need to be taken into account at prototyping stage (Phase 2 of this project) to ensure the architecture is compliant with Environment Agency systems.
Risk, Performance and Uncertainty in Flood and Coastal Defence – A Review (FD2302) (complete)	Provides general guidance; for example it indicates that uncertainty should be included where possible in an integrated flood map.
Refining the data quality and the methodology for mapping surface and groundwater flood risk (SC080029) (ongoing)	The resulting surface water map may be one of the sources of flood information to be combined.
Validation and Calibration of Probabilistic Flood Models (SC090008/WP1) (ongoing)	This project will develop methods for validating probabilistic flood maps (e.g. NaFRA outputs). These may be used to validate the outputs from the prototype tool, although the need to validate against all sources makes obtaining suitable validation data particularly difficult.

12 Conclusions and recommendations

The key project drivers are seen as the Floods Directive and the Pitt Review, in particular facilitating delivery of the Environment Agency's strategic overview responsibilities including flood mapping at a national scale. The project has developed a method (to be implemented in Phase 2 as a prototype tool) for mapping flooding from all sources. The software deliverable from Phase 2 will not be suitable for national roll-out for operational use because a further phase of trialling will be necessary. The project is delivering a software tool – not new flood maps or joint probability information.

The proposed method is a hybrid between a map combination approach, in which pre-existing flood map data are combined, and more detailed modelling where interactions and dependency between sources may be important. The method can identify areas where more detailed local modelling is required to represent significant interactions between sources. The method is therefore expected to be reliant on the availability of suitable pre-calculated flood maps from individual or previously combined sources, adding detail from local studies where necessary.

The sources of flooding that should be included within the scope of the project area are listed in Table 12.1.

Table 12.1 Sources of flooding to include within the scope of the project area.

Include and considered feasible	Include but significant issues exist	Exclude (but may still be able to be used and included in the future)
Fluvial (main river) [national]	Reservoir breach [national]	Sewer (blockage/failure)
Coastal [national]	Fluvial (small watercourses) [local]	Tsunami
Estuarine [national]	Groundwater [local]	[*]Canal overtopping
Surface water [local]	Canal breach[*(subject to review)]	Water supply failures
Breaching (fluvial/coastal)	Sewer (capacity exceedance)	
	FRM point infrastructure failure (pumps, barriers...)	

Notes:

The *national* and *local* descriptions of sources are based on the draft Flood and Water Management Bill which is subject to change. National scale mapping is taken as being the Environment Agency's responsibility, local is taken as the responsibility of local authorities. The Environment Agency will still take a strategic overview role for some sources mapped by local authorities (e.g. surface water). The significant issue with including reservoir inundation is defining the probability of the event – the Environment Agency is undertaking research on this issue and the proposed method will be able to integrate the reservoir inundation mapping once probabilities are assigned.

[*]Canal breaching and overtopping are likely to be classed as a 'Fluvial' source.

In Table 12.1 the sources of flooding in the first column must be able to be used with the method developed during the project. The method should also be developed to use the sources in the second column; however, it is acknowledged that there are issues with the data and/or physical attributes of these flooding phenomena which may make their inclusion difficult (e.g. their inclusion may degrade the utility of the results). The sources listed in the third column will be excluded from consideration during the project (but with availability of suitable data it may subsequently be possible to use these sources in the developed method).

The project must deliver a method that is able to generate a probabilistic national-scale flood map presenting a measure of the probability of flooding throughout England and Wales caused by the selected sources.

In achieving this goal the method will need to account for joint probabilities/dependencies, combine a range of information which may be at different scales and with different attributes, and deal with 'vague or soft' data (e.g. vulnerability maps) which are unlikely to have specific probabilities defined.

Desirable attributes of the outputs of the method are:

- Communicates uncertainty in the estimate of flooding (depth and/or probability).
- Communicates which sources have been included/excluded at specific locations.
- Communicates which sources contribute most to the flood hazard at specific locations.

Other attributes which may be considered useful include:

- Ability to include historical data.
- Ability to include future scenarios (climate change).
- Ability to use new DTM data to improve estimates from pre-existing mapping (e.g. resample from water surface elevation data).

12.1 Recommendations for Phase 2

It is recommended that Phase 2 of the project is undertaken to produce and pilot test a prototype tool for mapping flooding from all sources using the method presented in this report. A separate 'work plan' for Phase 2 has been developed and includes details of the following four recommended tasks:

- Review the functional design with a small group of expected end users, revise the design if necessary, develop the software and user guide and undertake software functionality testing.
- Undertake formal pilot testing of the prototype tool and assess the tool in terms of credibility of results, usefulness of results and usability of the tool.
- Update the prototype tool in response to the outcomes of the pilot testing.
- Final reporting and implementation planning.

12.2 Recommendations for future Environment Agency Science

The research undertaken for this project has identified the issues and questions listed in Table 12.2, which may help to guide the Environment Agency's future Science projects. The questions have been raised during both the consultation phase and development phase of the current project. (Note that some of the following may be addressed by ongoing or proposed projects.)

Table 12.2 Questions identified and links to Science projects.

Spatial Coherence Links to: SC060088 Spatial Coherence; FD2105 Improved methods for national spatial–temporal rainfall and evaporation modelling; DTI-SAM; SC070059 Developing the Next Generation of Surface Water Flood Risk Assessment	Treatment of probability at a point in space (e.g. flood probability for a single location) neglects the joint probabilities of nearby points being flooded also. The nature of these spatial joint probabilities has consequences for understanding risk at national (e.g. annual damage variances in NaFRA) down to local (e.g. emergency planning) scales. Phase 1 of this project developed and tested a method to investigate the likelihood and consequence of different places being flooded around the same time from rivers or the sea. It allows us to investigate the risk of widespread flooding, such as winter 2000/01. Spatial dependencies differ between sources (coastal sources highly dependent, surface water highly spatially variable). The main challenges are to understand these dependencies, and to communicate them in a way that can improve flood risk management. SC060088 has investigated the coherence between river flows at gauging stations; how this can be applied at other points/scales and sources is poorly understood.
Joint Probabilities Links to FD2308 Use of Joint Probability Methods in Flood Management; SC060064 Development and Dissemination of Information on Coastal and Estuary Extremes	FD2308 developed and described a number of methods for understanding joint probabilities for different flood sources (tide/surge, rainfall, fluvial flow etc.), focusing on the needs of local-scale modelling. Generalising these results to national scale is not straightforward given the spatial variability of the statistical processes. Other processes such as local rainfall and its dependency on catchment-scale rainfall are poorly understood, but are important for, for example, surface water modelling. It is not clear whether dependency between more than two sources is a significant source of risk.
Reservoir Risk Modelling Links to: RIM	The RIM study is generating risk maps conditional on reservoir failure, but failure probability is not well understood. Putting the risk from reservoirs in context with other risks requires a knowledge of failure probability.
Validation Links to: SC09008/WP1 Improving probabilistic Flood Risk Modelling capabilities	How do we assess the accuracy/reliability/ usefulness of combined maps? How should we draw sensible conclusions about risk after a flood event? Does this tell us anything useful?
Future Risk	Changes in some sources of flooding in response to a

Links to FD2113 Effect of climate change on catchment rainfall	changing climate are reasonably well understood (e.g. coastal), but for others are much more uncertain (e.g. convective rainfall). A holistic view of flood risk in the future requires understanding of how all risks are likely to change in the future.
Probability and Uncertainty in Risk	The different sources of uncertainty (natural variability and knowledge uncertainty) impact on decision making in different ways. For example, we can spend money on either reducing knowledge uncertainty or allowing for it in decision making (e.g. freeboard); we can do less about natural variability. Dealing with these two types of uncertainty in a consistent framework is a major challenge.
Groundwater flood risk	There are currently few methods for quantifying groundwater flood risk, and none available for operational use. Recent work by the BGS has developed a map of susceptibility based on thickness of the unsaturated zone (further detail given in Appendix F), but turning this into a map of flood probability is a further major challenge.

Abbreviations

1D, 2D	one Dimensional, two Dimensional
AEP	Annual Exceedance Probability
BGS	British Geological Survey
BW	British Waterways
CEH	Centre for Ecology and Hydrology
CEP	Combined Exceedance Probability
CFMP	Catchment Flood Management Plan
CIS	Corporate Information Systems
DEM	Digital Elevation Model
DTM	Digital Terrain Model
ERA	Extreme Rainfall Alert
FMAS	Flood Mapping from All Sources
FRA	Flood Risk Assessment
FREE	Flood Risk from Extreme Events
FRIS	Flood Reconnaissance and Information System
FRM	Flood Risk Management
FRMRC	Flood Risk Management Research Consortium
GEM	Groundwater Emergence Map
GIS	Geographic Information System
GWL	GroundWater Level
IDB	Internal Drainage Board
IPR	Intellectual Property Rights
JBA	Jeremy Benn Associates (Consulting Engineers)
LRF	Local Resilience Forum
MDSF	Modelling Decision Support Framework
MSfW	Making Space for Water
MWH	Montgomery Watson Harza (Consulting Engineers)
NaFRA	National Flood Risk Assessments
NFCDD	National Flood and Coastal Defence Database
PAR	Project Appraisal Report
RASP	Risk Assessment for System Planning
RIM	River Inundation Mapping
SEPA	Scottish Environment Protection Agency
SFRA	Strategic Flood Risk Assessment
SMP	Shoreline Management Plan
SQL	Structured Query Language
SWMP	Surface Water Management Plan
TE2100	Thames Estuary 2100
XML	eXtensible Markup Language

References

Environment Agency, 2009

Hawkes, P.J., 2005. *Use of Joint Probability Methods in Flood Management: A Guide to Best Practice*. Environment Agency technical report FD2308/TR2, HR Wallingford.

JBA, 2006. *Making Space for Water. Flooding from other sources* (HA4a).

Kirby, A.M. and Ash, J.R.V., 2000. *Fluvial Freeboard Guidance Note*. Environment Agency R+D technical report W187.

Appendices

Appendix A: Literature review

Scoping the development and implementation of Flood and Coastal RASP Models (SC050065)	**Summary:** Delivers a plan for the development and implementation of modelling tools for flood risk assessment to support FRM planning and decision making. Identifies and describes existing RASP (Risk Assessment for System Planning) methods, and proposes future RASP tools. **Relevance:** Advocates a coupled modelling approach to flooding from all sources. This may mean future RASP outputs may already combine additional sources.
MDSF2 development (SC050051)	**Summary:** Describes the updating of the Modelling Decision Support Framework (MDSF) to include RASP methods. **Relevance:** Specifies default RASP outputs as probability of inundation (based on 0 m exceedance).
Spatial coherence of flood risk (SC060088)	**Summary:** Describes development of a spatial model of the statistics of extreme flows and tide/surge. The statistical model is used to generate scenarios of high flows, which are fed into a simple damage regression model to produce damages. Hence a damage–return period plot can be produced, with the dependence between loads producing a result somewhere between the independent and completely dependent cases. **Relevance:** When combining different sources or maps of flooding, if the dependency is unknown, we can use the dependent/independent cases as upper/lower bounds for the combined probability. Reliance on gauged flows (rather than rainfall) means that the method is not suitable for looking at dependence between fluvial and surface water flooding.
Joint probability: dependence mapping and best practice (FD2308)	**Summary:** Describes spatially distributed dependencies between a number of variable pairs related to flood sources. Also includes software (spreadsheet tool), and methods for making climate change allowances. Concludes that dependence between river flow and daily maximum sea surge (significant at the 5% level) may be found at catchments on most south, west and northern UK coastlines, and higher dependence is generally found in hilly areas with a southerly to westerly aspect. Dependence is also strong for flows lagged one day after the surge. Dependence between precipitation and daily maximum surge is strongest when they occur on the same day, and not particularly strong for any lag. Dependency statistics are produced for up to 72 catchments around the UK. **Relevance:** Primary source of dependence information for many sources, and modelling results are likely to be in formats implied by the use of these methods. Variable pairs with dependencies reported are: • wave height and sea level • river flow and surge

	• hourly rainfall and sea level • wind-sea and swell • wave height and surge • tide and surge • daily rainfall and surge. Modelling outputs using this approach are likely to be in the form of water levels for a number of joint probability situations (e.g. 100 year flow + 5 year sea level). Maximum water levels are then taken for each location in the model study area.
Information on coastal extremes (SC060064)	**Summary:** Ongoing project aiming to provide improved information on extreme sea level and swell conditions. **Relevance:** No information about dependency is given in current project outputs. Format of extreme probabilities may influence future mapping products.
Making Space for Water: **Flooding from other sources (MSfW HA4a)**	**Summary:** Investigates the potential for mapping flood risk from sources other than those already covered by the Environment Agency's flood map, including groundwater, integrated urban drainage, overland flow and embankment breaching. Appendix B and Appendix C discuss groundwater flooding in detail. Reviews hard-rock flooding, alluvial flooding and groundwater rebound. Mainly literature review, but does include simple modelling studies that inform. **Relevance:** Provides comprehensive overview on data and modelling availability for mapping these flood risks. Key technical conclusions are: • Flood maps for other sources are likely to be less accurate than existing flood maps for fluvial and coastal flooding. • Undertaking a national flood risk assessment incorporating risk maps for the canal system has been demonstrated to be feasible. Such a map may therefore be available in the future. • A reservoir risk ranking approach (for reservoirs covered by the Reservoirs Act 1975 and others) may be useful in assessing reservoir failure probabilities. • Cost of generating urban (pluvial and sewer flooding) flood hazard maps of accuracy equivalent to fluvial/coastal maps may be prohibitively expensive. • 40% of the damages for the autumn 2000 floods were to properties outside the fluvial/coastal indicative floodplain (according to insurance industry estimates). • 14% of the properties flooded in the autumn 2000 floods were due to 'non-river' causes (according to the Environment Agency). • Canal breach events in Britain have occurred approximately once a year since 1750. This is the most comprehensive text on groundwater

	flooding that has been reviewed – very useful. Presents novel scoping methods for mapping groundwater flooding including using BFI HOST; and a methodology for mapping potential for groundwater emergence using a methodology similar to LDS23, with some improvements that take into account non-uniform changes of groundwater head across a catchment.
Making Space for Water: **Urban flood risk and integrated drainage – IUD pilots (MSfW HA2)**	**Summary:** Describes investigations into urban flooding issues from rainfall, sewer exceedance and foul water for 15 urban and semi-urban areas. **Relevance:** Highlighted the piecemeal approach to urban drainage adopted in the past. National or nationally consistent datasets are unlikely to be developed in the short term.
Making Space for Water: **Groundwater flooding records collation, monitoring and risk assessment (reference HA5) Extended Report (Chalk Aquifers) Final Report** **Jacobs, November 2006**	**Summary:** This report (Chalk aquifers) recommends that a national database collating records of flooding from all sources (including groundwater) is both desirable and feasible. Recommends that the existing Flood Reconnaissance and Information System (FRIS) is used as an intermediary database to collate records of groundwater flooding in all Areas while the National Flood and Coastal Defence Database (NFCDD) is being further developed. The national database should ultimately be populated with all known electronic databases of groundwater flood records. **Relevance:** Emphasised the importance of collation of data regarding groundwater flooding. Includes a survey of existing data.
Making Space for Water: **Groundwater flooding records collation, monitoring and risk assessment (reference HA5) Initial Statement (non-Chalk Aquifers) Final Report** **Jacobs, September 2006**	**Summary:** Discussions with Environment Agency hydrogeologists provided the basis for an assessment of the occurrence of groundwater flooding in non-Chalk aquifers in the absence of specific literature. Examples of groundwater flooding in unconsolidated intergranular floodplain gravels have been identified by several hydrogeologists. It is likely that groundwater flooding via these aquifer units in floodplains is common but is masked to a degree by misreporting as fluvial flooding, which often follows. Similarly, groundwater flooding in higher terrace gravels is often difficult to distinguish from other causes of flooding, although its impacts can be significant. Groundwater flooding in non-Chalk consolidated intergranular, fracture and karst aquifers is rarely reported and the hydrogeologists consulted only identified relatively localised mechanisms, with the exception of groundwater rebound. An assessment of these aquifer types based upon their hydraulic characteristics does not identify any mechanisms likely to produce significant groundwater flooding under the present climate. The report recommends the use of the same national database as for Chalk aquifers to hold data on non-Chalk and notes that flooding is generally local (not regional), hence requiring local record-keeping and understanding of hydrogeological systems. **Relevance:** Discusses further (i.e. non-Chalk) mechanisms of groundwater flooding. Emphasised the importance of collation of data regarding groundwater flooding.
Making Space for Water: **A preliminary risk assessment of the potential for groundwater**	**Summary:** Made an initial assessment of the likelihood of groundwater flooding in the Chalk following very high summer rainfall in 2007; based on past behaviour of groundwater levels as recorded in observation wells.

flooding during the winter of 2007/8 – an update. **J Finch, T Marsh, A McKenzie, October 2007 (CEH/BGS)**	Covered areas with recorded groundwater flooding, because of the need for observational evidence to provide predictions. Assessed rainfall anomalies, soil moisture deficit and groundwater levels in summer 2007 to predict potential flooding on regional basis the following winter. **Relevance:** Not high for map production, although useful for early warning.
Broad-Scale Modelling Scoping – a vision for flood modelling and risk science (FD2118)	**Summary:** Aims to provide a medium to long-term direction for the Modelling and Risk Thematic Programme, including the specific issue of the feasibility of an integrated modelling system. **Relevance:** The Drivers–Pressures–States–Impacts–Response broad-scale model may influence the choice of method for this project and the way in which sources or flood maps are combined.
Software requirements for Joint FCERM R&D programme modelling outputs and architecture specification for RASP family outputs (FD2121)	**Summary:** Provides guidance for research contractors for producing Environment Agency compatible software. **Relevance:** The guidance will need to be taken into account at prototyping stage (Phase 2 of this project) to ensure the architecture is compliant with Environment Agency systems.
Risk, Performance and Uncertainty in Flood and Coastal Defence – A Review (FD2302)	**Summary:** Advocates the Source–Pathway–Receptor model, and the definition of risk through probability and consequence. Performance is emphasised in flood risk management, rather than reliance on standard of protection. Recommends that uncertainty should be included with model outputs, and should be an integral part of decision making processes. **Relevance:** This indicates that uncertainty should be included where possible in an integrated flood map.
National flood risk assessments (NaFRA)	**Summary:** National estimate of annual flood damages from fluvial and coastal sources, based on RASP methodology. **Relevance:** NaFRA outputs will be available for all main rivers and coastal floodplains in England and Wales, and so could form the basis of fluvial/coastal map of flooding from all sources. The output formats will need to be compatible with the prototype tool. Focus is on defence systems, so may be less reliable in undefended areas. NaFRA calculates fluvial and tidal/coastal probabilities separately and then combines them with an assumption of independence.
Thames Estuary 2100 (TE2100): flood risk analysis work – ECO+ and IA8 methods and uncertainty analyses	**Summary:** Hydraulic modelling and flood impact calculations for baseline scenarios and a range of management options. Outputs are depth–probability grids combining overtopping and breaching for defended areas, direct damages to property and risk to life estimates. Uncertainty for the direct damages is estimated for each embayment. **Relevance:** The methodology for combination of breach and overtopping grids may be useful, as it is an example of combining probabilistic depth maps. Uncertainty is not available for depth–probability grids, but the damage estimate uncertainty could be used as a proxy for the quality of the depth information.
Mapping of Areas Susceptible to Surface Water Flooding	**Summary:** Report contains an evaluation of the JBA map indicative of surface water flooding, and recommendations

	for its use. **Relevance:** This surface water map may be one of the sources of flood information to be combined. This will be difficult to impossible as the map does not correspond to depth at a particular return period etc. Phase 2 of the project may produce a calibration of the risk (e.g. middle band corresponds to risk of 0.1% etc.).
Foresight Future Flooding	**Summary:** Project evaluates the effect of future climate change scenarios on flooding in the UK from coastal, fluvial and urban pluvial sources. **Relevance:** The relative impacts of changes in different drivers of flood risk in 2080 depend on climate change scenarios. There is therefore likely to be a change in the relative risks from different sources, but this can only be estimated with a large uncertainty.
Flood Risk Management Research Consortium (FRMRC and FRMRC2)	**Summary:** Phases 1 and 2 of academic research project to investigate underpinning science for flood risk management. Phase 1 produced a number of UFMOs (User Focused Measurable Outcomes). Phase 2 outcomes not available yet. **Relevance:** Flooding from multiple sources not covered. Some uncertainty outputs may be of use if uncertainty is to be taken into account in the prototype.
DTI SAM Modelling of Space–Time rainfall for System Based Analysis and Management of Urban Flood Risks	**Summary:** Describes the development and parameterisation of a statistical model of rainfall, based on radar observations, to represent the spatial and temporal structure. This will be used as input to urban drainage and flooding models. The study has aimed to develop a long synthetic time series of spatially varying rainfall data derived from observed radar sequences. However, the data do not compare very well with FEH rainfall statistics, particularly at shorter durations, principally as too few rainfall events were present in the radar archive data used to develop the synthetic time series. **Relevance:** It may be possible to gain information on the dependency between different sources (fluvial and surface) from the model, but not straightforward (beyond the scope of this project). Other outputs may prove useful, but are yet to be published on the website. Dissemination Seminar on 19 May should highlight prime use of outputs. In future, if this study were enhanced by a longer and more data-rich body of radar data, outputs could be particularly useful for examining spatial aspects of rainfall and flooding impact on urban environments – aspects such as rainfall event spatial extent, spatial distribution of high intensity cells within an event, speed of travel and event duration could all be examined.
FLOODsite **http://www.floodsite.net**	**Summary:** Major EU research project on integrated flood risk analysis and management methodologies. Potentially relevant work includes: joint probabilities/dependence mapping, failure modes/breaching, inundation modelling and uncertainty analysis. **Relevance:** The research mentioned above plus information on comparison on European practice.
Strategy for Flood and Coastal Erosion Risk Management: Groundwater Flooding Scoping	**Summary:** Scoping-level assessment of the scale, nature and distribution of groundwater flooding in England. Focuses on groundwater flooding from permeable hard-

Study (LDS23), Jacobs 2004	rock aquifers (particularly Chalk), groundwater rebound and mine water rebound. Estimates the number of properties at risk from groundwater flooding. **Relevance:** Includes the development of 'GEMs' (Groundwater Emergence Maps), which 'provide a readily accessible means of identifying those areas of England in which groundwater can rise to within the top 2 m of the ground surface in an unusually wet winter. The maps cover all major [consolidated] aquifers in England and have been calibrated on the exceptionally wet winter of 2000–2001. The GEMs provide a means of identifying broad areas at risk from groundwater emergence. There is no attempt at quantification of risk.
SE Morris, D Cobby and A Parkes, 2007. Towards groundwater flood risk mapping. Quarterly Journal of Engineering Geology and Hydrogeology 40, 203–211	**Summary:** A discussion and case study of developing groundwater emergence maps (GEMs), as presented in LDS23. **Relevance:** Nothing newer in here than the work reported in LDS23 but its publication lends the credibility of peer review to the approach.
SM Taylor, 1995. The Chichester Flood, January 1994, paper to Brit. Hydrol. Soc. 5th Symposium	**Summary:** Assesses causes and course of Chichester flood event in winter 1993/94. Chalk stream with alluvial sediments in coastal plain (flood area). Notes that normal Chalk stream behaviour (delayed response to rainfall of groundwater levels (GWLs) and discharge to river) was changed because of high GWLs. Suggested that this was due to high recharge in previous months, leading to: a) GWLs rising to saturate much or all of the available aquifer storage in both Chalk and superficial deposits; and b) GWLs rising into a zone of higher Chalk transmissivity. Both these factors lead to rapid discharge of infiltrated rainwater and increased runoff. Identified a 'threshold level' at a Chalk well in the catchment. **Relevance:** Identifies mechanisms for groundwater flooding in this Chalk catchment. Suggests early warning potential of GWL monitoring in strategically located wells.
RB Bradford and KM Croker, 2007. Application of head-flow responses to groundwater floods in Chalk catchments. Quarterly Journal of Engineering Geology and Hydrogeology 40, 67–74	**Summary:** Applies regression analysis to synchronous head-flow data to assess likely groundwater flooding thresholds. **Relevance:** May allow better quantification of risk if head-flow data can be correlated for a catchment.
BGS (A Hughes and J Bloomfield), CEH	**Summary:** Research project to develop Chalk catchment models, linking existing groundwater flow models to rainfall and runoff models (CEH JULES) via a spatially distributed unsaturated zone transfer function. Will also investigate the potential use of statistical techniques to link groundwater level and rainfall using historical data. This will provide quicker, dirtier, regional indications. **Relevance:** Statistical correlations could be used to indicate potential for groundwater emergence at high rainfall; could provide return periods, but unlikely to be accurate because the datasets will not contain many examples of the extremes required for flooding.
Najib et al. 2008	**Summary:** Focuses on water table rises that are induced by exceptional rainfall events in fractured and karstified carbonate aquifers. Methodology uses statistical analysis common in flood frequency analysis. The developed

	approach finally results in a simple methodology devoted to the assessment of extreme groundwater surge hazard. The main hydrogeological conditions under which the proposed approach can be applied are the following: (1) a thin unsaturated zone (shallow unconfined aquifer), (2) a high groundwater recharge potential and (3) a high hydraulic diffusivity. Head and rainfall data must be available for an extreme event of the scale being predicted, and over at least one hydrological cycle. **Relevance:** May be useful for quantifying risk.
AA Mckenzie, JP Bloomfield, A Hulbert and HK Rutter, undated. Confidence and Groundwater Flood Susceptibility Mapping	**Summary:** Assesses the confidence attributable to groundwater flood vulnerability maps for Chalk-type and 'PSD' (permeable superficial deposits) aquifers. Provides a method for producing confidence maps to go with the vulnerability maps. **Relevance:** Useful corollary to vulnerability maps like GEMs, but does not get any closer to return frequency.
Groundwater Flooding Susceptibility – BGS dataset	**Summary:** Map-based hazard assessment, using geological data and comparison of groundwater contour maps to DTM heights. National coverage. Not related to actual flooding records or to frequency of return. Developed (for permeable hard-rock aquifers) in the same way as GEMs. **Relevance:** Can provide useful comparison of risk within any given aquifer unit, but results are incomparable between aquifers.
Investigating the interdependencies between surface and groundwater in the Oxford area to help predict the timing and location of groundwater flooding and to optimise flood mitigation measures – MacDonald et al. undated	**Summary:** Intensive investigation project looked at influence of groundwater on surface water and vice versa, provided flow map for alluvial sediments in the Thames valley through Oxford. Project aims to link surface flow model (ISIS) with groundwater flow model (ZOOMQ). **Relevance:** May be too site-specific to provide generic information on groundwater flooding in alluvial aquifers, but is a useful case study.
B Adams, J Bloomfield, A Gallagher, C Jackson, H Rutter and A Williams, 2008. FLOOD1 final report	**Summary:** The project was set up to develop appropriate early warning systems for groundwater flooding in Chalk catchments. Three main objectives were addressed by the project: to understand the hydraulic behaviour of water flow in the unsaturated zone which leads to triggering of groundwater flood events; to develop unsaturated zone monitoring techniques (though this was done by BRGM, a French geological mapping enterprise, and is not reported here); and to produce more appropriate methodologies and tools for forecasting groundwater flood events capable of operating within a much longer timescale than is currently possible (i.e. days and weeks rather than hours). Various modelling approaches to simulate high groundwater levels were tried and the results summarised. A procedure for early warning was developed. **Relevance:** Development of early warning system with potential for wider use, if based on good local data.
http://www.hochwasser-dresden.de/MULTISURE	MULTISURE – Development of Multisequential Mitigation Strategies for Urban Areas with Risk of Groundwater Flood – this is a German research project, running to end 2009. There are other associated German research projects, e.g. at http://elise.bafg.de/?7336.

Tsunamis – Assessing the Hazard for the UK and Irish Coasts, HRWallingford, BGS, Proudman Oceanographic Laboratory, 2006	**Summary:** Assesses the impact of an earthquake similar to that off the coast of Lisbon in 1755, in terms of water levels. No assessment of the probability of such an event is given. **Relevance:** The report states 'Only the most southwesterly coast of the UK may incur sea level elevations marginally in excess of the 1:100 year extreme sea level predictions', indicating that tsunami risk is low in relation to other coastal flood risks.
FRACAS **http://badc.nerc.ac.uk/data/free/fracas.html**	**Summary:** Part of the NERC Flood Risk from Extreme Events programme, assessing future flood risk at national scale. Similar approach to NaFRA, but with future rainfall scenarios. Uncertainty included in the approach. **Relevance:** Estimating future flood risk will need to take these kinds of output into account.
Europe (EXCIMAP) Handbook on good practices for flood mapping in Europe	**Summary:** Collects best practice in production and communication of flood maps throughout Europe. Lists the types of flood map currently in use. **Relevance:** Presents a wide range of flood hazard communication methods which may be useful for a map of flooding from all sources.
EU Floods Directive	**Summary:** Outlines the responsibilities of governments in assessing and managing flood risks. **Relevance:** Defines floods with low, medium and high probability. Groundwater flooding and flooding of well-defended areas may be limited to low probability events.
Pitt Review	**Summary:** Review of lessons learned from the summer floods of 2007. **Relevance:** The report makes a number of recommendations, in particular: Recommendation 2: *The Environment Agency should be a national overview of all flood risk, including surface water and groundwater flood risk, with immediate effect.* This means the Environment Agency will have to deal with maps from different sources and agencies. Recommendation 5: *The Environment Agency should work with partners to urgently take forward work to develop tools and techniques to model surface water flooding.* Highlighting the need to tackle and understand surface water flooding mechanisms better and how to manage the risk. Recommendation 6: *The Environment Agency and the Met Office should work together, through a joint centre, to improve their technical capability to forecast, model and warn against all sources of flooding.* Means a more holistic approach to flood forecasting is required. Recommendation 57: *The Government should provide Local Resilience Forums with the inundation maps for both large and small reservoirs to enable them to assess risks and plan for contingency, warning and evacuation and the outline maps be made available to the public online as part of wider flood risk information.* Recommendation 61: *The Environment Agency should work with local responders to raise awareness in flood risk areas and identify a range of mechanisms to warn the public, particularly the vulnerable, in response to flooding.*

	Combined maps may help in raising public awareness.
PPS25	**Summary:** Guidance to 'ensure that flood risk is taken into account at all stages in the planning process'. Annex C describes 'forms of flooding' covering sources: rivers; sea; land (surface water); groundwater; sewers; reservoirs, canals and other artificial sources. The assessment of flood risk should 'consider and quantify the different types of flooding (whether from natural and human sources and including joint and cumulative effects)'. **Relevance:** Final output flood maps should cover, where possible, the sources listed here, and could be suitable for aiding planning decisions where the supporting flood hazard data are available.
Extreme Rainfall Alerts – for surface water flooding **Halcrow Group Ltd, Proposed Pluvial Flood Warning Trial Service, May 2008, report produced for the Met Office**	**Summary:** Report setting out the proposed format for pluvial flood alerts, which became known as the Extreme Rainfall Alert (ERA) service and went live as a trial operational service in July 2008. Details the problems behind forecasting for urban surface water flooding and sets out a proposed methodology for developing a service, including the meteorological requirements. Developed a set of rainfall trigger thresholds which, when forecasters predict may be exceeded, trigger the issuing of alert messages to Category 1 and 2 professional response organisations. **Relevance:** The ERA service became adopted as a fully operational service in January 2009 and from 1 April 2009 is delivered from the new national Flood Forecasting Centre (see http://www.environment-agency.gov.uk/105596.aspx). The ERA currently provides the only method for alerting organisations to the risk of surface water flooding, ahead of the rainfall falling. It therefore plays a significant role in flood risk management – particularly as surface water flooding was such a problem in the summers of 2007 and 2008. This report provides the theory behind the service's origin.
Extreme Rainfall Alerts – for surface water flooding **Halcrow Group Ltd, Post-event analysis of the ERA Trial Service between 2 July and 16 September 2008, November 2008, produced for the Met Office and Environment Agency**	**Summary:** A post-event analysis study of the ERA trial service over its first 7 weeks since inception. A period of numerous, country-wide surface water flooding incidents, this study reports on ERAs issued for ten separate/discrete rainfall events between 2 July and 16 August 2008. The report examines the reasoning behind ERA issues that were made (Met Office forecasting statistics), estimates of the rain that caused surface water flooding incidents and feedback from a representative sample of users of the trial service. Recommendations are made for service improvements. **Relevance:** Provides evidence of the use of surface water flooding alert messages by Category 1 and 2 response organisations and how actions taken have resulted in mitigating the impact of flooding in urban locations.
FREE Programme – Flood Risk from Extreme Events (NERC-funded research programme)	**General Summary:** Flood Risk from Extreme Events (FREE) is research to predict floods minutes to weeks and seasons to decades ahead. The programme uses environmental science to investigate the physical processes involved in generating extreme events, so they can be better forecasted. The FREE programme is researching what causes and propagates floods, so helping to forecast and quantify flood risk, and inform our society about the likely effects of climate change. FREE brings researchers in the hydrological, meteorological, terrestrial and coastal oceanography communities together in an integrated

	research programme for the first time. **General Relevance:** Many of the FREE Programme's eight objectives are concentrating on forecast uncertainty and dealing with risk and uncertainty in flooding. Hence this research programme could be of use in this project in providing guidance on uncertainty in flooding science – concerning pluvial, fluvial and coastal flooding sources principally. **Summary (GW):** There is a groundwater flood modelling project under FREE. The project runs 2007–2010, and looks particularly at the Chalk of South East England. The nature of groundwater flooding risk is poorly understood, and there is no adequate methodology to quantify it. The project proposes to integrate state-of-the-art models of the soil, unsaturated zone, groundwater and surface water to provide a new modelling tool for risk assessment, and investigate the use of simpler models for warning of the potential onset of flooding and regional assessment of risk. Historical data from affected areas will be used to test hypotheses and develop and validate the models. The models will be run for future climate states to assess current and future risk using an ensemble of climate models. No project reports yet available. **Relevance (GW):** Limited relevance to this project, as it is outside the time frame. But outputs should be of value, particularly concerning quantification of risk.
Forecasting Options for Rapid Response Catchments (Halcrow report for the Environment Agency)	**Summary:** Considers and evaluates potential options for flood forecasting in rapid response catchments (defined as T_p<3 hours). Recommends that rainfall thresholds at appropriate durations are developed for each catchment. **Relevance:** Small catchments are those for which fluvial and pluvial flooding are more likely to coincide, as these flood events are driven by the same types/locations of rainfall events, and there is less sensitivity to antecedent conditions in rapid response/urbanised catchments. The results could be used to infer joint probabilities for pluvial/fluvial flooding.
British Waterways Guidance: Hydraulic Design of Canal Works Good Practice Guide	**Summary:** Guidance for BW engineers on hydraulics of canals, including response to breaching. **Relevance:** BW has established an archive of historical breach failures along its waterways (available internally). BW is currently researching further the mechanisms of breach failures along with the identification of key indicators of high risk sites.
TE2100: Other Sources of Flooding and Effects of Interventions (Jacobs, for the Environment Agency)	**Summary:** Assesses hazard in the Thames basin from sources other than fluvial and coastal: surface water, sewer flooding and groundwater. Also assesses impact of flood risk management options on hazard from these other sources. Includes climate change scenarios. **Relevance:** Identifies two groundwater flood mechanisms: unconfined Chalk-basal sands; and permeable superficial aquifers hydraulically connected to in-bank river and tide/surge levels. High level vulnerability method for sewer flooding identified.

Appendix B: Consultation questionnaire responses

Response 1
What sources of flooding should this project include?
• Fluvial flooding • Coastal flooding – due to tide, storm surge, extreme wave action • Pluvial or surface water flooding • Groundwater flooding • Sewer flooding • Combined source flooding – especially *surface water* + *fluvial* and *fluvial* + *tidal*
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
• Fluvial flooding – lots: level and flow data from gauging stations, flood intelligence files (FIFs), rainfall data, aerial photography, wrack marks, archive media material. Yes metadata in FIFs etc., not much IPR problem • Coastal flooding – due to tide, storm surge, extreme wave action – tide level data, storm surge data, storm surge forecast data, extreme tide level estimates produced for the Environment Agency • Pluvial or surface water flooding – rainfall data, media reports of surface water flooding, Halcrow report on Post-event analysis of the ERA • Groundwater flooding – Borehole records, held by Environment Agency, SEPA and water companies • Sewer flooding – data mainly with water companies and DG5 records, captured in CFMPs to some extent • Combined source flooding – especially *surface water* + *fluvial* and *fluvial* + *tidal* – data sources as above, combined probability study for Defra FD2308
There are two methods proposed for combining flood maps from different sources: 1. The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2. The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
I would favour the flood maps approach since it will be more transparent as to the source of the flooding. Also, the sources option assumes that flood water will combine to give a combined volumetric total – but how do you combine groundwater and fluvial flood flows for example? Flood hazard is determined from both depth and velocity – the flood maps option stands the better chance of capturing hazard than the sources option I would think.
What will be the key technical challenges in delivering this project?
• Data collation – a lot to deal with • Data availability • Combining the sources in a realistic and hydraulically sound way to derive realistic maps of multiple flood sources • Assessing anything other than flood depth – e.g. velocities
How should the combined flood map be communicated? E.g. 1% flood depths, probability of exceeding 0.5 m, etc?
I would say it is best to relate to flood hazard via the Defra definitions for this. That should indicate what depth is critical for certain velocities so if velocities are not known they can be assumed (e.g. 1 m/s or something). Flood depths could then be mapped related to high and medium hazard perhaps. I think probability of exceeding xm depth (which is high hazard or whatever) is better than the % method as it communicates uncertainty to some degree.
What data are likely to become available in the next 5 years?

- More post-event data on surface water flooding incidents that couple up both rainfall statistics (intensities and durations) with on-the-ground impact. This should be an output from the recently developed Flood Forecasting Centre in order to improve the future effectiveness of surface water flooding forecasts from heavy rainfall.
- More accurate and higher resolution DTM data.

Response 2
What sources of flooding should this project include?
River flow, tidal flow (plus surge), tsunami, wave overtopping (local flooding).
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
All tidal data (excluding surge) can be generated in Halcrow's SANDS data system. British Met Office wave data for some locations are also available in SANDS.
There are two methods proposed for combining flood maps from different sources: 1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 –The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
A single probabilistic map associated with flood depth is appropriate.
What will be the key technical challenges in delivering this project?
1. Generate common boundary conditions for different sources of flooding. 2. Combine flood maps from different sources of flooding. 3. Consider three (or more) sources of flooding at the same time. 4. Do we consider the joint probability (for two events) or marginal probability as we consider the next point below?
How should the combined flood map be communicated? E.g. 1% flood depths, probability of exceeding 0.5 m, etc?
Please see the point 4 above.
What data are likely to become available in the next 5 years?
It is likely that: 1. Data of sea level rise should be available in the next 5 years. 2. Measured tidal level data should be available for certain locations in the next 5 years. 3. Wave data can be obtained from Met Office.
Response 3
What sources of flooding should this project include?
Considering NaFRA 2008/MDSF2: fluvial, tidal and coastal, overtopping and breaching.

What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
The NaFRA 2008 analysis requires input data for pre-processing tools activities as described in NaFRA 2008 training manual. The NPD3 dataset for economic pre-processing is also required for economic direct property damage calculation and MDSF2 is known to further require an Agricultural Land Classification dataset for the calculation of agricultural damages. Metadata status of datasets is not identified; it is understood that the Agency hold all datasets and appropriate IPR, with the exception of the Cumulative Catchment Area dataset (CCAR) which is CEH licensed and possibly datasets obtained from TE2100 models. Key flood mapping results available in the NaFRA/MDSF2 model databases (SQL server/Oracle) are as follows: • tblCellDepthProbability – providing the probability of flooding over specified depth thresholds • tblCellFloodProbability – providing the probability of flooding for a depth greater than 0 m • tblCellLocation – providing the X, Y location of each cell It is important to note that NaFRA (and MDSF2) produce two separate databases for a given 'catchment'; one fluvial database and one tidal/coastal database. These probabilities are combined (from tblCellFloodProbability from the fluvial and tidal/coastal databases) and an overall probability calculated in viewCellFloodProbabilityOverall (assuming independence of flood source drivers) and subsequently exported for use in mapping.
There are two methods proposed for combining flood maps from different sources: 1 –The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 –The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
The NaFRA and MDSF2 approach to estimating flood extent and estimated annual damage from direct property damage (and agricultural damage) is becoming well established in the Environment Agency and provides a suitable platform for considering fluvial, tidal and coastal flood extent mapping. It is understood that a requirement of Phase 1 of the current project is to integrate with existing RASP methods, and as such results from MDSF2 may be expected to form a central component of the project. Whether sources of flooding are combined into an MDSF2 style approach or whether resultant flood extent maps are combined may in part be best considered on a case by case basis, and whether a probabilistic/fragility curve approach to asset failure, as applied in NaFRA and MDSF2, is appropriate. The development of MDSF2 is thought to provide a suitable framework under which the current consideration of probabilistic failure of fluvial, tidal and coastal defence assets is extensible. It may further be appropriate to revise the current scope of the NaFRA/MDSF2 consideration of fluvial, tidal and coastal flooding to only that part of the system that is most suited to a probabilistic assessment of asset failure, i.e. do not apply the approach in undefended (fluvial) areas that may best be served by a deterministic approach. This would however require integration with competing models for (fluvial) flood risk mapping.
What will be the key technical challenges in delivering this project?

To integrate results from different sources of flooding producing a credible overall assessment for flood extent mapping. To ensure that each component produces credible results. E.g. NaFRA results are sensitive to the estimation of volume entering the floodplain. Are these estimates credible? The framework needs to allow for validation of components against other more detailed approaches (e.g. the fluvial flood mapping component could be considered deterministically to validate approach prior to consideration of the probabilistic failure of assets). There will be a clear need to establish the framework for the assessment of joint probability of different sources of flooding and the degree of dependence between results. To clearly identify whether the study is to consider flood hazard or flood risk. To clearly define the approach to considering spatial coherency of results, and how results may be aggregated.
How should the combined flood map be communicated? E.g. 1% flood depths, probability of exceeding 0.5 m, etc?
What data are likely to become available in the next 5 years?
MDSF2 framework, and results from application of MDSF2.
Please feel free to add any other information you feel is relevant to this project.

Response 4
What sources of flooding should this project include?
Pluvial Ordinary watercourse Main river Coast Groundwater Sewer exceedance
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
Water Companies can and do generate sewer exceedance maps – though they'd be very reluctant to share openly and are nervous about interpretation.
There are two methods proposed for combining flood maps from different sources: 1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 -The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
I'd work with the maps – better to accommodate modelling that has already been done. Overall map can be enhanced once improved data comes along. Historical data could be included too.
What will be the key technical challenges in delivering this project?
Separating likelihood and duration. Groundwater (very long duration) vs. surface water (very short duration). Is a common probability of occurrence meaningful?
How should the combined flood map be communicated? E.g. 1% flood depths, probability of exceeding 0.5 m, etc?
I think probability of depth reaching certain thresholds (e.g. 0.1 to 0.5 m and > 0.5 m) is useful – though if you're using existing maps it might not come in this format.
What data are likely to become available in the next 5 years?
SWMPs will map surface water flooding in many towns – hopefully in a consistent way!

Response 5
What sources of flooding should this project include?
Flooding due to dambreaks or overtopping of dams.
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
Flood inundation extents for all statutory reservoirs in England and Wales (ongoing project by JBA for the Environment Agency). Ongoing project to prepare inundation maps for all statutory reservoirs in England and Wales. Detailed dambreak studies on individual reservoirs undertaken by reservoir owners.
There are two methods proposed for combining flood maps from different sources: 1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 – The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
What will be the key technical challenges in delivering this project?
How should the combined flood map be communicated? E.g. 1% flood depths, probability of exceeding 0.5 m, etc?
What data are likely to become available in the next 5 years?
Detailed flood maps prepared by reservoir owners as directed by Water Resources Act 2003. Maps will be required by 2013.
Please feel free to add any other information you feel is relevant to this project.

Response 6
What sources of flooding should this project include?
- Fluvial   - Coastal   - Drainage surcharge   - Pluvial
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
I think detailed modelling is secondary for this type of high level assessment. The main data required is an accurate (± 200 mm) DEM and a reliable estimate of the volume inflow. In the majority of cases the assumption the water level is horizontal and that flood water ponds in low lying areas will be perfectly adequate. There will, of course, be more complex areas where higher level models will be necessary but these will be the exception rather than the rule.
There are two methods proposed for combining flood maps from different sources: 1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 – The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
I would prefer the second as I think the second will be easier for non-technical stakeholders to understand.
What will be the key technical challenges in delivering this project?
Acceptably accurate estimates of flood flows from surcharging drainage systems. If this can be estimated then spreading it over the surface is a relatively trivial modelling exercise; however, the surcharging volume is likely to be the result of complex hydraulic interactions which may require detailed modelling to resolve.
How should the combined flood map be communicated? E.g. 1% flood depths, probability of exceeding 0.5 m, etc?
Not sure.
What data are likely to become available in the next 5 years?
Greater high resolution DEM coverage.
Please feel free to add any other information you feel is relevant to this project.

Response 7

What sources of flooding need to be considered in delivering a single integrated map of flood hazard?

In my view the following are the main sources of flooding which should be considered as part of this project:

Fluvial
Coastal
Urban
Groundwater
reservoirs

What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?

Fluvial: EA flood map
Coastal: EA flood map
Urban: EA surface water map mk 1 and 2 (plus any maps from detailed studies conducted as part of SWMPs)
Groundwater/reservoirs: currently under preparation

Only the EA fluvial and coastal flood maps are currently cleared for external publication.

There are two methods proposed for combining flood maps from different sources:

1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain

2 – The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth)

Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?

Number 2 is much more feasible and will be able to give much better results within the timescale envisaged for this project.

Before extending the RASP/NaFRA/MDSF2 approach it needs to be investigated why such a complex approach may be necessary and which benefits it may have. This would only make sense if the EA e.g. were to prioritise investment to tackle urban flooding nationally as part of the inland overview. To my knowledge is not yet planned.

The existing RASP tools are already fairly complex systems and still need further development to accurately model river and coastal flooding. To add additional functionality at this stage would increase the complexity even further and introduce risks.

In addition, significant changes with regards to the way (urban) defences are represented would be required. It may also be necessary to move to a rainfall runoff type approach to allow for the representation of urban flooding which is harder to represent through the use of in-river water levels.

Note that there is also a proposed second project in the EA R&D project (aka as the pluvial method) which could pick up point 1 if necessary).

However, for both approaches a sound method for assessing the joint probability of flooding from different sources will be required. In my view this will be the most important and challenging part of the project. EA R&D project SC060088 further developed a method by Hefferman and Tawn (2004) which may provide a good starting point for this.

What will be the key technical challenges in delivering this project?

A sound method for assessing the joint probability of flooding from different sources will be required. In my view this will be the most important and challenging part of the project. EA R&D project SC060088 further developed a method by Hefferman and Tawn (2004) which may provide a good starting point for this. The other major challenge would be to define at which scales the project outputs can be used (national, regional, catchment, site…) and to ensure the proposed approaches are fit for purpose and are not used at scales they are not aimed at.
How should the combined flood map be communicated? E.g. 1% flood depths, or probability of exceeding 0.5 m, others?
The most obvious approach would be to visualise probabilities that flooding occurs (e.g. flood depths is more than 0 m). Other thresholds (e.g. flood depths exceeding a generic property threshold would also be useful) but may need to be aimed more towards a specialist audience (e.g. professional partners) and should be accompanied by relevant guidance. In addition, maps of certain velocities (or other measures of hazard) being exceeded would also be useful.
What data are likely to become available in the next 5 years?
Improved surface water maps, groundwater and reservoir flooding maps. Improved river and coastal flood maps based on 2D hydraulic models. Coastal erosion maps….
Please feel free to add any other information you feel is relevant to this project. I think this is a very worthwhile project. But it is important to maintain a clear focus on what the outputs are meant to achieve. Developing and applying a robust method to assess joint probabilities and defining the areas and limits of application will be key for this.

Response 8
What sources of flooding need to be considered in delivering a single integrated map of flood hazard?
Rather than talking about just sources of flooding, we should consider sources and pathways of flooding. This is because sometimes the same pathway can flood from several different sources (e.g. astronomical tides, strong waves and a tsunami are all sources that cause flooding on the coastal pathway). Equally the same source can flood different pathways (e.g. a convective thunderstorm could cause sewer flooding, river flooding and surface water flooding). Using this concept would give a modular approach to mapping all sources of flooding where you could reuse pathway data for multiple sources and vice versa. This would have clear cost and resource benefits for those involved in FCRM. It would also be a holistic approach rather than just a bolt on to the existing work we do on river and sea flooding. As such, it also offers opportunities to improve our existing modelling allowing us to distinguish between convective thunderstorms and winter flood events, instead of just bundling them up together in the hydrology. The sources and pathways that should be considered are: **Pathways** – Rivers, the coast, sewers, aquifers, land surface **Sources** – Convective thunderstorm, winter storms, snowmelt, reservoir failure, burst water pipe, astronomical tides, tidal surge, waves, tsunami
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
Detailed river network – Properly archived – EA and OS owned **Flood outlines in NFCDD** (May include other sources of flooding) – properly archived – EA owned **LiDAR** for most of England and Wales – Properly archived – EA owned **Geological maps** – BGS owned (but EA has licence to use) **Snowmelt events** – SEPA may be able to help here **Sewers/Drains** – Local authorities or water authorities may be able to help here but likely to be incomplete data
There are two methods proposed for combining flood maps from different sources: 1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 – The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?

What RASP is good for is: being able to tweak parts of the system to test out different management options. However, for something like groundwater flooding there's not really a lot we could do other than warn people away from vulnerable areas. A 1% flood outline would be more than suitable for this task. Conversely, RASP for urban flooding would be incredibly useful as we have much more control over this type of flooding. Some types of flooding require a RASP approach and others require only a 1% flood map. The approach will depend on how much control we have over that type of flooding and the magnitude of threat that type of flooding poses. This means of course that the final overall map will have to be just a 1% flood map. This level of detail is perfectly acceptable for the jobs we want to do with an overall flood map – such as communicating about flood risk, and informing development planning. Anybody needing to do something more involved than this with the flood map will seek out more detailed data. One feature that must occur in the final overall map is the ability to tell which areas of the map are threatened by which types of flooding. This request is based on experience from the historical flood map that the EA publishes. The historical map shows an outline covering every area we have ever recorded as flooding. The first question anyone asks after having seen it is 'yes, but when did it flood there?'. We should try to avoid this situation with the all flooding map. The map should be able to tell people which type of flooding poses the greatest risk in each area of the map.
What will be the key technical challenges in delivering this project?
1. Focusing on the high risk types of flooding – We'll need to develop screening tools that allow us to focus on only the highest risk types of flooding because modelling every type of risk in every location is not feasible. **2. Communicating uncertainty and limitations of the data** – This project will bring together some very sophisticated methods and data, such as RASP, and use them with some very basic techniques, such as those currently available for surface water flooding. The danger is that we're not comparing like for like. The final overall map will need to show intuitively that some areas are high certainty and others are not. One solution to this could be to pixellate the map. Where the size of the pixel represents the level of certainty.
How should the combined flood map be communicated? E.g. 1% flood depths, or probability of exceeding 0.5 m, others?
A 1% flood outline is sufficient for the overall flood map but there must be more detailed data available behind the map to support it. (For example the RASP data for fluvial flooding.)
What data are likely to become available in the next 5 years?
▪ More detailed and better coverage LiDAR data. ▪ Improved Earth observation data from satellites which may help with mapping flood events (of any type) in real time. ▪ Better data sharing on sewers and drains (as called for by Pitt). ▪ Models of surface water flooding to support Surface Water Management Plans. ▪ Improved models of the impacts of climate change.

Response 9

What sources of flooding need to be considered in delivering a single integrated map of flood hazard?

Fluvial flooding

Coastal flooding

Surface water flooding from extreme rainfall including ponding, pathways and interactions with small watercourses where levels are too high for water to escape. This includes exceedance flooding where the underground drainage system is overwhelmed and water cannot escape into the underground drainage system.

One difficulty will be data and model acquisition for more frequent flooding from sewers – internal 'frequent' flooding caused by underperformance, or 'other causes' (blockage or collapse) is likely to be much harder to model as integrated above/below ground modelling is generally needed.

Groundwater flooding: unconfined aquifer, flooding through alluvials and rebound in post-industrial cities can all be important, but are regional.

Flooding from impoundments clearly has severe potential adverse consequences, but understanding the associated probability is far from easy, and I feel this should be represented separately.

What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?

All sources: JBA are collating a national geo-database of historic flooding records to meet interim solution of the Pitt Review. Some of these records include sewer flooding. [see also Areas susceptible to surface water flooding maps]
Groundwater – there are some detailed outlines from historic events available, JBA has also produced a risk map for flooding from aquifers.

There are two methods proposed for combining flood maps from different sources:

1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain

2 – The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth)

Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?

Option 2) is more feasible, but 1) would allow quantification of risk. Does quantification of risk help people understand? I think aspects of both 1) and 2) would be best. With RASP there is a danger of hiding behind a black-box and using very uncertain parameters such as fragility curves, and giving 'probabilistic' output without confidence intervals on those probabilities. People might understand measures such as there are 50 residential properties at risk on an annual average basis in their town from fluvial flood damage, which may be compounded by surface water flooding. Do they understand that there is a 1.33% AEP probability of flooding to a depth of 0.6 m from any source? Does it better prepare them? Correlation between pluvial and fluvial will be difficult to tease out.

I would also be careful that using integrated measures or probability does not smear out climate change sensitivity – we've been finding extreme pluvial outlines very sensitive to climate change.

What will be the key technical challenges in delivering this project?

Firstly deciding what to include – there will be a lot of expectations from different people and organisations wanting different sources/mechanisms included. See above sections for other comments. Key challenges will be understanding: 1) Dependency between fluvial and surface water flooding. 2) Spatial relationships between fluvial from multiple sources. 3) The uncertainties attached to all the outlines, and especially where breach is considered. 4) How/whether to include more frequent flooding.
How should the combined flood map be communicated? E.g. 1% flood depths, or probability of exceeding 0.5 m, others?
1% flood depths has the advantage of being what people are familiar with, but would not satisfy EA aim of predicting quantitative risk. The only problem with displaying probability of exceedance is comprehension, and setting multiple depths. Also what is the point of trying to integrate very rare events with catastrophic loss of life events and smearing very different things together? Despite aiming for a universal measure (probability*consequence), this does not always give the full sense of a problem.
What data are likely to become available in the next 5 years?
I think more Water Company data may become shared. The historic flooding database should become more widely used and updated. Improved Areas susceptible maps. UKCIP08 data and its interpretation.
Please feel free to add any other information you feel is relevant to this project.

Response 10
What sources of flooding need to be considered in delivering a single integrated map of flood hazard?
Fluvial, coastal, surface water (pluvial) including failure of pump systems, sewer exceedance and flash floods, groundwater, dam and canal breach All sources! Tsunami which was specifically mentioned in the tender is probably a step too far.
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
Useful to breakdown datasets according to source–pathway–receptor. All of the receptor datasets exist (at least for simple depth-based damage calculations). Pathway datasets are topography, NFCDD, national flood outlines and water surface datasets (e.g. JFLOW) etc. Sewer and drainage assets are the main problem. Until there is a quality controlled and widely available national sewer layer then all sources mapping is going to be a problem. Sources: problems with availability of local short time-step rainfall data. Also groundwater data. FEH flow data could do with updating. Wave data are costly. Surge data are ok. Then there is the question of whether future change in any/all of these quantities is included in the analysis.
There are two methods proposed for combining flood maps from different sources: 1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 – The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
The problem with the latter is assessing the joint probability of different sources. This will be very difficult to do on the basis of maps, as the sources of dependence can only be disentangled ‘upstream’. Approach 2 is easier, but I don’t think it will be taken seriously.
What will be the key technical challenges in delivering this project?

• Sorting out the dependence issues. • Integrating flooding processes that operate on different spatial/temporal scales. • Acquiring and dealing with the relevant datasets. • Pluvial flooding: representing the relevant processes and integrating over the spatial–temporal variability in some computationally feasible way. • Uncertainty. • Dealing with long-term change, e.g. climate change (Defra/EA are bound to ask for this soon if they haven't done already). • Computational architecture and flexibility of amending calculation components Transparency of calculation procedure.
How should the combined flood map be communicated? E.g. 1% flood depths, or probability of exceeding 0.5 m, others?
Both. Multiple maps to attribute risks to different sources and then demonstrate how they aggregate. Care must be taken in communication of the uncertainty associated with these probabilistic estimates.
What data are likely to become available in the next 5 years?
An interesting recent development has been the common metadataset for analysis of urban water/flooding issues in London, developed by MWH for GLA. Would be nice to see this sort of thing appearing for the UK… and even nicer if it was data rather than metadata. National rainfall-runoff modelling is more or less with us, but I expect it will become more physically based over the next 5 years. I hope that NFCDD will continue to improve.
Please feel free to add any other information you feel is relevant to this project.

Response 11
What sources of flooding need to be considered in delivering a single integrated map of flood hazard?
Rivers (main and non-main), sea, groundwater, lakes, reservoirs, canals, surface water, tsunami. See also a draft Floods Directive working group document, attached. Not all of these are relevant in England and Wales.
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
There are two methods proposed for combining flood maps from different sources: 1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 – The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
1. The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain. Probably better results in time; they certainly should be, and I think this is where we should be aiming long term; but I'm not sure that we're ready for it now; I suspect we'll lose the detail from some of the source-specific modelling that we currently do and that it will all be lost in a fog of uncertainty. 2. The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth). More feasible in the shorter term; also perhaps easier to break down the final result again (which is something that I think we need to do) so that we can say 'you've x chance of flooding; you're most likely to be flooded from source y; the uncertainty in source z is…' etc.
What will be the key technical challenges in delivering this project?
How should the combined flood map be communicated? E.g. 1% flood depths, or probability of exceeding 0.5 m, others?
Probably both! Priority goes to 'probability of exceeding 0.5 m'. It's a big change from what we do now, but it's more compatible with a probabilistic approach and can be modified to a simple 'what's my chance of flooding?'.
What data are likely to become available in the next 5 years?
Please feel free to add any other information you feel is relevant to this project.

Response 12
What sources of flooding need to be considered in delivering a single integrated map of flood hazard?
ALL – but in more detail: Fluvial, Pluvial, Coastal should be included. Combined probability and combined sources (i.e. flooding occurring at the same time) need to be modelled. Groundwater will be difficult to include – different timescale to other processes. Water supply infrastructure will be difficult to include – only information about subsurface water is currently available.
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
There is some information on combined pluvial/fluvial flooding in *An analysis of the combined consequences of pluvial and fluvial flooding* (conference paper supplied).
There are two methods proposed for combining flood maps from different sources: 1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 – The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
Combining flood maps is more straightforward, and allows different modelling methods to be used for different (non-interacting) sources. Combining sources might give better results (see conference paper).
What will be the key technical challenges in delivering this project?
The different levels of detail (spatial, probabilistic, depth) in each source/different models. For example it will be hard to combine models at different scales.
How should the combined flood map be communicated? E.g. 1% flood depths, or probability of exceeding 0.5 m, others?
Difficult to say – communicate to whom? Depends on the recipient of the information, from general public to expert. E.g. for the public, simple 1% flood depth probably enough; for professionals this will be too simple, they will require more information. Different things may need to be communicated: Hazard maps/risk maps/flood risk to people/damage.
What data are likely to become available in the next 5 years?
No new types. May be higher resolution modelling – but we're probably reaching the limit where further increases in resolution will produce little improvement in models. Resolution of rainfall (spatial, temporal) is likely to improve – this may change the types of outputs from pluvial models.

Response 13
What sources of flooding need to be considered in delivering a single integrated map of flood hazard?
Groundwater flooding in Chalk aquifers, and maybe alluvial aquifers. I identify Chalk and alluvial aquifers in particular because on these there is significant risk of groundwater flooding, and the driving processes are moderately well understood so that we may be able to quantify risk. Other aquifers tend not to generate groundwater floods so there is very limited data availability, and I don't believe that there is a driver to consider groundwater flooding on these.
What data are available for these sources? Are these associated with appropriate metadata? Are there IPR or security issues associated with these data?
Limited. Data tend to be informally stored at the Area/Region level. GW flood events are sporadic and not always identified as relating to GW, so records are inconsistent. MSfW HA5 made recommendations for data storage in FRIS, so maybe things are better – I haven't seen this database though.
There are two methods proposed for combining flood maps from different sources: 1 – The **sources** of flooding can be combined and used in a RASP type approach, adding the water volumes from the different sources and routing these over the floodplain 2 – The **flood maps** from different sources can be combined to give a single probabilistic map (e.g. 1% probability flood depth) Of these two methods, which do you think is: the more feasible; will give better results? Should any other options be considered?
I suggest that for groundwater flooding, 1 would best be used. This, however, will require modelling of the groundwater system under different recharge/rainfall conditions to predict flows and points of groundwater emergence. Useable models are available for most of the Chalk aquifer, but the effort may be considerable. Areas of susceptibility have been identified by various means but these have no consideration of risk/probability. These might be used as a surrogate for groundwater model results to show areas of groundwater emergence, but flows and return periods would have to be derived in another way. Maps of groundwater flooding are very limited, and have no associated probabilities.
What will be the key technical challenges in delivering this project?
Knowledge of groundwater flooding is far behind that of other types of flooding. Aquifer heterogeneity is very important in controlling locations and flow rates of groundwater emergence: this can only be understood at the catchment scale. Groundwater head response to rainfall is highly non-linear and depends on both long-term and short-term rainfall intensity; therefore quantifying risk may need a joint probability approach. There are no established methods for predicting groundwater flooding. I suggest trialling a variety of techniques on a few catchments where there are good data.
How should the combined flood map be communicated? E.g. 1% flood depths, or probability of exceeding 0.5 m, others?

No opinion on this.
What data are likely to become available in the next 5 years?
As systematic collection of groundwater flooding data has only been going for about 3–4 years, the dataset should double in size (assuming that it rains enough). I think that knowledge and modelling of groundwater flooding processes in the Chalk is going to advance considerably (e.g. by the FREE project).
Please feel free to add any other information you feel is relevant to this project.
High groundwater levels, while perhaps not causing groundwater emergence, can significantly increase the runoff coefficient of a catchment. Hence return periods for surface water flooding may be tied in with return periods of high groundwater levels. In this project it is important that surface water and groundwater flooding are considered together to avoid double counting of flows, and for surface water flood return periods to be tied in with those for high groundwater levels (for certain catchments).

Appendix C: Internal and external consultation list

Internal consultation list		Response included in Appendix B
Name	*Company/Organisation*	
Murray Dale (Pluvial)	Halcrow	Y
Elliot Gill (Surface and Sewer)	Halcrow	Y
Shaun Yeoh (Dam Breach)	Halcrow	Y
Bin Li (Tidal, Surge, Tsunami)	Halcrow	Y
Matthew Scott (NaFRA and MDSF2)	Halcrow	Y
Steve Buss (Groundwater)	ESI	Y
Garry Pender (All Sources)	Heriot Watt	Y
External consultation list		
Suresh Surendran	Environment Agency	
Adam Baylis	Environment Agency	Y
Stefan Laeger	Environment Agency	Y
Mike Steel	Environment Agency	
Kate Marks	Environment Agency	
Bill Donovan	Environment Agency	
Shirley Greenwood	Environment Agency	Y
Ian Andrews	Environment Agency	
Keith Beven	Lancaster University	
Jim Hall	Newcastle University	Y
Slobodan Djordjevic	Exeter University	Y
Barry Hankin	JBA	Y
Richard Dun	British Waterways	
Clare Savvas	Environment Agency	
Gary Tustin	Environment Agency	
Nathan Muggeridge	Mouchel	
David Balmforth	MWH	
Paul Shaffer	CIRIA	
Richard Kellagher	HR Wallingford	

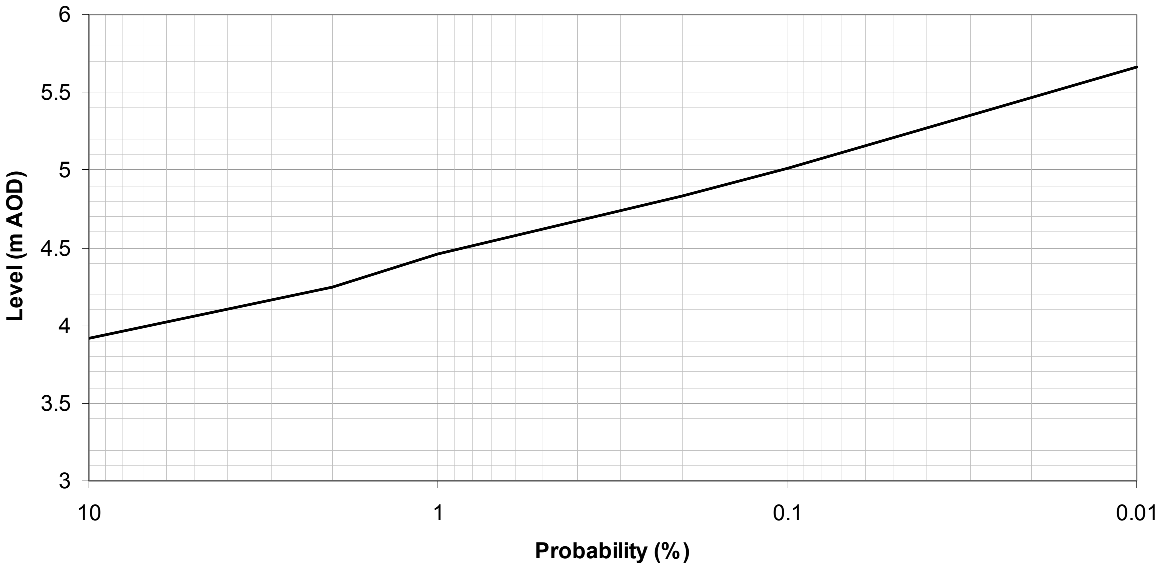
Appendix D: Terminology guide

A brief guide to probability, uncertainty and systems-based modelling

The Environment Agency's policy is to move towards probabilistic modelling to support the decisions it has to make to manage flood risk. The Environment Agency Science project 'Developing a Tool for Mapping Flooding from All Sources' is supporting this aim by taking a probabilistic approach to combining flood risk maps from many sources, and as part of this project we have developed a brief guide to the terminology associated with probabilistic and systems-based modelling and uncertainty. This guide will help both experts and non-experts to understand how these terms are used in the 'Mapping Flooding from All Sources' project.

Natural variability

Environmental systems respond to processes that are random, or so complex that they appear random. An example is a river catchment responding to rainfall patterns in time and space. We can't predict the future behaviour of such systems (for example when floods will occur in 2015), but we can describe their behaviour statistically using a probability distribution. For the river example, we might express this by saying that the probability of the flow exceeding 93 m^3/s in any one year is 0.1%. Probabilities such as these are used to describe river flows, tide/surge levels (see figure below) and rainfall.



Natural variability expressed as a water level varying with probability. The use of a single line indicates that for each return probability we have perfect knowledge of the water level, and for each water level we have perfect knowledge of the probability.

Knowledge uncertainty

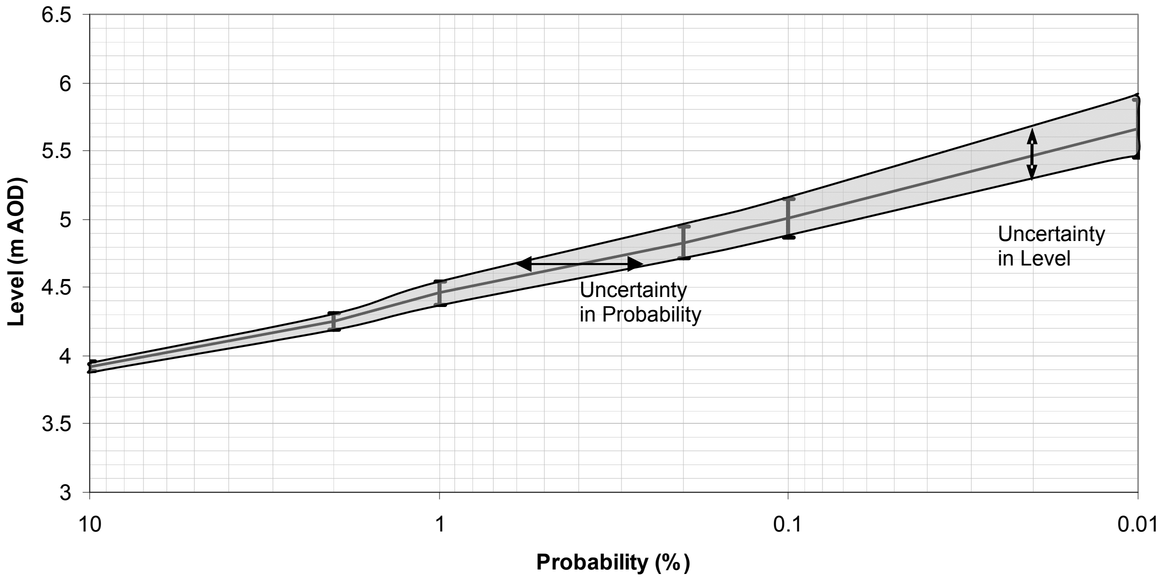
The probability distributions used to describe natural variability can be derived from series of measurements over a period of time (e.g. river flow gauge records), or from modelling (using our understanding of how the system behaves). Neither of these methods give a perfect understanding of the probability: our models are not perfect, and we only have a finite length of recorded measurements. This imperfect understanding of the natural processes, due either to limitations of science, the way we use the science in modelling, or lack of observations, can be expressed as knowledge uncertainty.

Knowledge uncertainty is equivalent to the concept of epistemic uncertainty, whereas natural variability corresponds to aleatory uncertainty.

This is illustrated in the figure below, where the probability/water level relationship is associated with an uncertainty, which expresses our incomplete knowledge of the probability distribution. This uncertainty can be expressed as either a range of water levels for each probability, or a range of probabilities for each water level. We assume that the true probability or water level lies somewhere in this range. This range can also be expressed as a probability distribution.

The natural variability and knowledge uncertainty can be combined into a single probability distribution. For example, a water level may be exceeded because a storm of a certain probability occurs, or because a smaller storm occurs but the true water levels are much higher than our model predicts. The probabilities for these occurrences can be combined mathematically, giving a single probability distribution.

It is however, useful sometimes to keep the natural variability and knowledge uncertainty separate. As we improve our models, or acquire more records over time, our knowledge uncertainty can decrease, so we gain a better idea of the probabilities. Natural variability, in contrast, can't be reduced in this way.



Natural variability and knowledge uncertainty expressed as a range of water levels varying with probability. For each probability we have imperfect knowledge of the water level, and for each water level we have imperfect knowledge of the probability.

Deterministic modelling

Deterministic modelling accounts for some of the natural variability of an environmental process, but not the knowledge uncertainties associated with our measurements or modelling. Typically we represent the natural variability at a small number of points, for example flows at 2–3 probabilities (say 1%, 0.5% and 0.1%).

This is the traditional approach to estimating flood risk. We can estimate, for example, the river flow corresponding to a 1% probability – this has a 1 in a 100 chance of being exceeded in any given year. This flow can then be converted to water levels and flood extents using a model or hydraulic calculations. The output is a single water level for each point in space, or a flood extent, corresponding to this single 1% probability event.

Probabilistic modelling

Probabilistic modelling aims to provide a better understanding of the probability distribution at the receptors of flooding. Probabilistic modelling combines both natural variability and knowledge uncertainty, outputting, for example, a map of probability of a depth threshold being exceeded. This is typically calculated using a Monte Carlo method, where a large number of model simulations are driven by randomly generated boundary conditions (e.g. representing natural variability) and model parameters (e.g. representing our knowledge uncertainty). This generates a large number of results, with the probability distribution derived from these.

Representing the combination of natural variability and knowledge uncertainty in this way gives more 'honest' model results, as it expresses our lack of knowledge of the environmental processes that generate flood risk. A drawback is that it is difficult to 'disentangle' natural variability and knowledge uncertainty from a single probability distribution. It is hard to see therefore how much improving the model can help in narrowing down the range of model outputs.

Systems-based thinking

The flood risk system is considered in terms of source–pathway–receptor–consequences components. The analysis should include all components of the system which are significant for the specific decision being supported. The system may change for different decisions and the modelling used to support them. An example is the river channel: for defence breaching models this may be regarded as a source; whereas for a dynamic flood mapping model the channel is regarded as a pathway.

Risk-based approach

Risk-based modelling uses the risk assessment process of problem formulation, risk assessment, option appraisal and risk management planning, targeting limited resources (time and money) to achieve maximum benefit (tangible and intangible). It can quantify and explicitly incorporate knowledge uncertainty and natural variability within a probabilistic analysis framework with risk defined as consequences multiplied by probability. This approach recognises that some elements of risk might arise from knowledge uncertainty (e.g. there is a probability that our flood models are wrong, and hence we design our defence crest levels too low), as well as natural variability (for any defence height, there is a probability they will be overtopped or fail). A risk-based approach can ensure we make the most robust decisions in the face of both types of uncertainty.

Hierarchical/tiered approach

The risk analysis method should be proportionate to the risk, decisions and spatial and temporal scale, while making best use of the available data. A hierarchical or tiered approach formalises which processes and/or process model detail are appropriate for application for different scales and types of decision making. For example, a broad-scale set of process descriptions may apply at a national level risk assessment, an intermediate level of detail at a catchment scale and more detailed process descriptions at a local level. For example, national-scale risk estimates can be used to inform funding levels, without requiring a detailed understanding of where the risk is located. Local modelling is, however, required to manage that risk at specific locations (e.g. flood defence design or warning).

A generally useful attribute of a hierarchical framework is to allow more accurate (local-scale) results to inform or replace the broader-scale assessment (where the local-scale data are available).

RASP approach

RASP (Risk Assessment for System Planning) describes a hierarchical, risk-based analysis framework. It is essentially a concept based on the following three core principles:

- **Systems-based thinking**. The flood risk system is considered in terms of source–pathway–receptor–consequences components. The analysis includes all components of the system which are significant for the specific decision being supported.
- **Risk-based approach.** Use of a standard risk assessment process of problem formulation, risk assessment, option appraisal and risk management planning, targeting limited resources (time and money) to achieve maximum benefit (tangible and intangible). Quantify and explicitly incorporate knowledge uncertainty and natural variability within a probabilistic analysis framework with risk defined as consequences multiplied by probability.
- **Hierarchical approach**. The risk analysis method is proportionate to the risk, decisions and spatial and temporal scale, while making best use of the available data.

Although RASP is not a software product or computer model, RASP-based methods have been implemented in a number of software tools such as MDSF2 and the computational engine used to generate NaFRA results. It should be noted that these current implementations do not represent all significant components of the flood risk system and are not fully probabilistic in that many sources of uncertainty are not incorporated. The RASP method currently used in NaFRA and MDSF2, for example, incorporates the natural uncertainty of loading water levels and defence failure states, but not knowledge uncertainty related to the input parameters and accuracy of sub-models used within RASP.

Final comments

In practice, methods and data rarely reside fully in just one of the above classifications. For example, analysis methods will combine probabilistic and deterministic elements. It is, perhaps, more useful to consider where an individual approach lies on a relative scale.

Also, it is important to use appropriate language when communicating methods and results with different audiences. For example, it is generally not beneficial to use terms such as probabilistic and deterministic when communicating with the general public, but these terms are useful during discussions between experts.

Appendix E: Question log

This appendix provides a record of questions that have been raised of the project during the lifetime of the project. This list is intended to help those involved in communication of follow-on projects and products.

- How will the tool and approach integrate with Surface Water Management Plans?
- Can the tool tell me how hazard/risk will change with climate change?
- How can the tool be used to help inform the design of interventions and to support decisions which will help to manage flood hazards (e.g. depth, velocity) and consequences?
- Can't NaFRA/RASP already do all this?
- How accurate are the results?
- At what scale should the results be used?
- How have you dealt with joint probabilities/dependence?
- How have you combined short duration (e.g. pluvial) and long duration (e.g. groundwater) flooding phenomena?
- Can we use this type of approach to combine defended and undefended datasets?
- How do we resolve the conflict between datasets with the same uncertainty/confidence?
- What are the best types of datasets to use in producing maps of all sources?
- How should we treat areas with very high uncertainty? How can we reliably communicate a 'don't know' answer?
- How can we use uncertainty estimates to direct future modelling to increase our confidence in the outputs?
- How can mapping from all sources include future datasets on dependency between sources?
- How will the method cope with climate change datasets?
- Can we include blockage and other asset failures as a source of risk? Can this be considered separately or does it need to be included in other datasets?

Appendix F: Groundwater flood risk assessment

1 Why is groundwater different (what are the potential challenges)?

Fluvial and coastal flooding are complex phenomena, affected by the interplay of numerous factors. However, these flood risk assessments are made more precise by the availability of detailed long-term datasets for key influencing factors. Such datasets are less commonly available for groundwater, analysis of which is also made more difficult by effects of heterogeneity between, and within, aquifers, which cannot be directly observed. The following comments expand upon these challenges.

a) Different aquifers respond in different ways to high rainfall, and their response depends on antecedent conditions, for example:

   i) Under normal conditions Chalk aquifers can absorb most incident rainfall and runoff typically does not occur. However, during or after a wet winter, when groundwater levels are already high, shallow subsurface pathways rapidly move rainwater through the system to cause flooding. This is followed by increased baseflow from the aquifer, which prolongs the flood event. Therefore the occurrence of groundwater flooding in Chalk catchments relies on an interaction between long-term and short-term rainfall amounts (e.g. Taylor 1995, Bradford and Croker 2007, Adams *et al*. 2008).

   ii) Groundwater levels in alluvial aquifers respond to changes in river level during fluvial flood events, as river water moves into the adjacent aquifer. Therefore the incidence and extent of the groundwater flood is controlled by the stage of the adjacent river as well as antecedent groundwater conditions (Macdonald *et al*. 2007).

b) In layered geological strata, aquifer properties are anisotropic and heterogeneous, sometimes leading to the creating of subsurface preferential flow paths. This leads to discharge occurring from discrete locations (i.e. springs), which may not be identifiable under conditions other than at times of flood (e.g. Finch *et al*. 2004). In addition, aquifer heterogeneity can make a considerable difference to the directions of groundwater flow and therefore the locations of discharge zones within a catchment. The nature and extent of superficial deposits over the aquifer will also influence discharge locations and quantities.

c) Groundwater level datasets are, more often than not, obtained at locations distant from any flooding (although in certain circumstances this can be seen as an advantage). Regional drawdown due to groundwater abstraction can ameliorate flooding, and mask its 'natural' frequency.

d) Groundwater floods may be masked by co-incident, more severe, fluvial flooding; therefore their occurrence may be unknown. Sometimes groundwater flooding only comes to light once defences from surface water flooding have been constructed. Or groundwater flooding may be a minor, initial part of an overall more serious fluvial flood event (e.g. Macdonald *et al*. 2007).

e) There is also an issue of acceptance of any flood maps developed by this project by the wider hydrogeological community. While there are often sufficient data to enable hydrologists to develop empirical flood models, hydrogeologists often find, for the reasons listed above, that there are not enough data. To generalise, therefore, hydrogeologists typically feel uncomfortable with empirical models and prefer calibrated physics-based models.

Groundwater flooding also arises from post-groundwater abstraction rebound of water tables and minewater emergence. These mechanisms are not directly controlled by antecedent conditions, but mostly by historical influences.

A search for international approaches to groundwater flood risk mapping has been undertaken, but there appear to be few. Most that were identified (US studies) were for specific areas, where the flooding occurs frequently and data are available. A more relevant example was identified in Germany but has not yet reported (http://www.hochwasser-dresden.de/MULTISURE).

2 Assessment of groundwater flooding

2.1 Flood mapping

Reliable recent and historical records of groundwater flooding events provide the best source of identifying areas susceptible to future groundwater flooding. However, due to under-reporting and uncertainty in identifying groundwater as the source of the flooding, these may not be reliable or thorough. Particularly with changing groundwater abstraction regimes, lack of evidence for groundwater flooding is not a good indicator of the likelihood of flooding in the future.

However, there are some reliable datasets (principally for Chalk catchments) that can be incorporated in to a flood mapping scheme. The main existing databases identified by Jacobs (2006) include the following:

Database	Ownership	Coverage
FRIS	Flood incident management	South West Region
NFCDD	Flood risk mapping and data management	National
Springs and Sources	Groundwater and contaminated land team	Thames Region
WISKI	Field monitoring and data	National

Jacobs (2006) recommended the inclusion of groundwater flooding incidents in the existing FRIS system. Since then, and particularly since the publication of the Pitt Review, the Environment Agency has been compiling groundwater flooding data. However, the outcome of the data collation has not yet been seen for this project. In some cases, the extent of flooding is indicated by line data (identifying flooded reaches) rather than flood envelopes.

It is unlikely that any of these data will be able to be directly used to derive a return period (except in catchments that experience groundwater floods frequently: in which case the flood mapping may already be part of the fluvial outline).

2.2 Screening assessments for hazard

Screening assessments that present groundwater flooding hazard have been widely produced. These combine available data on geology, groundwater levels and topography and provide regional overviews based on calculated depth to groundwater. However, they can only show where more detailed assessment may be warranted for local areas. As regional assessments, many of the data used in the assessments are sparsely distributed or poorly known – these include hydraulic conductivity (if used), the conceptual model and rest water levels. Methodologies used, or proposed, include the following:

a) JBA (2006) suggested that the average BFI HOST value of a given catchment may indicate its susceptibility to groundwater flooding, but this was not tested.

b) BGS GeoSure data includes a groundwater flooding susceptibility classification for the whole of the UK. Six classes range from 1: 'relatively high susceptibility' to 5: 'relatively low susceptibility', plus a class that indicates no susceptibility. There is no attempt to quantify risk in the dataset. Comparison of relative risk between aquifer types is not possible, as there is a classification of 1 in all aquifers, despite the Chalk of South East England being accepted as most prone to groundwater flooding. However, within a given aquifer type, the classification may be useful. Unlike many of the following methods, the BGS dataset includes predictions of areas potentially susceptible to alluvial groundwater flooding.

c) Jacobs (2004) developed a methodology for deriving groundwater emergence maps (GEMs) for major aquifers in England, which generates estimated groundwater contours for winter 2000/01, and compares these with the level of the ground surface to estimate areas of groundwater emergence. The method was published as Morris *et al.* (2007) so has undergone peer review, and GEMs are in use with some planning authorities to assess risk of groundwater flooding. Since groundwater contours are uplifted evenly across the catchment, this methodology is prone to excessive conservatism in river valleys, and will underestimate flood susceptibility on the interfluves.

d) JBA (2006) also developed a methodology for deriving groundwater emergence maps. While Jacobs (2004) used groundwater level contours, the method in JBA (2006) uses borehole data to assess in more detail the variation in groundwater head change across the catchment. This leads to lower groundwater level rises in river valleys and larger rises beneath interfluves, removing some of the potential error in the Jacobs (2004) method. Otherwise the two methods are similar. JBA has since mapped the whole Chalk aquifer for an indicative 1:100 year flood.

e) Jacobs (2008) used an assessment of high in-bank river levels to map areas of permeable superficial deposits adjacent to the Thames through London, which may be susceptible to groundwater flooding (unseen – reported in Cobby *et al.* 2009).

f) ESI (2009a) and ESI (2009b) use two further methods for assessing locations of groundwater emergence. ESI (2009a) uses a calibrated groundwater resources model of the Lower Mersey area to generate naturalised groundwater level contours (i.e. with no groundwater abstractions) to assess where groundwater rebound (due to cessation of abstraction) may cause surface inundation. (Note that these conditions are, however, outside the calibration range for the model.) ESI (2009b) generates groundwater flood susceptibility for an aquifer in north Yorkshire with sparse monitoring data using a spatially distributed analytical solution. Again, in each model study predicted heads are compared against the ground surface to estimate areas of groundwater emergence or enhanced runoff.

None of the methodologies identified relate the susceptibility of groundwater emergence to a return period or frequency, and therefore cannot be used directly to determine risk. However any of them might be combined with statistical methods (see Section 2.5) to develop a measure of risk.

It should also be noted that these methods all relate to the prediction of groundwater emergence – not groundwater flooding. For reliable prediction of groundwater flooding to be made, two further aspects need to be estimated: the volumetric rate of groundwater emergence and the effectiveness of surface drainage to remove that flow. (However, for the Chalk of southern England, the Jacobs [2004] GEMs show good correspondence between areas of groundwater emergence and known groundwater flooding.)

2.3 Integration of existing models

Existing models rarely effectively combine the simulation of both surface water and groundwater. Groundwater models are generally unsuitable for modelling flood events because they tend to be discretised coarsely over area and time: typically the Environment Agency's regional models have 200–250 m grid cells and are run with 10–30 day time-steps. There also tends to be no specific representation of the unsaturated zone, or of high near-surface hydraulic conductivity, both of which are critical in initiating floods in Chalk catchments (Adams *et al*. 2008). It is also questionable how well such models will predict extreme events if they have few real data to validate relevant simulations, but even using results outside of the range of calibration is better than making generic assumptions (as in many of the methods given in this section). Surface water models can be weak at representing groundwater emergence because the water can emerge as a diffuse seepage rather than a point discharge (spring), and they may not consider transient changes in runoff coefficient of the ground.

Off-the shelf solutions for integrated surface water–groundwater modelling are available, and there is no reason why they could not be used, given adequate resolution of the issues listed above. Some software potentially capable are ISIS (linked to an OpenMI compliant groundwater model or via the prototype ISIS-MODFLOW module) and the MIKE SHE suite. These appear to be largely untested for the prediction of groundwater flooding, however, and this topic is still very much in the research arena. GIS-based runoff-routing may provide a simplistic approach.

Imperial College, BGS and CEH are attempting to provide a new modelling tool to predict groundwater flooding in the Chalk of South East England (as part of the FREE project: http://www.groundwaterflooding.org.uk) by integrating existing models for surface and near-surface flow and groundwater flow. This will be used to investigate the use of simpler models for warning of the potential onset of flooding and regional assessment of risk. If successful, the approaches may be appropriate for catchment-scale investigation, but would be site-specific. The project is planned to deliver results in 2010, so will be too late for Phase 1 of this project.

Since groundwater flooding is essentially a one-way process of groundwater discharge to the surface, it may not be necessary to use fully coupled models. For example, a groundwater model (or GEM-like maps: see above) can be used to generate antecedent conditions (high water table and spring flows etc.) following seasonally high rainfall. Outputs would then be fed into a runoff-routing and surface water flood prediction tool, that takes into account spring flows and the increased runoff coefficient of the predicted saturated ground, to deal with high intensity–low duration rainfall and above-ground conditions that combine to cause flooding.

2.4 Flood frequency

While mapping approaches are available, the calculation of a return period may be the biggest challenge in this project, as there are no published approaches. Jacobs (2006) proposed the following five possible approaches which would enable estimated probabilities to be assigned to a groundwater flood:

a) *Assess the frequency of observed flooding.* Statistical methods (e.g. Najib 2008) can be used to predict flood return periods in similar ways to those established for river flooding. They depend, however, on the availability of medium to long-term datasets that cover extreme events. It is unlikely that these will be routinely available except for areas that are prone to frequent flooding by groundwater (e.g. Chichester [Taylor 1995]).

   For many identified groundwater floods on record there is one date only (i.e. winter 2000–01), and a return frequency could not be estimated. Newspaper archives would provide information on the frequency of groundwater flooding in areas where fluvial or surface water flooding does not occur, or where reports discuss the emergence of groundwater via springs or seeps, or the movement of spring lines (e.g. Taylor 1995).

b) *Assess the frequency from the drivers.* Rainfall datasets are of better quality and duration than any records on groundwater flooding. Statistical inferences might be identified from historical rainfall data. It is unlikely that a reliable method could be derived for prediction of groundwater flooding frequency based on a single rainfall variable because the relationship of groundwater level with rainfall is highly non-linear and poorly understood at present. It may be found to be valid to correlate river stage data with the incidence of groundwater flooding in alluvial aquifers.

   In low-storage, fissured aquifers such as the Chalk it has been observed that groundwater flooding requires firstly unusually high long-term rainfall (resulting in high groundwater levels) and then high short-term rainfall, which cannot be accommodated in the available storage (Adams *et al*. 2008). Therefore there is the opportunity to use joint probability techniques (Defra and Environment Agency 2005) to give an indication of the frequency of their simultaneous occurrence. Likewise, predictions of alluvial flooding might be refined by using river flow data and rainfall data series in a joint probability approach.

c) *Assess the frequency from surrogates.* Long-term datasets of groundwater level are available, although they often vary in quality and frequency. Likewise data on spring discharges and river flows are good. These surrogate datasets might be reliably used, when calibrated with known groundwater flooding events, as surrogate measures for analysing return periods (e.g. Bradford and Croker 2007). On a regional basis these might be appropriate, but there are likely to be insufficient monitoring locations to discriminate between individual catchments.

d) *Assess the frequency from a combination of drivers and surrogates.* Again, this is an opportunity to use the joint probability approach to refine estimates of flood frequencies by combining groundwater level or river flow with rainfall data. This suffers from the same limitations as considering each individually.

e) *Mathematical modelling.* As with flood mapping, groundwater models could be used to generate series of groundwater levels in order to create peaks over threshold time series. This approach is limited by the same issues as discussed above.

3. Existing datasets

Section 2.1 discusses the availability of groundwater flood mapping. These datasets, along with limited literature and newspaper archives, are probably the only sources of information on groundwater flooding. However, the following datasets might be of use for predicting groundwater flooding where no data are available..

a) Data on aquifer geology and superficial deposits are available from the British Geological Survey (BGS). All of England and most of Wales is covered by digital 1:50,000 mapping. They also hold a dataset of superficial deposit thickness, although the resolution of this is variable across the country. 3D geological models have also been developed of selected areas. A map of susceptibility to groundwater flooding has been produced for Great Britain. These datasets are all copyright NERC but many 3D models were paid for by the Environment Agency to feed into groundwater models.

b) Aquifer properties are for principal and secondary aquifers collated in Allen *et al*. (1997 and Jones *et al*. (2000). These are averaged on an aquifer-by-aquifer basis, rather than presented as distributed properties. Distributed aquifer properties are available for many principal aquifers in calibrated groundwater models, which are mostly held by Environment Agency regions, although some are in the possession of water companies.

c) Surrogate data: groundwater heads, spring flows and rainfall are all held by the Environment Agency.

d) Acquisition of GEMs would be valuable. These were generated by Jacobs (2004) but the copyright will probably be held by Defra.

e) JBA has proprietary mapping of 1:100 year groundwater flood events in the Chalk aquifer. Review of these would be valuable.

4. Summary and possible methodology

4.1 Summary

To summarise, the key points made above are as follows:

a) Knowledge and data regarding groundwater flooding is currently relatively behind that relating to most types of surface water flooding. This is partly because it had not become an issue, except locally, until the widespread groundwater floods of 2000/01. Also the processes are difficult to quantify due to aquifer heterogeneity and the complex non-linear relationship between rainfall and flooding. Therefore special consideration needs to be made when putting groundwater flooding into a framework developed mainly for aspects of surface water flooding.
b) Data on groundwater flood envelopes is very patchy and held informally by Environment Agency area and region offices.
c) Many methods have been proposed and used to develop maps of susceptibility to groundwater flooding, or areas of potential groundwater emergence. All have limitations, not least that there is no quantification of risk involved. Also, flood flows cannot be predicted by any of these methods.
d) Coupled modelling of groundwater and surface water for the prediction of groundwater flooding is still a research topic. Off-the-shelf models are available, but these suffer from limitations, both regarding the groundwater component and the surface water component.
e) There is no established methodology for deriving groundwater flood frequency, partly because of the lack of data. There are opportunities to use measures relating to the drivers (i.e. rainfall return periods) or surrogate measures (e.g. groundwater heads or spring flows).

4.2 Possible methodology

The following recommendations relating to the progress of this project are made, to enable this project to incorporate groundwater flood risk within the timescales required. Further work on mapping groundwater flood risk is recommended later.

a) As a demonstration, and to focus on real risk, the project should concentrate on mapping groundwater flood risk only from Chalk aquifers. It is on Chalk aquifers that most events are reported (see Table 2-A in Jacobs 2004) and damage to property occurs. Most existing data on groundwater flooding is related to Chalk aquifers, and most of the literature relates to groundwater flooding from Chalk aquifers. Therefore we can be confident to derive a robust, testable methodology for Chalk groundwater flooding.

b) Given that there is inadequate existing map data, it is recommended that the flood maps be developed using a modelling technique. It is envisaged that generation of flood envelopes would comprise two stages: firstly generate flows of emerging groundwater, then apply a RASP type approach to generate flood envelopes. Return periods would be developed by using surrogate time-series datasets within a joint probability approach.

c) The methodology should be trialled on a test catchment where there is sufficient data to validate it. The Kennet Valley is ideal for this purpose because: it comprises mostly Chalk outcrop, and it experiences regular groundwater flooding (e.g. in Newbury); there is an existing groundwater model, held by the Environment Agency, that is known to simulate extreme high groundwater levels well (Travis Kelly, Environment Agency, personal communication); and there is literature that uses the Kennet Valley, or adjacent catchments, as case studies (Finch *et al.* 2004, Bradford and Croker 2007, Morris *et al.* 2007, Adams *et al.* 2008). As an alternative, or in addition, the Lavant Catchment (including Chichester) would make a good case study for many of the same reasons.

d) Mapping of flood envelopes should use existing groundwater emergence maps (Jacobs 2004) to identify areas where, during a groundwater flood event, the water table is within 2 m of the ground surface. At these locations, the baseflow index can be assumed to be zero. A suitable 2D fluvial flood model can then be used to calculate the flow (100% of rainfall x area of GEM) and incorporate the effects of surface water drainage. A RASP method will be used to define the flood envelope for all types of flooding.

e) Return periods should be developed by using a surrogate measure with a joint probability approach. Suggested time series that can be used are: long-term (e.g. 90 day) and short-term (1 day) rainfall totals, or groundwater head and short-term rainfall total. The exact combination that could be used should be calibrated using the trial catchment(s). If the former combination is used, the period for totalling the long-term rainfall should be varied according to the catchment if possible, as this would incorporate some geological heterogeneity (i.e. length of flow path and storage effects).

f) It is important to involve an Environment Agency hydrogeologist on the review team for this project, as the alternative viewpoint will be of benefit to the overall product.

Once the methodology is tested and incorporated for Chalk aquifers, further refinements can be added. These should be prioritised according to perceived risk and complexity of the approach (i.e. timescale of implementation), or left for a later phase of mapping.

a) Develop catchment-specific GEMs for Chalk aquifers using calibrated groundwater models. This will allow explicit incorporation of aquifer heterogeneity into the system. Although the models might not be calibrated at extreme groundwater levels, this is still an improved approach to using generic GEMs.

b) Alluvial aquifers should be incorporated into the mapping. It is suggested that only areas of reported groundwater flooding be incorporated, as a broad-brush screening-type approach would produce large areas to deal with, without any gain. Simplistically, the flood envelope can be determined as those areas of the floodplain beneath the stage of the associated river at any given time (account needs to be taken of the peak duration, which will control how far the river water can intrude into the alluvial aquifer), and return periods would be linked to the re-occurrence of that river stage.

   Note that most of the areas that will be identified as at risk from groundwater flooding from alluvium will also be part of a fluvial flood envelope, although the frequency of the groundwater flood will be higher than that of the fluvial flood. Note also that, to take proper account of flood defences, their penetration of the alluvial aquifer must be ascertained as groundwater floods in alluvial aquifers can propagate beneath flood defences.

c) If there is a driver to do so (i.e. a significant number of reported groundwater flooding events), other fractured limestone aquifers can be incorporated into the mapping, using the same approach as Chalk aquifers. Karstic limestone aquifers (e.g. Carboniferous Limestone) will demonstrate different groundwater flooding processes, so should not be incorporated at this stage.

d) Other principal aquifers could be mapped in due course, if there are a significant number of reported groundwater flooding incidents arising. Flood processes will be different from those in the Chalk, so an amended methodology would need to be developed.

Appendix F references

Adams, B., Bloomfield, J., Gallagher, A., Jackson, C., Rutter, H. and Williams, A., 2008. *FLOOD1 Final Report.* BGS Open Report OR/08/055.

Allen, D.J., Brewerton, L.J., Coleby, L.M., Gibbs, B.R., Lewis, M.A., MacDonald, A.M., Wagstaff, S.J. and Williams, A.T., 1997. *The Physical Properties of Major Aquifers in England and Wales.* British Geological Survey Technical Report WD/97/24. Environment Agency R&D Publication 8.

BGS (British Geological Survey), 2007. *Susceptibility to Groundwater Flooding – User Guide*.

Bradford, R.B. and Croker, K.M., 2007. Application of head-flow responses to groundwater floods in Chalk catchments. *Quarterly Journal of Engineering Geology and Hydrogeology*, 40, 67–74.

Cobby, D., Morris, S., Parkes, A. and Robinson, V., 2009. Groundwater flood risk management: advances towards meeting the requirements of the EU Floods Directive. *Journal of Flood Risk Management*, doi: 10.1111/j.1753-318X.2009.01025.x.

Defra and Environment Agency, 2005. *Use of Joint Probability Methods in Flood Management. A Guide to Best Practice*. R&D Technical Report FD2308/TR2.

ESI, 2009a. *Lower Mersey and North Merseyside Water Resources Study: Final Report*. Report reference 6588R4.

ESI, 2009b. *Mapping Groundwater Flooding between Helmsley and Scarborough*. Report reference 60002R2.

Finch, J.W., Bradford, R.B. and Hudson, J.A., 2004. The spatial distribution of groundwater flooding in a chalk catchment in southern England. *Hydrological Processes*, 18, 959–971.

Finch, J., Marsh, T. and McKenzie, A., 2007. *A Preliminary Risk Assessment of the Potential for Groundwater Flooding During the Winter of 2007/8 – An Update*. Centre for Ecology and Hydrology.

Jacobs, 2004. *Strategy for Flood and Coastal Erosion Risk Management: Groundwater Flooding Scoping Study (LDS 23)*.

Jacobs, 2006. *Making Space for Water. Groundwater flooding records collation, monitoring and risk assessment (reference HA5). Extended Report (Chalk aquifers)*.

Jacobs, 2008. *Thames Estuary 2100: Other Sources of Flooding and Effects of Interventions*. Second revision.

JBA, 2006. *Making Space for Water. Flooding from other sources* (HA4a).

Jones, H.K., Morris, B.L., Cheney, C.S., Brewerton, L.J., Merrin, P.D., Lewis, M.A., MacDonald, A.M., Coleby, L.M., Talbot, J.C., McKenzie, A.A., Bird, M.J., Cunningham,

J. and Robinson, V.K., 2000. *The Physical Properties of Minor Aquifers in England and Wales*. British Geological Survey Hydrogeology Group Technical Report WD/00/4, Environment Agency R&D Publication 68.

Macdonald, D., Hall, R., Carden, D., Dixon, A., Cheetham, M., Cornick, S. and Clegg, M., 2007. Investigating the interdependencies between surface and groundwater in the Oxford area to help predict the timing and location of groundwater flooding and to optimise flood mitigation measures. *Paper to Defra 42nd Flood and Coastal Management Conference*, York.

Morris, S.E., Cobby, D. and Parkes, A., 2007. Towards groundwater flood risk mapping. *Quarterly Journal of Engineering Geology and Hydrogeology*, 40, 203–211.

Najib, K., Jourde, H. and Pistre, S., 2008. A methodology for extreme groundwater surge predetermination in carbonate aquifers: groundwater flood frequency analysis. *Journal of Hydrology*, 352, 1–15.

Taylor, S.M., 1995. The Chichester Flood, January 1994. *Paper to the British Hydrological Society 5th Symposium*.

Appendix G: Proceedings of stakeholders' workshop

A workshop was organised by the project team with the aim of:

- raising awareness of the project among specialists and potential end users;
- facilitating the exchange of ideas on mapping flooding from all sources.

The workshop took place on 14 July 2009 and the participants reflect a wide spectrum of interested parties.

Name	Organisation
Adam Bayliss	Environment Agency
John Blanksby	Sheffield University
Steve Buss	ESI
Sally Daniels	Environment Agency
Slobodan Djordjevic	Exeter University
Richard Dun	British Waterways
Shirley Greenwood	Environment Agency
Liana Hamilton-King	Environment Agency
Matt Horritt	Halcrow
Rahman Khatibi	Halcrow
Stefan Laeger	Environment Agency
Luke Lovell	Halcrow
Gareth Pender	Heriot Watt
Guest of G. Pender	Heriot Watt
Mervyn Pettifor	Flood Management Support Services
John Ray	Environment Agency
Syd Simpson	Bradford Council
Mike Steel	Environment Agency
Gary Tustin	Environment Agency
Jim Walker	Environment Agency
Simon Waller	JBA
Alan Warren	Halcrow
Jon Wicks	Halcrow
Brian Wilkinson	UKWIR
Mike Wood	Torbay Council

The format of the meeting comprised three sets of presentations, each followed by discussions in the morning session. In the afternoon sessions, two breakout workshops were run, and both reported back an overview of their discussions. Discussions were constructive and there seemed to be an overall endorsement of the emerging thinking in the projects. This appendix presents a summary of each presentation followed by a record of the discussions, and a record of the breakout workshops. Further comments received after the workshop are summarised at the end of this appendix.

Presentation 1 by Shirley Greenwood, Environment Agency, project executive

Shirley's high level presentation covered the following topics.

Why do we need to start considering Flood Risk from All Sources – drivers?

- Defra Making Space for Water; Future Water
- PPS25 – encouraging LAs to consider all sources
- Pitt Review – range of mapped risks increased
- Floods Directive – All sources mapped by Dec 2013

How will we consider Flood Risk from All Sources of Flooding?

- Combine sources for complete picture of flooding?
- Decisions based on the whole system approach?
- One map?

Key challenges

- Consider joint probability
- Show combined risk in a probabilistic format
- Avoid a 'cumulative approach' that leaves 'the whole country coloured blue'
- Current R&D Project: Produce Prototype Methods for Mapping All Sources of Risk (this is a first step)

Presentation 2 by Jim Walker, **Environment Agency, project manager**

Jim's presentation covered the following topics:

What we have now Fluvial and Tidal Risk: the Flood Map

- 10 years since first published to the web: £Millions additional investment to refine Flood Map via 'Section 105' and Strategic Flood Risk Mapping Programme
- Resulting in a comprehensive picture of risk from these main sources
- Use: Strategic Planning
- Leading Europe in development of such National Flood Maps

What we have now: NaFRA (National Flood Risk Assessment)

- Flood Map Zones 2 and 3 = Undefended risk
- Fluvial and tidal risk including the influence of defences
- Needed for catchment scale flood risk management (CFMPs): Capital investment and maintenance
- Needed for probabilistic consideration of risk – insurance industry

List of Datasets Being Considered

- Fluvial
- Tidal
- Estuarine

- Surface Water
- Reservoir Inundation
- Groundwater
- Canal Breach
- Failure of Pumps and Barriers
- Tsunami

Consideration of scale, primarily national datasets – but method flexible enough to incorporate local data in future.

Presentation 3 by Matt Horritt, Halcrow

After scene setting by Shirley and Jim, and before embarking on the technical details of the project, Matt defined terminology appropriate to the project to ensure a common understanding – particularly for deterministic and probabilistic modelling and their inherent uncertainty and variability. The first session was complemented by discussion and generally there was an overall agreement but different opinions were also expressed, especially on whether to regard MDSF and NaFRA as fully probabilistic. The discussion covered the following main points:

1. **The problem of communicating risk in probabilistic terms**

   There are lots of causes of uncertainty/lack of confidence and this may affect communicating risk in probabilistic terms (SL); differing views expressed were: bookmakers already use probabilistic terms (AB) [and many people are comfortable with such language]; if probabilistic approaches are hard to communicate, so are deterministic approaches, are deterministic results properly understood? (SG); by tailoring the communication of risk to spectrum of needs, the difficulties in communication can be reduced (MP); when communicating risk, complexity can be avoided by not using such terms as deterministic or probabilistic (SG); we must focus on the important sources of uncertainty; a communication of flood risk must be suitable for triggering appropriate action (i.e. keep end user needs in mind (MP)).

2. **The measure of probability**

   There is also an issue with how to report probabilistic measures of risk or just use central tendency values (SB); another issue is whether one should pass on uncertainties associated with probabilistic measures (AB); probability is important but more important is the depth or the probability associated with depth (MP); communicating risk in terms of whether a property will get wet or will remain dry (JR); natural frequencies can be used, e.g. 1 in 1000, as used in medical practices for a long time, but also probabilistic ratio of 0.001 would be understood, others say that decimal probabilities are hard to understand (e.g. RD); it is better to use probability values (GP); conventional practices for probability measures are either in terms of return period or in terms of probability of annual exceedance with the Environment Agency promoting the latter (SG).

3. **Risk perception**

   An opinion was expressed on people's perception of flood risk. One opinion emphasised depth of flooding and argued that homeowners would want to know the depth of flooding that their properties would be exposed to (MP); another opinion was that homeowners would be interested only whether their properties will get wet or will remain dry (JR).

4. **End users**

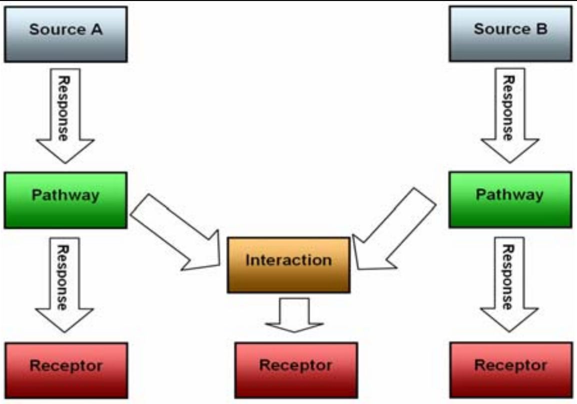
The difficulties in communicating flood risk in probabilistic terms can be better addressed by focusing on different end users (e.g. a spectrum of professional users and laymen) and tailoring communication messages to trigger appropriate actions (MP); when appropriate actions are triggered, communication is regarded to be fit for purpose (MP).

5. **Omission**

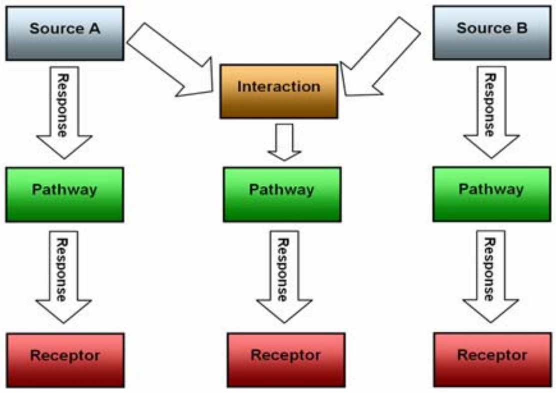
   It was pointed out that the definitions did not cover spatial coherence in addressing the likelihood associated with particular events, e.g. the summer 2007 event (SW).

Technical Presentation 4 by Matt Horritt, Halcrow

This presentation introduced the broad method options available as summarised by the following figure.

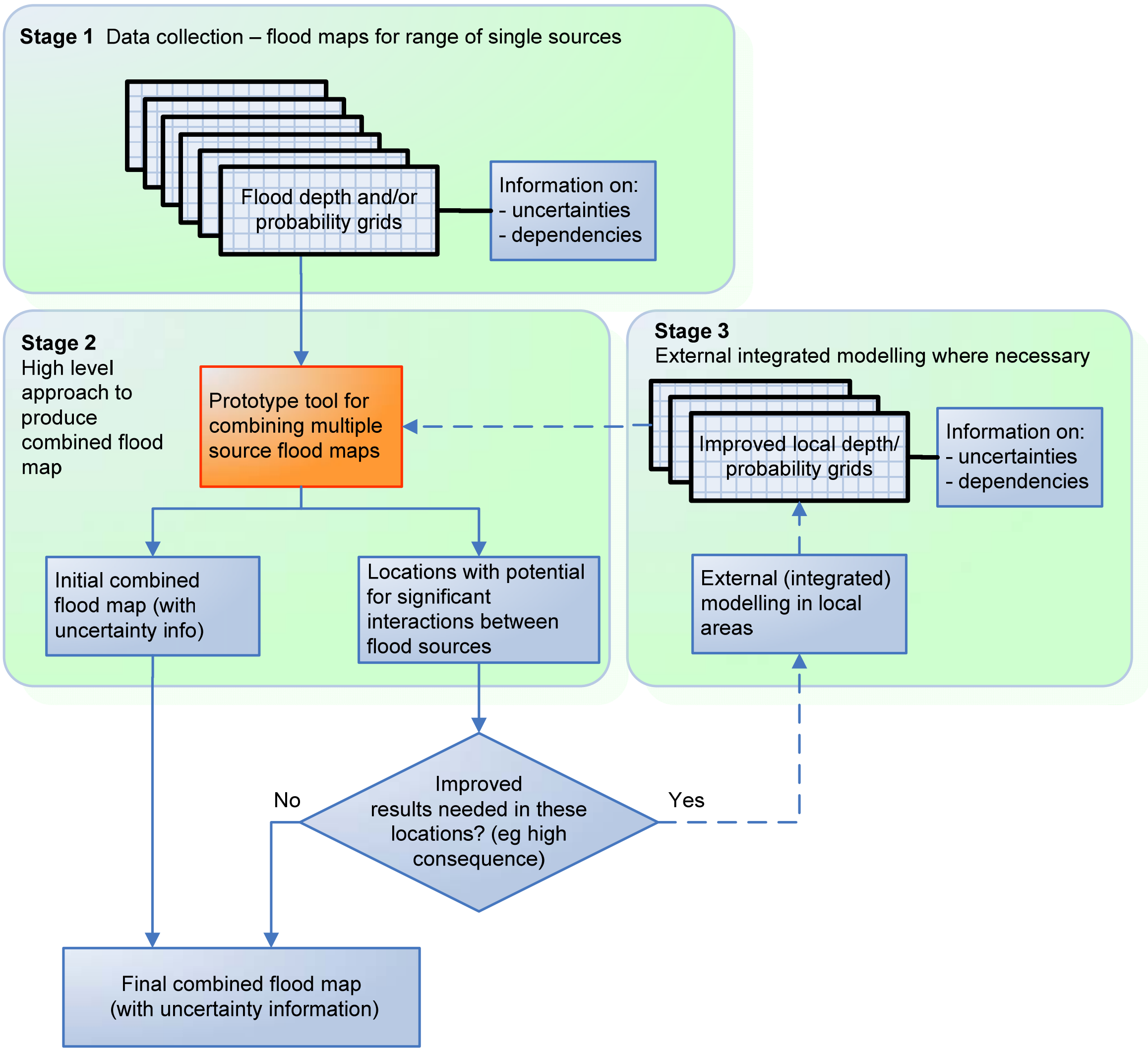


Option A – Integration <u>after</u> routing along pathways (map combination approach)



Option B – A fully integrated approach

The workshop attendees were in general in agreement with the currently selected approach which uses option A, with option B used locally where necessary. (Although it was pointed out that this is a Science project and we should, where appropriate, propose new and potentially complex new methods.) This 'risk-based' approach is summarised in the following figure.



The presentation included examples derived from the above procedure. The subsequent discussions are summarised below:

1. **Consequential events**

   Certain sources, such as reservoir dam failure, are difficult to assign a probability of recurrence to and are therefore unlikely to provide meaningful probability depth grids (RD). However, an Environment Agency Science project will be investigating this issue (GT). The use of non-quantitative likelihood bands was discussed (e.g. low, medium and high) – could these be used in the method and could they be mixed or displayed along with quantitative probabilities? (RD).

2. **Output data resolutions**

   Mapping associated with each source has an appropriate resolution and therefore mapping from all sources needs to resolve the problem of differing resolutions (SD). Resampling to a common resolution (probably small grid size) can be a way to resolve the problem (MH).

3. **Integrity of results assessing the risk from all sources**

   In a hierarchy of decisions [towards mapping from all sources], there are data-rich areas and those with sparse data and therefore the example of Hull used in the presentation (which has detailed data coming from the relevant SFRAs) may not paint the full picture (MP); the interactions between different sources-pathways are likely but the Option A method in the presentation may not capture these interactions and therefore the outcome of mapping from all sources should be communicated with a ‘health warning’ by specifying the uses – we want to try to ensure the results are not miss-used (JB); modelled results need to be scrutinised to ensure that they are fit-for-purpose (SG); constraining to existing data can lead to underestimation and possibly ignoring interactions (SW); we should not ‘lock out’ the use of local data (SG).

4. **Scope**

   Should we rule out other measures of hazard, such as velocity and duration, speed of inundation (MP); there is a need to include infrastructure reliability in the mapping (BW), (MH suggested this can be done as an additional layer); there are benefits to green infrastructure to identify areas flooded frequently [hotspots] and areas that do not flood – i.e. to identify areas which could be flooded more often (MP).

5. **Knowledge management**

   We should learn from the process of checking and roll-out of NaFRA. Think about what the Environment Agency will do with the map, how it will be checked, who will have access, how will it be updated etc.

6. **Additional emerging issues**

   There are flood mapping datasets under SFRA and new sets may be compiled by Internal Drainage Boards (IDBs); also there is an initiative to collect all the historical flood data.

Breakout Group Sessions

Two breakout group sessions were held:

- Session 1: Methods and Technical Issues.
- Session 2: Datasets and User Requirements; this session involved a presentation on the compiled information on categorising mapping datasets for the various sources of flood risk, which are covered in Chapter 7 of this report and a copy was distributed to the participants seeking their views.

Following the group gatherings, the discussions were consolidated and are given below.

Main feedbacks from Session 1

- **‘Don’t know’ is good** – we need to be honest about situations where our modelling doesn’t tell us much about flood risk.
- **Source interaction and uncertainty** – it may be that uncertainty swamps any signal from source interaction, so that the interaction can be ignored.
- **Depth ratio** – when combining two sources can we screen them by looking at the ratio depth, and ignoring the smaller? This will be equivalent to taking the maximum except when they are approximately the same, when interaction may be significant.
- **Copula approach** – using the joint probability information at source level (e.g. river flows and tide/surge) is probably OK for water levels on the floodplain too.
- **Interaction screening** – one possible approach is to look at where hazard maps overlap, e.g. surface water flooding in fluvial zone 2 or 3.
- **Interaction areas** – these can also be used to prioritise future modelling or reuse of old models.

Main feedbacks from Session 2

- **Main uses of mapping of flooding risk from all sources**
  - Primary uses: facilitate the compliance with the Floods Directive, support strategic planning and investment planning activities, and public awareness raising.
  - Secondary uses: pre-event flood warning planning, future generations of CFMPs and SMPs, emergency or contingency planning for real-time flood risk management, critical infrastructure strategies.
  - Other uses: insurance, performance indicators, SFRAs.
- **Issues and challenges** (biggest issues first)
  - Cost, time, resources to implement local checking/improvements
  - Quality of input data and input data coverage
  - A programme of updating the maps
  - Security can be an issue, e.g. flooding by reservoir breaches
  - The ownership of the maps: the Environment Agency or another organisation (avoiding duplication) – where will they reside (Environment Agency website?)
  - How will the maps be used in the planning system?
  - Producing a common national standard including legends, formats, attributes
  - Grid sizes and resolutions
  - Capability to produce maps for a diversity of flood variables, e.g. depth, duration, velocity
  - Producing maps appropriate to the public (household) and professional uses – there may be different levels of detail for these two user groups
  - There may also be technology issues (e.g. local authorities tend to use MapInfo rather than ArcView)
  - The relationship between the emerging new maps from all sources and the existing maps from (generally) one source – a test on whether the new maps are appropriate for a specific end use will be whether the new maps actually improve decision making for that use.

Post-workshop comments

Key points of post-workshop comments by Gary Tustin, Environment Agency

Gary confirmed some of the details on Reservoir Inundation Mapping (RIM) project commissioned by the Environment Agency. It is carried out on a national scale in England and Wales. Phase 2 involves more detailed mapping exercise. All together, this will involve mapping flood risk from 600 dam failures of the highest risk reservoirs. The timetable for this work is to have both phases complete by 31 December 2009.

Before the Draft Floods and Water Management Bill becomes law, the Environment Agency is likely to extend the RIM programme to extend the register to reservoirs with a minimum capacity of 10,000 m^3.

The remit of the RIM project does not include formulating a procedure for probability of annual exceedance of the risk associated with reservoir breaches. However, a rule of thumb approach can be formulated based on the median probability of dambreak failure and this will be defined shortly by the project with some supporting evidence.

Refining the uncertainty requires a more detailed risk assessment exercise and this could be achieved by either (a) undertaking individual risk assessments, which is likely to be very costly or (b) undertaking a desk exercise against available data to assign probability.

Currently strict security issues are associated with RIMs and no mapping can be made publicly available which brings together RIMs and other sources. Security also extends to draft reports.

Key points of post-workshop comments by Brian Wilkinson (UKWIR)

There are uncertainties at the many stages of flood risk assessment, i.e. in rainfall data, pipe and surface flow models, catchment details. Mapping of flood risk areas should include only those areas where confidence in the assessment of flood risk is above a minimum threshold, say, >50%. Otherwise the information based on the maps from all sources will not be useful to users and decision makers.

A single mapping system to meet the needs of all users seems somewhat ambitious. Professional users such as the Environment Agency, planners, developers, insurers, utilities etc. are able to interpret and make use of complex risk information whereas members of the public need straightforward guidance otherwise they will misinterpret the information or not use it at all. Perhaps selective access to an overall mapping system would be capable of meeting the varied needs of users.

Mapping to date has focused mainly on river catchments which change relatively slowly. As we seek to move similar principles of mapping into the peri-urban and urban areas, where changes of catchments and infrastructures are more frequent, this will throw up more issues for the validity of flood mapping such as:

- Can a mapping system represent adequately the localised nature of flooding?
- In modelling to assess risk, what assumptions are to be made about the availability, reliability and standard of maintenance of flow paths and infrastructure during storm events? Blockage of flow paths or failure of plant at critical times substantially changes localised risk.
- In a regularly changing urban environment, how will risk assessments and mapping information be kept up to date? Whose responsibility will this be? If the maps become outdated, the mapping facility will be discredited and fall into disuse.

Key points of post-workshop comments by Mike Steel, Environment Agency

The importance of the management of underlying datasets. Mike concurs with the vision of the project of producing raster database of probability from various sources (probability of depth, velocity and duration – even speed of onset where available). After establishing this unique and consistent base dataset, the maps can be interrogated in terms of exceedance, uncertainty, joint probability and dependence, temporal change, audit, and combined with a National Receptor Database to evaluate Risk.

Mike points out additional benefits associated with one clear and consistent national (but also going down to local detail) understanding of likelihood from all sources. This would make it feasible to consider (and audit) changes over time.

It would also allow presentation/communication to be tailored to different stakeholders; rather than one map fitting all, the one map that users see could be different depending on whether they were a member of the public, an insurer, local authority spatial or emergency planner, or LRF.

Time horizon and spatial extent of storms/precipitation. Mike comments that storms have different characteristics in terms of local intensity (e.g. quick storms or flash floods), longer and catchment-wide events and seasonal periods of high rainfall regionally or nationally which produce groundwater flooding. The benefits of such a wide focus are that events are no longer pigeon-holed into narrow categories, but this might allow for a different approach to assessing what types of input are critical for a particular location.

Appendix H: Application of FD2121 guidance

H1 Introduction

This appendix contains a copy of the FD2121 'software architecture notes' forms completed as far as possible for the Prototype tool for mapping flooding from all sources.

H2 General Project Information

Project Ref	SC080050
Project Title	Prototype tool for mapping flooding from all sources
Contact name	Jon Wicks
Company	Halcrow Group Limited
Tel	01793 812479
Email	wicksjm@halcrow.com
Target Audience	Environment Agency and the consultants engaged upon tasks on behalf of the Environment Agency
No. of Environment Agency users (approx) if applicable	40 Environment Agency users 20 Consultant users
Project overview	
Halcrow has been commissioned by the Environment Agency to develop a prototype tool for mapping flooding from all sources. The purpose of the tool is to combine information on flooding from different sources (e.g. river, coastal, surface water etc.) into a single map communicating the probability of flooding. See Science Report SC080050/SR1 for more information.	
Architectural Diagram	
N/A	
Other relevant information	
There are non-Environment Agency users to consider and the full spectrum of users has not been finalised.	

H3 Environment Agency Software Platforms

Develop software to run harmoniously on existing Environment Agency systems (hardware/network/software) and non-Environment Agency systems to maximise user acceptance.

Software developed to be run on Environment Agency systems needs to run on the platforms the Environment Agency already uses (or will have at the time of delivery) and can support (for a synopsis of Environment Agency platforms, correct at the time of writing see the latest, 'Enterprise Architecture: Technical Reference Model').		
Response		Rationale
Will run Will not run N/A	[X] [ ] [ ]	Software will be designed to run on current Environment Agency platforms.
Software to be run on non-Environment Agency machines should be written to maximise uptake of the software by these third parties, i.e. write the software to run on the most commonly used platforms.		
Response		Rationale
Implemented Not implemented N/A	[X] [ ] [ ]	We will write the software to best fit with the target audience (primarily Windows 2000/XP).
Software to be run on both Environment Agency machines and non-Environment Agency machines should marry the requirements of the two in the best way possible. This issue must be discussed with CIS.		
Response		Rationale
Plan in place Not considered Agreed with CIS N/A	[X] [ ] [ ] [ ]	To be agreed with CIS. Our proposed solution is designed to maximise take-up of the software by non-Environment Agency entities. Other sections of the document explain some of the steps taken to achieve this.

H4 Hardware Platforms

Software developed to be run on Environment Agency machines must be developed to run on existing hardware platforms at the Environment Agency.		
Response		Rationale
Will run Will not run N/A	[X] [ ] [ ]	Will be designed from the beginning to do so.
Any network bandwidth usage by the software must be communicated to CIS, including details of average and burst activity (see Appendix H for further information).		
Response		Rationale
Communicated Not communicated N/A	[ ] [X] [ ]	Full details not known at this stage. It is expected, however, that most files will be stored on the local machine and so network bandwidth is not expected to be a major issue.
(Server) Processor usage should also be communicated to CIS (see Appendix D.4 for further information).		
Response		Rationale
Communicated Not communicated N/A	[ ] [ ] [X]	The currently proposed solution involves running the application entirely on desktop PCs and as such server processor usage is not an issue.

The need for any peripherals will need to be agreed with CIS.

Response		Rationale
Agreed with CIS No agreement N/A	[ ] [ ] [X]	None required.

H5 Database Usage

Develop 'Enterprise' database based software to run on the Environment Agency standard database. Develop 'Desktop database' software in a way that doesn't require client installs and is not locked to a proprietary format.

Database based solutions must run on the standard Environment Agency database (currently Oracle) if the program is to be run at the Environment Agency. Databases other than the standard enterprise database will not be allowed onto Environment Agency systems.

Response		Rationale
Will run Will not run N/A	[ ] [ ] [X]	There is no database component to the proposed solution.

If the developed software is to be run at both Environment Agency and non-Environment Agency sites then if possible develop for the standard Environment Agency database. If this is not possible or will harm uptake by the non-Environment Agency users then write database agnostic software which will run on both the standard Environment Agency databases and those in use by the non-Environment Agency entities (a recommended approach in general).

Response		Rationale
E. Agency standard DB agnostic Other N/A	[ ] [ ] [ ] [X]	

Where software is to be developed for use at the Environment Agency and 'enterprise' databases are inappropriate for the task, desktop/embedded databases may be required. In these cases native access from within the application would be required, with no application or client installs on Environment Agency desktop PCs. It is also required that output to a non-proprietary format (e.g. XML) is easily available from the database.

Response		Rationale
Will comply Will not comply N/A	[ ] [ ] [X]	

H6 Non-Database Data

Do not create new proprietary data formats; store ancillary data, such as program settings, using XML file formats (see Appendix D.1 for more details).

Software developed should not write to/read from its **own** proprietary format; in general XML should be used for **new** formats. The only justification for creating proprietary formats in extreme cases might be due to performance issues, but this would have to be agreed with CIS beforehand. Where binary formats are proposed the Environment Agency would expect to receive documentation as to the format of these and also expect some ability to handle/produce XML input/output. It is acceptable to use the de facto 'standard' file formats that the Environment Agency itself uses for things such as GIS systems.

Response		Rationale
No proprietary formats	[X]	We will adopt XML as our standard format for any new file

Proprietary formats N/A	[ ] [ ]	format requirements identified. Industry standard native formats (such as ESRI GIS formats) will be used where necessary.

Where **ancillary** data (program settings etc.) is required to be stored you are expected to use XML as the format for this data.

Response		Rationale
XML used XML NOT used N/A	[X] [ ] [ ]	Ancillary data such as program settings will be stored using XML.

H7 Application Architectural Compliance

Develop applications using an n-tier, server side logic, thin client browser based approach, wherever this can satisfy the project requirements

New software developments to run on Environment Agency machines should follow the Environment Agency standard application architecture – an n-tier approach, utilising a 'business logic' server side in conjunction with a browser based thin client. Where this approach cannot satisfy the project requirements you will need to agree an alternate strategy with CIS, strong justification will be required (see Appendix D.6 for one possible alternative – Citrix).

Response		Rationale
Will comply Will not comply N/A	[ ] [X] [ ]	Unfortunately, it is anticipated that these standards would harm the project. An n-tier browser based approach is not planned, because: • There is no database access layer to the application • GIS functionality in web browser interfaces is typically less responsive, giving a poorer user experience, as well as being more difficult to develop • Installation issues – following Environment Agency CIS guidelines would involve users needing to install a web/application server, either onto servers in their organisation or their own desktop PCs. This is a barrier to use of the program, particularly for users external to the Environment Agency.

Software developed to run both at non-Environment Agency sites and on Environment Agency machines should follow the above Environment Agency application architecture wherever possible. If this is not practical for non-Environment Agency entities then a dual interface approach (using the same basic code base) is preferred, e.g. a rich client application at non-Environment Agency sites and standard Environment Agency application architecture for Environment Agency machines.

Response		Rationale
E. Agency standard Dual interface Other N/A	[ ] [ ] [X] [ ]	This has been considered, but cannot be justified primarily due to cost issues and time constraints. The development will be done so as to separate the user interface from the processing code, making a dual interface technically feasible as a possible future development. However, as stated earlier, the browser based approach is unlikely to provide the level of functionality and responsiveness in terms of GIS functionality.

H8 Development Tools & Languages

Write software using Environment Agency standard development tools.

Ideally, all new development should take place in the standard development language (currently Java). Any deviation from this requires justification. If the software is not to run on Environment Agency machines then justification will be easier.		
Response		Rationale
Will comply Will not comply N/A	[X] [ ] [ ]	While the choice of development language has not been decided definitively at this stage – it is anticipated that Java will be chosen assuming it can satisfy the requirements of the tool. This will then comply with Environment Agency CIS standards.
Where the exact CIS standards cannot be met, you should provide justification. Note that an architecturally compliant solution (i.e. thin client browser based) is preferable to a strict adherence to specific tools.		
Response		Rationale
Standards fully met Architectural compliance Other N/A	[X] [ ] [ ] [ ]	
Where the Environment Agency standard development language cannot satisfy the project requirements, for example modelling applications, then the use of a different language could be justified – the Environment Agency standard for modelling applications is currently C++.		
Response		Rationale
E. Agency standard Non-standard N/A	[X] [ ] [ ]	

H9 Modular, Sustainable Development

Develop modular, easily extensible and reusable software to obtain maximum value from the Environment Agency's investment.

Contractors should develop their software in as modular a fashion as possible, using loosely coupled functions/methods probably via Object Oriented development.		
Response		Rationale
Done/will do Not appropriate Not considered N/A	[X] [ ] [ ] [ ]	The software will be designed in this way.
Develop software so that user interfaces are decoupled from program logic as much as possible.		
Response		Rationale
Done/will do Not appropriate Not considered N/A	[X] [ ] [ ] [ ]	This will be an integral part of the overall architecture of the program and would allow for the future migration to a n-tier application architecture in the future if required.
The use of design patterns and other modern programming techniques should be considered. Use techniques such as inheritance and encapsulation appropriately and to their best advantage.		
Response		Rationale
Done/will do Not appropriate Not considered N/A	[X] [ ] [ ] [ ]	Design patterns where applicable will be used (in particular the 'Gang of Four' patterns) as well as general object oriented techniques.

Developed software may contain functionality that itself will be useful for reuse in other software perhaps by another contractor or the Environment Agency itself. You should make this as easy to achieve as possible and should endeavour to make it possible regardless of development environment.		
Response		Rationale
Done/will do Not appropriate Not considered N/A	[X] [ ] [ ] [ ]	This will be partially achieved by the separation of application logic from the user interface. It is expected that the good practices being employed will naturally lead to this, but development will be carried out with this in mind.
Achieve maximum interoperability by following the CIS standards along with various methods such as creating 'wrappers' around software, separating code into libraries/componentisation, open communication and data exchange via SOAP and XML.		
Response		Rationale
Done/will do Not appropriate Not considered N/A	[X] [ ] [ ] [ ]	Most of this is not applicable to the development of the tool. However, if there is a suitable opportunity to develop standalone libraries of code, this will be done.
External interfaces and available functionality should be clearly documented.		
Response		Rationale
Done/will do Not appropriate Not considered N/A	[X] [ ] [ ] [ ]	This should be done as a matter of course and internal code will be documented using a tool such as JavaDoc.

You should consider the use of coding standards and provide details and/or references to these.
We will use Halcrow coding standards.

H10 Security (User & Data)

Only implement application level security where absolutely necessary.

Comments
There will be no data storage requirements pertinent to the Data Protection Act. It is unlikely that security will need to be implemented beyond restricted access to associated network storage locations to Environment Agency users. Access to the system will be via existing rights management systems, i.e. we will not create our own – access will be secured using standard Windows security. Some data may have licensing restrictions and the Environment Agency may need to grant licences to use its data to external consultants. In addition it there may be security issues in relation to reservoir data, in particular location of reservoirs.

H11 Testing and Acceptance

Plan testing from the beginning of the project and follow the Environment Agency testing model.

Comments
It is currently planned to follow the testing guidelines fully.

H12 Implementation Planning

Follow the Environment Agency deployment procedures, determine who is responsible for support and enable a smooth implementation with no 'nasty surprises'.

H13 Software Deployment

For software to be installed on Environment Agency machines, the contractor must follow the Environment Agency standards for software deployment/install. Suppliers of proposed systems must provide documented support for the application integration task to the standardised Environment Agency desktop.

Comments
We currently plan to create an install program to be run on user machines. For instance, assuming Java as the development language, this will install the program, Java Runtime Engine (if required) and any example data. The program could also be provided in the form of a folder copy type of operation

H14 Transition to Support and Maintenance

Who will support the application?		
Response		Rationale
Environment Agency CIS Contractor Other Not yet known	[ ] [ ] [ ] [X]	TBC (but most likely first line support will be Environment Agency CIS)

A2-4.3 Storage Requirements

Storage requirements	
Frequency of backup required	Weekly (Backup should be to Environment Agency standards for project related data).
Recovery time	Same day (Recovery should be to Environment Agency standards for project related data).
Amount of storage required (now)	Unknown, very dependent on no. of users and datasets used. It is expected that data be stored locally in which case storage requirements will be low.
Amount of storage required (future)	The number of datasets available is likely to increase, so the storage requirements are likely to increase. Again this will depend on where Environment Agency users are to store data.

Appendix J: Pseudocode

The following pseudocode (actually python scripts) will be further commented for the final version.

```
def prob_from_depth_single2(p1_lb,p1_ub,d1,p2_lb,p2_ub,d2_lb,d2_ub,depths):
# Interpolate to common set of depths
# Second data set is a single point with lb/ub for both probability and depth
# e.g. dam break data, surface water data

    import numpy

    p1_int_lb=numpy.zeros(len(depths))
    p1_int_ub=numpy.zeros(len(depths))
    p2_int_lb=numpy.zeros(len(depths))
    p2_int_ub=numpy.zeros(len(depths))

    for i, d in enumerate(depths):
        (p1_int_lb[i],p1_int_ub[i])=interpolate_bounds_log(d1,p1_lb,p1_ub,depths[i])

        if d<d2_lb:
            p2_int_lb[i]=p2_lb
            p2_int_ub[i]=p2_ub
        elif d<d2_ub:
            p2_int_lb[i]=0.
            p2_int_ub[i]=p2_ub
        else:
            p2_int_lb[i]=0.
            p2_int_ub[i]=0.

    p_comb_lb=p1_int_lb+p2_int_lb-p1_int_lb*p2_int_lb
    p_comb_ub=p1_int_ub+p2_int_ub-p1_int_ub*p2_int_ub

    return p_comb_lb, p_comb_ub


def prob_from_depth(p1_lb,p1_ub,d1,p2_lb,p2_ub,d2,depths):
# Interpolate to common set of depths
    import numpy
    import math
    import array

    p1_int_lb=numpy.zeros(len(depths))
    p1_int_ub=numpy.zeros(len(depths))
    p2_int_lb=numpy.zeros(len(depths))
    p2_int_ub=numpy.zeros(len(depths))

    for i, d in enumerate(depths):
                    (p1_int_lb[i],p1_int_ub[i])=interpolate_bounds_log(d1,p1_lb,p1_ub,depths[i])
                    (p2_int_lb[i],p2_int_ub[i])=interpolate_bounds_log(d2,p2_lb,p2_ub,depths[i])
    p_comb_lb=p1_int_lb+p2_int_lb-p1_int_lb*p2_int_lb
    p_comb_ub=p1_int_ub+p2_int_ub-p1_int_ub*p2_int_ub

    return p_comb_lb, p_comb_ub


# perform look up using log interpolation for the P variable
# input is arrays of depth and probabilities (or vice versa)
# dependent variable is given as upper and lower bounds
# assume monotonically increasing function

def interpolate_bounds_log(xarg,yarg_lb,yarg_ub,xi):
    import numpy
    import math
    import array

# Sort if >1 element
```

```
    if len(xarg)>1:
                    index=xarg.argsort()
                    x=xarg[index]
                    y_lb=yarg_lb[index]
                    y_ub=yarg_ub[index]
    else:
                    x=xarg
                    y_lb=yarg_lb
                    y_ub=yarg_ub

# Extrapolation below
    if xi<x[0]:
                                    ub=1.
                                    lb=y_lb[0]
                                    return lb, ub

# Extrapolation above
    if xi>x[-1]:        # -1 is last element (index wrapping)
                                    lb=0.
                                    ub=y_ub[-1]
                                    return lb, ub

# Spot on single value
    if xi==x[0]:
        lb=y_lb[0]
        ub=y_ub[0]
        return lb, ub

# Interpolation
    for j, xj in enumerate(x):
        if xj >= x[j] and xj <= x[j+1]:
                    grad=(numpy.log(y_lb[j+1])-numpy.log(y_lb[j]))/(x[j+1]-x[j])
                    lb=numpy.exp(numpy.log(y_lb[j])+(xi-x[j])*grad)
                    grad=(numpy.log(y_ub[j+1])-numpy.log(y_ub[j]))/(x[j+1]-x[j])
                    ub=numpy.exp(numpy.log(y_ub[j])+(xi-x[j])*grad)

                    return lb, ub

# If we get here, somethings wrong
                    print("Error in interpolate - stopping.")
                    sys.exit()

def grids():
    import numpy
    import ascii_raster
    import math
    import CombineIndP

    (tide_200,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_200.asc")
    (tide_1000,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_1000.asc")
    (fluvial_100,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\f_nd_rp100.asc")
    (fluvial_200,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\f_nd_rp200.asc")


    p1_tide=1./200
    p2_tide=1/1000.
    p1_tide_error=0.001
    p2_tide_error=0.0002

    p1_fluvial=1./100
    p2_fluvial=1/200.
    p1_fluvial_error=0.001
    p2_fluvial_error=0.0007

    NaN=-9999.

    depth_thresh=numpy.array([0.3,1.0,2.0])
    nd=len(depth_thresh)

    p_out_ub=numpy.zeros((ni,nj,nd))
```

```
    p_out_lb=numpy.zeros((ni,nj,nd))
    p_out_cep=numpy.zeros((ni,nj,nd))
    p_out_crv=numpy.zeros((ni,nj,nd))
    p_out_err=numpy.zeros((ni,nj,nd))
    p_out_ent=numpy.zeros((ni,nj,nd))

#math.ceil(max(tide_200.max(),tide_1000.max(),surface_200.max()))
    maxd=3.0
    step=0.1

    d_out=numpy.array(range(0,maxd/step+1,1))*step

    for i, col  in enumerate(tide_200):
    for j, h in enumerate(col):

            p1_ub=0
            p1_lb=0
            p2_ub=0
            p2_lb=0

# Cases for different combinations of non-zero data. We can have no data, one value
# or both.
# Both values
            if tide_200[i,j]!=NaN and tide_1000[i,j]!=NaN:
                p1_ub=numpy.array([p1_tide+p1_tide_error,p2_tide+p2_tide_error])
                p1_lb=numpy.array([p1_tide-p1_tide_error,p2_tide-p2_tide_error])
                d1=numpy.array([tide_200[i,j],tide_1000[i,j]])
# One value only. Below this depth, the probability must be between the values
# associated with the two grids. We fudge this by inserting a data point below
# the p2 value to ensure we don;t interpolate linearly (see report section XXXX)
            elif tide_1000[i,j]!= NaN:
                p1_ub=numpy.array([p1_tide+p1_tide_error,p1_tide+p1_tide_error,p2_tide+p2_tide_error])
                p1_lb=numpy.array([p2_tide-p2_tide_error,p2_tide-p2_tide_error,p2_tide-p2_tide_error])
                d1=numpy.array([0,tide_1000[i,j]-0.001,tide_1000[i,j]])
            else:
# No data at all - we can still say the probability of depth being greater than
# zero is less than the probability of the most extreme event grid
                p1_ub=numpy.array([p2_tide+p2_tide_error])
                p1_lb=numpy.array([0.])
                d1=numpy.array([0.])

# Cases for different combinations of non-zero data. We can have no data, one value
# or both.
# Both values
            if fluvial_100[i,j]!=NaN and fluvial_200[i,j]!=NaN:
                        p2_ub=numpy.array([p1_fluvial+p1_fluvial_error,p2_fluvial+p2_fluvial_error])
                        p2_lb=numpy.array([p1_fluvial-p1_fluvial_error,p2_fluvial-p2_fluvial_error])
                        d2=numpy.array([fluvial_100[i,j],fluvial_200[i,j]])
# One value only. Below this depth, the probability must be between the values
# associated with the two grids. We fudge this by inserting a data point below
# the p2 value to ensure we don't extrapolate incorrectly (see report section
# XXXX for plots and a description of how this works)
            elif fluvial_200[i,j]!= NaN:
                p2_ub=numpy.array([p1_fluvial+p1_fluvial_error,p1_fluvial+p1_fluvial_error,p2_fluvial+p2_fluvial_error])
                p2_lb=numpy.array([p2_fluvial-p2_fluvial_error,p2_fluvial-p2_fluvial_error,p2_fluvial-p2_fluvial_error])
                d2=numpy.array([0,fluvial_200[i,j]-0.001,fluvial_200[i,j]])
# No data at all - we can still say the probability of depth being greater than
# zero is less than the probability of the most extreme event grid
            else:
                p2_ub=numpy.array([p2_fluvial+p2_fluvial_error])
                p2_lb=numpy.array([0.])
                d2=numpy.array([0.])

            p_comb_lb, p_comb_ub =CombineIndP.prob_from_depth(p1_lb,p1_ub,d1,p2_lb,p2_ub,d2,d_out)


            for k,d in enumerate(depth_thresh):
                        p_out_lb[i,j,k]=p_comb_lb[list(d_out-depth_thresh[k]).index(min(abs(d_out-depth_thresh[k])))]
                        p_out_ub[i,j,k]=p_comb_ub[list(d_out-depth_thresh[k]).index(min(abs(d_out-depth_thresh[k])))]
```

```
        print "Processed column ", i, " out of ", ni

    p_out_cep=0.5*(p_out_lb+p_out_ub)
    p_out_crv=numpy.log(1./p_out_cep-1)

    p_out_err=100.*0.5*(p_out_ub-p_out_lb)/p_out_cep
    p_out_ent=-numpy.log2(p_out_cep)*p_out_cep-numpy.log2(1.-p_out_cep)*(1.-p_out_cep)

    for i,d in enumerate(depth_thresh):
                    file_tail="_%3.1f.asc"%d

ascii_raster.write_ascii(p_out_lb[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_fluvial_p_lb"+file_tail)
ascii_raster.write_ascii(p_out_ub[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_fluvial_p_ub"+file_tail)
ascii_raster.write_ascii(p_out_cep[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-FMAS\TestCases\Hull\\tide_fluvial_p_cep"+file_tail)
ascii_raster.write_ascii(p_out_err[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050- FMAS\Test Cases\Hull\\tide_fluvial_p_err"+file_tail)
ascii_raster.write_ascii(p_out_ent[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050- FMAS\Test Cases\Hull\\tide_fluvial_p_ent"+file_tail)
ascii_raster.write_ascii(p_out_crv[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050- FMAS\Test Cases\Hull\\tide_fluvial_p_crv"+file_tail)


# Ming: Update points to use the same calculation method as in grids above


def points():
(tide_200,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_200.asc")
(tide_1000,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_1000.asc")
(surface_200,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\s_rp200.asc")

                    p1_tide=1./200
                    p2_tide=1/1000.
                    p1_tide_error=0.001
                    p2_tide_error=0.0002

                    p1_surface=1/200.
                    p1_surface_error=0.002

                    NaN=-9999.

                    maxd=3.0
                    step=0.1

                    d_out=numpy.array(range(0,maxd/step+1,1))*step

test_points=numpy.array([[505901.076251,430023.993597],[512545.290454,430675.995925],[508431.466240,433361.624563],[5
11023.951688,433610.006403]])

                    test_points_ij=numpy.zeros((4,2),dtype=int)
                    for i, tp in enumerate(test_points):
                                    test_points_ij[i,0]=int((tp[0]-xll)/dx)
                                    test_points_ij[i,1]=int((tp[1]-yll)/dx)

                    point_no=1

                    for i,j in test_points_ij:

                    if tide_200[i,j]!=NaN and tide_1000[i,j]!=NaN:

                                    p1_ub=numpy.array([p1_tide+p1_tide_error,p2_tide+p2_tide_error])
                                    p1_lb=numpy.array([p1_tide-p1_tide_error,p2_tide-p2_tide_error])
                                    d1=numpy.array([tide_200[i,j],tide_1000[i,j]])
                    elif tide_200[i,j]!= NaN:
                                    p1_ub=numpy.array([p1_tide+p1_tide_error])
                                    p1_lb=numpy.array([p1_tide-p1_tide_error])
                                    d1=numpy.array([tide_200[i,j]])
                    else:
                                    p1_ub=numpy.array([p2_tide+p2_tide_error])
                                    p1_lb=numpy.array([0.])
                                    d1=numpy.array([0.])

                    if surface_200[i,j]!=NaN:
```

```
                p2_ub=numpy.array([p1_surface+p1_surface_error])
                p2_lb=numpy.array([p1_surface-p1_surface_error])
                d2=numpy.array([surface_200[i,j]])
        else:
                p2_ub=numpy.array([p1_surface+p1_surface_error])
                p2_lb=numpy.array([0.])
                d2=numpy.array([0.])
                p_comb_lb, p_comb_ub = \

        CombineIndP.prob_from_depth(p1_lb,p1_ub,d1,p2_lb,p2_ub,d2,d_out)

        print "Point ", point_no, i, j
        print "Surface 200:", surface_200[i,j]
        print "Tidal 200:", tide_200[i,j]
        print "Tidal 1000:", tide_1000[i,j]

        for k, d in enumerate(d_out):
                p_out_cep=0.5*(p_comb_lb[k]+p_comb_ub[k])
                p_out_crv=(1./p_out_cep-1)
                p_out_err=100.*0.5*(p_comb_ub[k]-p_comb_lb[k])/p_out_cep
                p_out_ent=-numpy.log2(p_out_cep)*p_out_cep-numpy.log2(1.-p_out_cep)*(1.-p_out_cep)

                print "%8.6f %8.6f %5.2f %8.6f %8.6f %5.3f"%(p_comb_lb[k], p_comb_ub[k], d, p_out_cep ,
p_out_err, p_out_ent)
                point_no=point_no+1

                print "All done."

 def grids(chi):

 import numpy
 import ascii_raster
 import math
 import CombineIndP

 nodata_value=-9999.0

 (tide_200,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_200.asc")
 (tide_1000,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_1000.asc")
 (fluvial_100,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\f_nd_rp100.asc")
 (fluvial_200,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\f_nd_rp200.asc")

 fluvial_100[numpy.nonzero(fluvial_100<0.01)]=nodata_value
 fluvial_200[numpy.nonzero(fluvial_200<0.01)]=nodata_value

 nz=numpy.nonzero(fluvial_200>=0.01)
 fluvial_200[nz]=fluvial_200[nz]+0.1

 p1_tide=1./200
 p2_tide=1/1000.
 p1_tide_error=0.001
 p2_tide_error=0.0002

 p1_fluvial=1./100
 p2_fluvial=1/200.
 p1_fluvial_error=0.001
 p2_fluvial_error=0.0007

 depth_thresh=numpy.array([0.3,1.0,2.0])
 nd=len(depth_thresh)

 p_out_ub=numpy.zeros((ni,nj,nd))
 p_out_lb=numpy.zeros((ni,nj,nd))
 p_out_cep=numpy.zeros((ni,nj,nd))
 p_out_crv=numpy.zeros((ni,nj,nd))
 p_out_err=numpy.zeros((ni,nj,nd))
 p_out_ent=numpy.zeros((ni,nj,nd))
 p_JP=numpy.zeros((ni,nj,nd))
 p_JP_class=numpy.zeros((ni,nj,nd),dtype='int')
```

```
    maxd=3.0
    step=0.1

    d_out=depth_thresh

    for i, col  in enumerate(tide_200):
    for j, h in enumerate(col):

          p1_ub=0
          p1_lb=0
          p2_ub=0
          p2_lb=0
          p1_be=0
          p2_be=0

          JP_zero_flag1=0
          JP_zero_flag2=0

# Cases for different combinations of non-zero data. We can have no data, one value
# or both. As well as upper and lower bounds, we need to process "best estimate"
# values (somewhere between the lower and upper bounds) to calculate joint
# probabilities

# SOURCE #1

# Both values
          if tide_200[i,j]!=nodata_value and tide_1000[i,j]!=nodata_value:
             p1_ub=numpy.array([p1_tide+p1_tide_error,p2_tide+p2_tide_error])
             p1_lb=numpy.array([p1_tide-p1_tide_error,p2_tide-p2_tide_error])
             p1_be=numpy.array([p1_tide,p2_tide])
             d1=numpy.array([tide_200[i,j],tide_1000[i,j]])

             JP_zero_flag1=1

# One value only. Below this depth, the probability must be between the values
# associated with the two grids. We fudge this by inserting a data point below
# the p2 value to ensure we don;t interpolate linearly (see report section XXXX)
          elif tide_1000[i,j]!= nodata_value:
             p1_ub=numpy.array([p1_tide+p1_tide_error,p1_tide+p1_tide_error,p2_tide+p2_tide_error])
             p1_lb=numpy.array([p2_tide-p2_tide_error,p2_tide-p2_tide_error,p2_tide-p2_tide_error])
# Use flattest growth curve as "best estimate" - this will tend to overestimate
# effects of dependency
             p1_be=numpy.array([p1_tide+p1_tide_error+1e-6,p1_tide+p1_tide_error,p2_tide])
             d1=numpy.array([0.,tide_1000[i,j]-0.001,tide_1000[i,j]])
          else:
# No data at all - we can still say the probability of depth being greater than
# zero is less than the probability of the most extreme event grid
             p1_ub=numpy.array([p2_tide+p2_tide_error,p2_tide+p2_tide_error])
             p1_lb=numpy.array([1e-6,1e-6])   # Small +ve value - will take log later!
             p1_be=0.5*(p1_ub+p1_lb)    # Use CEP as best estimate
             p1_be[0]=p1_be[0]+1e-6
             d1=numpy.array([0.,1.])

# SOURCE #2

# Both values
          if fluvial_100[i,j]!=nodata_value and fluvial_200[i,j]!=nodata_value:
                   p2_ub=numpy.array([p1_fluvial+p1_fluvial_error,p2_fluvial+p2_fluvial_error])
                   p2_lb=numpy.array([p1_fluvial-p1_fluvial_error,p2_fluvial-p2_fluvial_error])
             p2_be=numpy.array([p1_fluvial,p2_fluvial])
                   d2=numpy.array([fluvial_100[i,j],fluvial_200[i,j]])

             JP_zero_flag2=1
# One value only. Below this depth, the probability must be between the values
# associated with the two grids. We fudge this by inserting a data point below
# the p2 value to ensure we don't extrapolate incorrectly (see report section
# XXXX for plots and a description of how this works)
          elif fluvial_200[i,j]!= nodata_value:
             p2_ub=numpy.array([p1_fluvial+p1_fluvial_error,p1_fluvial+p1_fluvial_error,p2_fluvial+p2_fluvial_error])
             p2_lb=numpy.array([p2_fluvial-p2_fluvial_error,p2_fluvial-p2_fluvial_error,p2_fluvial-p2_fluvial_error])
```

```
# Use flattest growth curve as "best estimate" - this will tend to overestimate
# effects of dependency
            p2_be=numpy.array([p1_fluvial+p1_fluvial_error+1e-6,p1_fluvial+p1_fluvial_error,p2_fluvial])
            d2=numpy.array([0.,fluvial_200[i,j]-0.001,fluvial_200[i,j]])
# No data at all - we can still say the probability of depth being greater than
# zero is less than the probability of the most extreme event grid
         else:
            p2_ub=numpy.array([p2_fluvial+p2_fluvial_error,p2_fluvial+p2_fluvial_error])
            p2_lb=numpy.array([1e-6,1e-6])
            p2_be=0.5*(p2_ub+p2_lb)    # Use CEP as best estimate
            p2_be[0]=p2_be[0]+1e-6
            d2=numpy.array([0.,1.])

         p_comb_lb, p_comb_ub =CombineIndP.prob_from_depth(p1_lb,p1_ub,d1,p2_lb,p2_ub,d2,d_out)

# Store results to upper/lower bound probability grids
         for k,d in enumerate(depth_thresh):
            p_out_lb[i,j,k]=p_comb_lb[list(d_out-d).index(min(abs(d_out-d)))]
            p_out_ub[i,j,k]=p_comb_ub[list(d_out-d).index(min(abs(d_out-d)))]

            p_out_cep[i,j,k]=0.5*(p_out_lb[i,j,k]+p_out_ub[i,j,k])

# Can only calculate joint probability if both grids are non-zero for both
# sources
            if JP_zero_flag1 and JP_zero_flag2:
               for k,d in enumerate(depth_thresh):
                  p_JP[i,j,k],p_JP_class[i,j,k]=JPchi(p1_be,d1,p2_be,d2,d,chi)
            else:
               p_JP[i,j,k]=-9999.
               p_JP_class[i,j,k]=-9999

      print "Processed column ", i, " out of ", ni

   p_out_crv=numpy.log(1./p_out_cep-1)
   p_out_err=100.*0.5*(p_out_ub-p_out_lb)/p_out_cep
   p_out_ent=-numpy.log2(p_out_cep)*p_out_cep-numpy.log2(1.-p_out_cep)*(1.-p_out_cep)

for i,d in enumerate(depth_thresh):
 file_tail="_%3.1f.asc"%d

ascii_raster.write_ascii(p_out_lb[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_fluvial_p_lb"+file_tail)
ascii_raster.write_ascii(p_out_ub[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_fluvial_p_ub"+file_tail)
ascii_raster.write_ascii(p_out_cep[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-FMAS\TestCases\Hull\\tide_fluvial_p_cep"+file_tail)
ascii_raster.write_ascii(p_out_err[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_fluvial_p_err"+file_tail)
ascii_raster.write_ascii(p_out_ent[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-FMAS\Test Cases\Hull\\tide_fluvial_p_ent"+file_tail)
ascii_raster.write_ascii(p_out_crv[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050- FMAS\Test Cases\Hull\\tide_fluvial_p_crv"+file_tail)
ascii_raster.write_ascii(p_JP[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_fluvial_p_JP"+file_tail)
ascii_raster.write_ascii(p_JP_class[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-
FMAS\TestCases\Hull\\tide_fluvial_p_JPclass"+file_tail)


def JPchi(p1,d1,p2,d2,d_out,chi):
# Use look up table to determine for a given value of
# depth X what value of uniform variate u gives X=x1(u)+x2(u)
   import numpy
   import Interpolation

# Choose set of probabilities to include all those given, plus points where
# the two curves reach zero depth
   P_int=numpy.concatenate((p1,p2),axis=0)

# Use log log interpolation works better at small depths - this is quite important
   P_zero1=numpy.exp(-numpy.exp(Interpolation.interpolate(numpy.array(d1),numpy.log(-numpy.log(p1)),0)))
   P_zero2=numpy.exp(-numpy.exp(Interpolation.interpolate(numpy.array(d2),numpy.log(-numpy.log(p2)),0)))

   P_zero1=min([P_zero1,1])
   P_zero2=min([P_zero2,1])
   P_zero3=0.5*(P_zero1+P_zero2)  # This point sometimes helps in interpolation

   P_int=numpy.append(P_int,[P_zero1,P_zero2,P_zero3])
```

```
# Remove duplicates and sort
    tmp=[]
    for v in P_int:
if not v in tmp: tmp.append(v)

    P_int=numpy.array(tmp)
    P_int=P_int[P_int.argsort()]

#    print "P_int: ", P_int

# Interpolate depths for the two sources at the P_int values, ensure these are
# >0. Again use log log interpolation
    x1_int=Interpolation.int_prob(-numpy.log(p1),numpy.array(d1),-numpy.log(P_int))
    x2_int=Interpolation.int_prob(-numpy.log(p2),numpy.array(d2),-numpy.log(P_int))

     x1_int=numpy.maximum(x1_int,0)
    x2_int=numpy.maximum(x2_int,0)

    x1x2=x1_int+x2_int

    alpha=numpy.log(2-max(min(chi,0.99),1e-6))/numpy.log(2)

# Find uniform variate u for which X1(u)+X2(u)=d_out
# Single log interpolation - works OK for this
    u12=numpy.exp(Interpolation.interpolate(x1x2,numpy.log(P_int),d_out))
    u1=numpy.exp(Interpolation.interpolate(x1_int,numpy.log(P_int),d_out))
    u2=numpy.exp(Interpolation.interpolate(x2_int,numpy.log(P_int),d_out))

# Use copula to find probability for this uniform variate
    P_x1x2=numpy.exp(-((-numpy.log(u12))**(1/alpha)+(-numpy.log(u12))**(1/alpha))**alpha)
   JP_ratio=P_x1x2/(u1+u2-u1*u2)

    if JP_ratio<0.5:
        JP_class=0
    elif JP_ratio>1.0:
        JP_class=2
    else:
        JP_class=1

    return P_x1x2, JP_class

def grids(rho):

    import numpy
    import ascii_raster
    import math
    import CombineIndP

    nodata_value=-9999.0

    (tide_200,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_200.asc")
    (tide_1000,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_1000.asc")
    (surface_100,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\s_rp100.asc")
    (surface_200,ni,nj,xll,yll,dx)=ascii_raster.read_ascii("C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\s_rp200.asc")

    surface_100[numpy.nonzero(surface_100<0.01)]=nodata_value
    surface_200[numpy.nonzero(surface_200<0.01)]=nodata_value

    nz=numpy.nonzero(surface_100>=surface_200)
    surface_100[nz]=nodata_value
    surface_200[nz]=nodata_value

    p1_tide=1./200
    p2_tide=1/1000.
    p1_tide_error=0.001
    p2_tide_error=0.0002

    p1_surface=1./100
    p2_surface=1/200.
```

```
    p1_surface_error=0.001
    p2_surface_error=0.0007


    depth_thresh=numpy.array([0.3,1.0,2.0])
    nd=len(depth_thresh)

    p_out_ub=numpy.zeros((ni,nj,nd))
    p_out_lb=numpy.zeros((ni,nj,nd))
    p_out_cep=numpy.zeros((ni,nj,nd))
    p_out_crv=numpy.zeros((ni,nj,nd))
    p_out_err=numpy.zeros((ni,nj,nd))
    p_out_ent=numpy.zeros((ni,nj,nd))
    p_JP=numpy.zeros((ni,nj,nd))
    p_JP_class=numpy.zeros((ni,nj,nd),dtype='int')

    maxd=3.0
    step=0.1

    d_out=depth_thresh

    for i, col  in enumerate(tide_200):
    for j, h in enumerate(col):

        p1_ub=0
        p1_lb=0
        p2_ub=0
        p2_lb=0
        p1_be=0
        p2_be=0

        JP_zero_flag1=0
        JP_zero_flag2=0

# Cases for different combinations of non-zero data. We can have no data, one value
# or both. As well as upper and lower bounds, we need to process "best estimate"
# values (somewhere between the lower and upper bounds) to calculate joint
# probabilities

# SOURCE #1

# Both values
        if tide_200[i,j]!=nodata_value and tide_1000[i,j]!=nodata_value:
          p1_ub=numpy.array([p1_tide+p1_tide_error,p2_tide+p2_tide_error])
          p1_lb=numpy.array([p1_tide-p1_tide_error,p2_tide-p2_tide_error])
          p1_be=numpy.array([p1_tide,p2_tide])
          d1=numpy.array([tide_200[i,j],tide_1000[i,j]])

          JP_zero_flag1=1

# One value only. Below this depth, the probability must be between the values
# associated with the two grids. We fudge this by inserting a data point below
# the p2 value to ensure we don;t interpolate linearly (see report section XXXX)
        elif tide_1000[i,j]!= nodata_value:
          p1_ub=numpy.array([p1_tide+p1_tide_error,p1_tide+p1_tide_error,p2_tide+p2_tide_error])
          p1_lb=numpy.array([p2_tide-p2_tide_error,p2_tide-p2_tide_error,p2_tide-p2_tide_error])
# Use flattest growth curve as "best estimate" - this will tend to overestimate
# effects of dependency
          p1_be=numpy.array([p1_tide+p1_tide_error+1e-6,p1_tide+p1_tide_error,p2_tide])
          d1=numpy.array([0.,tide_1000[i,j]-0.001,tide_1000[i,j]])
        else:
# No data at all - we can still say the probability of depth being greater than
# zero is less than the probability of the most extreme event grid
          p1_ub=numpy.array([p2_tide+p2_tide_error,p2_tide+p2_tide_error])
          p1_lb=numpy.array([1e-6,1e-6])   # Small +ve value - will take log later!
          p1_be=0.5*(p1_ub+p1_lb)    # Use CEP as best estimate
          p1_be[0]=p1_be[0]+1e-6
          d1=numpy.array([0.,1.])

# SOURCE #2
```

```
# Both values
        if surface_100[i,j]!=nodata_value and surface_200[i,j]!=nodata_value:
                p2_ub=numpy.array([p1_surface+p1_surface_error,p2_surface+p2_surface_error])
                p2_lb=numpy.array([p1_surface-p1_surface_error,p2_surface-p2_surface_error])
          p2_be=numpy.array([p1_surface,p2_surface])
                d2=numpy.array([surface_100[i,j],surface_200[i,j]])

          JP_zero_flag2=1
# One value only. Below this depth, the probability must be between the values
# associated with the two grids. We fudge this by inserting a data point below
# the p2 value to ensure we don't extrapolate incorrectly (see report section
# XXXX for plots and a description of how this works)
        elif surface_200[i,j]!= nodata_value:
          p2_ub=numpy.array([p1_surface+p1_surface_error,p1_surface+p1_surface_error,p2_surface+p2_surface_error])
          p2_lb=numpy.array([p2_surface-p2_surface_error,p2_surface-p2_surface_error,p2_surface-p2_surface_error])
# Use flattest growth curve as "best estimate" - this will tend to overestimate
# effects of dependency
          p2_be=numpy.array([p1_surface+p1_surface_error+1e-6,p1_surface+p1_surface_error,p2_surface])
          d2=numpy.array([0.,surface_200[i,j]-0.001,surface_200[i,j]])
# No data at all - we can still say the probability of depth being greater than
# zero is less than the probability of the most extreme event grid
        else:
          p2_ub=numpy.array([p2_surface+p2_surface_error,p2_surface+p2_surface_error])
          p2_lb=numpy.array([1e-6,1e-6])
          p2_be=0.5*(p2_ub+p2_lb)    # Use CEP as best estimate
          p2_be[0]=p2_be[0]+1e-6
          d2=numpy.array([0.,1.])

        p_comb_lb, p_comb_ub =CombineIndP.prob_from_depth(p1_lb,p1_ub,d1,p2_lb,p2_ub,d2,d_out)

# Store results to upper/lower bound probability grids
        for k,d in enumerate(depth_thresh):
          p_out_lb[i,j,k]=p_comb_lb[list(d_out-d).index(min(abs(d_out-d)))]
          p_out_ub[i,j,k]=p_comb_ub[list(d_out-d).index(min(abs(d_out-d)))]

          p_out_cep[i,j,k]=0.5*(p_out_lb[i,j,k]+p_out_ub[i,j,k])

# Can only calculate joint probability if both grids are non-zero for both
# sources
          if JP_zero_flag1 and JP_zero_flag2:
            for k,d in enumerate(depth_thresh):
              p_JP[i,j,k],p_JP_class[i,j,k]=JPrho(p1_be,d1,p2_be,d2,d,rho)
          else:
            p_JP[i,j,k]=-9999.
            p_JP_class[i,j,k]=-9999

      print "Processed column ", i, " out of ", ni


  p_out_crv=numpy.log(1./p_out_cep-1)

  p_out_err=100.*0.5*(p_out_ub-p_out_lb)/p_out_cep
  p_out_ent=-numpy.log2(p_out_cep)*p_out_cep-numpy.log2(1.-p_out_cep)*(1.-p_out_cep)

  for i,d in enumerate(depth_thresh):
  file_tail="_%3.1f.asc"%d

ascii_raster.write_ascii(p_out_lb[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-FMAS\TestCases\Hull\\tide_surface_p_lb"+file_tail)
ascii_raster.write_ascii(p_out_ub[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-FMAS\TestCases\Hull\\tide_surface_p_ub"+file_tail)
ascii_raster.write_ascii(p_out_cep[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-
FMAS\TestCases\Hull\\tide_surface_p_cep"+file_tail)
ascii_raster.write_ascii(p_out_err[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-FMAS\TestCases\Hull\\tide_surface_p_err"+file_tail)
ascii_raster.write_ascii(p_out_ent[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-
FMAS\TestCases\Hull\\tide_surface_p_ent"+file_tail)
ascii_raster.write_ascii(p_out_crv[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-
FMAS\TestCases\Hull\\tide_surface_p_crv"+file_tail)
ascii_raster.write_ascii(p_JP[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050 - FMAS\Test Cases\Hull\\tide_surface_p_JP"+file_tail)
ascii_raster.write_ascii(p_JP_class[:,:,i],xll,yll,ni,nj,dx,"C:\D\Projects\SC080050-FMAS\Test
Cases\Hull\\tide_surface_p_JPclass"+file_tail)
```

```
def JPrho(p1,d1,p2,d2,d_out,rho):
# Use look up table to determine for a given value of
# depth X what value of uniform variate u gives X=x1(u)+x2(u)
    import numpy
    import Interpolation
    import CDFs

# Choose set of probabilities to include all those given, plus points where
# the two curves reach zero depth
    P_int=numpy.concatenate((p1,p2),axis=0)

# Use log log interpolation works better at small depths - this is quite important
    P_zero1=numpy.exp(-numpy.exp(Interpolation.interpolate(numpy.array(d1),numpy.log(-numpy.log(p1)),0)))
    P_zero2=numpy.exp(-numpy.exp(Interpolation.interpolate(numpy.array(d2),numpy.log(-numpy.log(p2)),0)))

    P_zero1=min([P_zero1,1])
    P_zero2=min([P_zero2,1])
    P_zero3=0.5*(P_zero1+P_zero2)  # This point sometimes helps in interpolation

    P_int=numpy.append(P_int,[P_zero1,P_zero2,P_zero3])

# Remove duplicates and sort
    tmp=[]
    for v in P_int:
                                                if not v in tmp: tmp.append(v)

    P_int=numpy.array(tmp)
    P_int=P_int[P_int.argsort()]

# Interpolate depths for the two sources at the P_int values, ensure these are
# >0. Again use log log interpolation
    x1_int=Interpolation.int_prob(-numpy.log(p1),numpy.array(d1),-numpy.log(P_int))
    x2_int=Interpolation.int_prob(-numpy.log(p2),numpy.array(d2),-numpy.log(P_int))

    x1_int=numpy.maximum(x1_int,0)
    x2_int=numpy.maximum(x2_int,0)

    x1x2=x1_int+x2_int

#    alpha=numpy.log(2-max(min(chi,0.99),1e-6))/numpy.log(2)

# Find uniform variate u for which X1(u)+X2(u)=d_out
# Single log interpolation - works OK for this
    u12=numpy.exp(Interpolation.interpolate(x1x2,numpy.log(P_int),d_out))
    u1=numpy.exp(Interpolation.interpolate(x1_int,numpy.log(P_int),d_out))
    u2=numpy.exp(Interpolation.interpolate(x2_int,numpy.log(P_int),d_out))

# Convert uniform variates to normal variates
    u12_n=CDFs.UVN_inv(1-u12)
    u1_n=CDFs.UVN_inv(1-u1)
    u2_n=CDFs.UVN_inv(1-u2)

# Use copula to find probability for this normal variate
    a=CDFs.BVN(1000,u12_n,rho)
    b=CDFs.BVN(u12_n,1000,rho)
    c=CDFs.BVN(u12_n,u12_n,rho)

    P_x1x2=1-(a+b-c)
    JP_ratio=P_x1x2/(u1+u2-u1*u2)

    if JP_ratio<0.5:
        JP_class=0
    elif JP_ratio>1.0:
        JP_class=2
    else:
        JP_class=1

    return P_x1x2, JP_class
```

```
def JPchi(p1,d1,p2,d2,d_out,chi):
# Use look up table to determine for a given value of
# depth X what value of uniform variate u gives X=x1(u)+x2(u)
    import numpy
    import Interpolation

# Choose set of probabilities to include all those given, plus points where
# the two curves reach zero depth
    P_int=numpy.concatenate((p1,p2),axis=0)

# Use log log interpolation works better at small depths - this is quite important
    P_zero1=numpy.exp(-numpy.exp(Interpolation.interpolate(numpy.array(d1),numpy.log(-numpy.log(p1)),0)))
    P_zero2=numpy.exp(-numpy.exp(Interpolation.interpolate(numpy.array(d2),numpy.log(-numpy.log(p2)),0)))

    P_zero1=min([P_zero1,1])
    P_zero2=min([P_zero2,1])
    P_zero3=0.5*(P_zero1+P_zero2)  # This point sometimes helps in interpolation

    P_int=numpy.append(P_int,[P_zero1,P_zero2,P_zero3])

# Remove duplicates and sort
    tmp=[]
    for v in P_int:
    if not v in tmp: tmp.append(v)

    P_int=numpy.array(tmp)
    P_int=P_int[P_int.argsort()]

# Interpolate depths for the two sources at the P_int values, ensure these are
# >0. Again use log log interpolation
    x1_int=Interpolation.int_prob(-numpy.log(p1),numpy.array(d1),-numpy.log(P_int))
    x2_int=Interpolation.int_prob(-numpy.log(p2),numpy.array(d2),-numpy.log(P_int))

    x1_int=numpy.maximum(x1_int,0)
    x2_int=numpy.maximum(x2_int,0)

    x1x2=x1_int+x2_int

    alpha=numpy.log(2-max(min(chi,0.99),1e-6))/numpy.log(2)

# Find uniform variate u for which X1(u)+X2(u)=d_out
# Single log interpolation - works OK for this
    u12=numpy.exp(Interpolation.interpolate(x1x2,numpy.log(P_int),d_out))
    u1=numpy.exp(Interpolation.interpolate(x1_int,numpy.log(P_int),d_out))
    u2=numpy.exp(Interpolation.interpolate(x2_int,numpy.log(P_int),d_out))

# Use copula to find probability for this uniform variate
    P_x1x2=numpy.exp(-((-numpy.log(u12))**(1/alpha)+(-numpy.log(u12))**(1/alpha))**alpha)

    JP_ratio=P_x1x2/(u1+u2-u1*u2)

    if JP_ratio<0.5:
        JP_class=0
    elif JP_ratio>1.0:
        JP_class=2
    else:
        JP_class=1

    return P_x1x2, JP_class
# Univariate normal CDF from error function
def UVN(a):
    from scipy import special
    from math import sqrt

    return 0.5*(1+special.erf(a/sqrt(2)))

def UVN_inv(a):
    from scipy import special
```

```
    from math import sqrt

    return sqrt(2)*special.erfinv(2*a-1)

# Calculate cumulative bivariate normal cumulative distribution function
# Three parameters: variates a, b, and correlation rho

# f is a utility function
def f(x,y,aprime,bprime,rho):
    from math import exp
    r = aprime*(2*x-aprime) + bprime*(2*y-bprime) + 2*rho*(x-aprime)*(y-bprime);
    return exp(r)

def BVN(a,b,rho):

    from math import sqrt
    from scipy import sign

    pi=3.14159265358979323846264 3

    if (a<=0.0) & (b<=0.0) & (rho<=0.0):
        aprime = a/sqrt(2.0*(1.0-rho*rho))
        bprime = b/sqrt(2.0*(1.0-rho*rho))
        A=[0.3253030, 0.4211071, 0.1334425, 0.00637423]
        B=[0.1337764, 0.6243247, 1.3425378, 2.2626645]

        sum=0

        for i in range(4):
            for j in range(4):
                sum=sum+A[i]*A[j]* f(B[i],B[j], aprime,bprime,rho)


        sum = sum * ( sqrt(1.0-rho*rho)/pi)
        return sum

    elif ( a * b * rho <= 0.0 ):
        if ( a<=0.0 ) & (b>=0.0 ) & (rho>=0.0):
            return UVN(a) - BVN(a, -b, -rho)
        elif (a>=0.0) & (b<=0.0) & (rho>=0.0):
            return UVN(b) - BVN(-a, b, -rho);
        elif (a>=0.0) & (b>=0.0) & (rho<=0.0):
            return UVN(a) + UVN(b) - 1.0 + BVN(-a, -b, rho);

    elif a * b * rho >= 0.0:   # Actually do the calculation
        denum = sqrt(a*a - 2.*rho*a*b + b*b)
        rho1 = ((rho * a - b) * sign(a))/denum
        rho2 = ((rho * b - a) * sign(b))/denum
        delta=(1.0-sign(a)*sign(b))/4.0
        return BVN(a,0.0,rho1) + BVN(b,0.0,rho2) - delta

    return -99.9 # should never get here

# Approximation for on diagonal values
def BVNdiag(p,rho):
    pi=3.14159265358979323846264 3

    C_rho=(1+rho)**1.5*(1-rho)**(-0.5)*(4*pi)**(-rho/(1.+rho))

    return C_rho*p**(2/(1+rho))
```

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