Department for Environment Food & Rural Affairs

Llywodraeth Cymru
Welsh Government

Cyfoeth Naturiol Cymru
Natural Resources Wales

Environment Agency

delivering benefits through evidence

Spatial joint probability for flood and coastal risk management and strategic assessments

Method report

Report – SC140002/R1

Flood and Coastal Erosion Risk Management Research and Development Programme

We are the Environment Agency. We protect and improve the environment.

Acting to reduce the impacts of a changing climate on people and wildlife is at the heart of everything we do.

We reduce the risks to people, properties and businesses from flooding and coastal erosion.

We protect and improve the quality of water, making sure there is enough for people, businesses, agriculture and the environment. Our work helps to ensure people can enjoy the water environment through angling and navigation.

We look after land quality, promote sustainable land management and help protect and enhance wildlife habitats. And we work closely with businesses to help them comply with environmental regulations.

We can't do this alone. We work with government, local councils, businesses, civil society groups and communities to make our environment a better place for people and wildlife.

Published by:
Environment Agency, Horizon House, Deanery Road, Bristol, BS1 9AH
www.environment-agency.gov.uk

ISBN: 978-1-84911-394-6

Email: fcerm.evidence@environment-agency.gov.uk

Further copies of this report are available from our publications catalogue:
http://www.gov.uk/government/publications

or our National Customer Contact Centre:
T: 03708 506506

Email: enquiries@environment-agency.gov.uk

Author(s):
Tim Aldridge, Richard Allan, Ben Gouldby, Oliver Gunawan, Neil Hunter, Rob Lamb, Jonathan Tawn, Elizabeth Wood

Keywords:
Widespread, inland, flood, spatially coherent events, hazard, risk

Research Contractor:
JBA Consulting, South Barn, Broughton Hall, Skipton, North Yorkshire BD23 3AE
Tel: +44 (0) 1756 799919

Environment Agency's Project Manager:
Sue Manson, Evidence Directorate

Project Executive
Stefan Laeger, Improvements and Service Manager

Collaborator(s):
Health & Safety Laboratory
HR Wallingford
JBA Risk Management
Lancaster University
University of Reading

Project Number:
SC140002

Evidence at the Environment Agency

Scientific research and analysis underpins everything the Environment Agency does. It helps us to understand and manage the environment effectively. Our own experts work with leading scientific organisations, universities and other parts of the Defra group to bring the best knowledge to bear on the environmental problems that we face now and in the future. Our scientific work is published as summaries and reports, freely available to all.

This report is the result of research commissioned and funded by the Joint Flood and Coastal Erosion Risk Management Research and Development Programme. The Joint Programme is jointly overseen by Defra, the Environment Agency, Natural Resources Wales and the Welsh Government on behalf of all Risk Management Authorities in England and Wales:
http://evidence.environment-agency.gov.uk/FCERM/en/Default/FCRM.aspx.

You can find out more about our current science programmes at:
https://www.gov.uk/government/organisations/environment-agency/about/research.

If you have any comments or questions about this report or the Environment Agency's other scientific work, please contact research@environment-agency.gov.uk.

Professor Doug Wilson
Director, Research, Analysis and Evaluation

Executive summary

This report outlines the process and methodology of generating and selecting reasonable worst case scenarios used to inform the evidence base for the 2016 update to the National Risk Assessment of inland flooding risks. The method uses carefully collated historical river flow and rainfall data from the Heffernan and Tawn joint probability model (previously recommended in the Environment Agency research project SC060088 which investigated the spatial coherence of flood risk) to create a large number of extreme but statistically plausible events.

A set of extreme events defined at river flow and sub-daily rainfall gauges across England and Wales were developed into inland flood scenarios based on the statistical simulations.

The likelihood associated with each of the new scenarios can be assessed in terms of a hydrological proxy for the aggregate severity of the event, based on the average of the extreme flow or rainfall values over the network of gauges as well as a simplified measure of exposure. Although this is a very simple metric chosen as a proxy to indicate the relative severity of each event, this approach was considered a reasonable basis to inform and support the development of the scenarios through a blend of statistical analysis and expert judgement. The hydrometeorological plausibility of each event was also considered within this process, based on qualitative analysis of the large-scale climatological and meteorological drivers for flooding in the British Isles. A number of events were selected for consideration to provide example scenarios of extreme rainfall and river flow events that would cause flooding. One example joint fluvial–coastal scenario was also selected.

A narrative accompanying each scenario gives an overview of the hydrometeorological conditions under which these events may occur, accompanied by historical context comparing the scenario to past flooding events. To calculate the consequences of flooding, hazard modelling for each scenario was conducted to create a hazard footprint using 2D hydrodynamic modelling. The hazard footprints generated were then used in impact analysis by the Health and Safety Laboratory to convert receptor data to relevant metrics for flood risk assessment. The metrics give an indication of flood severity for each of the selected scenarios. The methods described here were used to inform evidence base for flooding risks in the National Risk Assessment 2016.

Acknowledgements

We would like to thank the Project Executive, Stefan Laeger, Project Board (Jodie Lofthouse, Graham Wilson, Dan Haigh and Anthony Hammond), Project Manager (Sue Manson) and consultees for their input.

We are grateful to the Centre for Ecology and Hydrology, Environment Agency, Met Office and Natural Resources Wales for supplying data for the project. We would also like to thank Stephen Blenkinsop at Newcastle University for sharing processed rainfall data derived from Environment Agency archives.

Contents

1 Introduction

1.1 Context

It is difficult to assess the probability of widespread or multiple source flooding events such as those in summer 2007 or winter 2013 to 2014. Recently developed joint probability methods offer more flexibility in representing such events. This project addresses the need for realistic probabilistic scenarios that account for spatially extensive events across England and Wales, and also multiple source events.

The model for spatially and temporally coherent extreme events developed by Lamb et al. (2010) and Keef et al. (2013), based on Heffernan and Tawn's (2004) conditional model for extremes, was recommended in a previous Environment Agency research project (SC060088, which investigated the risk of widespread flooding) as a flexible and practical data-based joint probability model (Environment Agency 2011a). This can be applied to meet this project's requirements.

Building on the approach used in SC060088, this project applied these methods to produce evidence and national scale modelling to inform the evidence base and scenarios for the 2016 update to the National Risk Assessment (NRA).

1.2 Project objectives

The project had 2 main objectives:

- To develop scenarios to inform the evidence for the 2016 update to the NRA to meet the needs of government and civil contingency planning at a national or broad scale
- To provide information on flood hazard joint probability methods, supported by updated tools and guidance, for use by flood and coastal risk management specialists

1.3 User requirements

The main end users of the methods produced during this project are the Department for Environment, Food and Rural Affairs (Defra) and the Environment Agency's Flood and Coastal Risk Management (FCRM) staff.

Representatives from Defra, Natural Resources Wales, the Environment Agency's National Capital Programme Management Services and Operational Modelling and Forecasting teams, and the Flood Forecasting Centre were consulted about user requirements.

This consultation process identified the following important findings.

- Users specified a need for analysis at a number of timescales – both shorter timescales (3–4 days) and longer timescales (3 months or seasonal).

- A desire was communicated for both local and broad scale outputs encompassing anything from a few gauges up to catchment wide analysis and national scale for the NRA.
- Aspirations were expressed for both detailed technical outputs and simpler, intuitive outputs so as to be able to address a range of operational questions identified by Natural Resources Wales and Environment Agency modelling and forecasting staff.

The needs of the users have been incorporated into the design of the scenarios, along with supporting tools and case studies. The full findings of the user requirements consultation can be found in Report SC140002/D1 (Environment Agency 2015).

1.4 National Risk Assessment

The scenario for inland flooding in the current NRA is expressed in terms of events with a likelihood (probability) of occurring at least once in the next 5 years of between 1 in 200 and 1 in 20. This is sometimes referred to as the ‘encounter probability’, defined as the exceedance probability of some prescribed ‘design’ level or event within a specified period.

This project improved the science underpinning the probability assessment of scenarios for the NRA and the quality of evidence through better quantification of scenario probability using existing tried and tested statistical methods. From the events generated using the Heffernan and Tawn joint probability method, a number of scenarios are selected to meet the requirements of the NRA (see Section 2.2).

The scenarios selected for the NRA will be ‘reasonable worst case scenarios’. These are events that have not yet occurred but are plausible, and provide an upper limit on the risks for which the government plans. It is also the situation against which infrastructure owners and operators can reasonably be expected to build resilience.

For the 2016 update to the NRA, the Cabinet Office used the following definition of a ‘reasonable worst case’:

> ‘It is a plausible and challenging expression of a risk to inform scalable and agile emergency planning. Subject-matter experts should consider that the scenario is:
>
> a) a civil emergency as defined by the Civil Contingency Act (2004);
>
> b) a challenge for central Government – this could be because it would overwhelm local/departmental resources and/or because it would require cross-government coordination for planning, response or recovery; and
>
> c) a credible risk the next 5 years – for naturally or accidentally occurring scenarios, industrial action or public disorder it should have at least a 1 in 20,000 chance of occurring in the next five years’.

2 Scenario generation

2.1 Method overview

The conditional exceedance method (Heffernan and Tawn 2004, Lamb et al. 2010, Keef et al. 2013), which uses a flexible and practical data-based joint probability model, was used to generate scenarios to meet the project's requirements. The method was reviewed and tested as part of project SC060088 (Environment Agency 2011a). It is particularly suitable for flood risk applications for 2 crucial reasons.

- It is based on theory that can capture the complex joint distributions and dependence structures observed in real environmental datasets, rather than imposing false assumptions. This can be particularly important in assessing the probability of extreme multivariate events including flood scenarios.
- It is a practical approach that can be applied to combinations of a few variables, or scaled up to handle problems of high dimensionality. The latter was demonstrated by JBA Consulting in an SC060088 case study which assessed event probabilities for notable events of widespread river flooding, and in subsequent developments of UK and European risk models for the reinsurance industry.

A full explanation of the method used in this project and applied for the scenario generation can be found in Lamb et al. (2010) and Keef et al. (2013).

This report discusses the method used for inland flooding. The equivalent method used for the coastal flooding scenarios is discussed in the coastal project method report (Defra 2016).

2.1.1 Fluvial and surface water scenarios

Fluvial and surface water scenarios were generated using data from river flow and rainfall gauges respectively. A large amount of flow and rainfall gauge data from the Environment Agency, Natural Resources Wales and the Met Office were quality checked and assessed for suitability for use in the statistical model. Gauges to be included in the model were selected based on the concurrency of record lengths, completeness of data and hydrometric quality. Further discussion of the quality checking measures used and issues encountered can be found in Appendix A. Events within the historical data were identified using the most appropriate timescale. For both fluvial and surface water, events were identified within a 7-day window.

Having prepared the data, a large number of statistically plausible events were generated by applying the methodology for coherent extreme events recommended in Environment Agency (2011a). Fluvial events were generated using data from 682 flow gauges (both daily mean and 15 minute). Short duration rainfall accumulations (across 3 hours) at 190 tipping bucket rain gauges (TBRs) were used to generate events representative of surface water flooding.

2.1.2 Joint fluvial–coastal scenario

Simulated events were generated by linking JBA analysis of inland fluvial extremes with HR Wallingford modelling of coastal extremes developed for the parallel coastal flooding scenario. A total of 32 coastal locations were used in the analysis. For each coastal location, the nearest inland gauge (without tidal influence) was identified. A further 10 inland gauges, distributed across England and Wales, were selected to capture any events occurring concurrently further inland. A total of 33 inland flow gauges were used in the event set generation.

2.2 Scenario selection process for the NRA

A high-level overview of the scenario development process is shown in Figure 2.1.

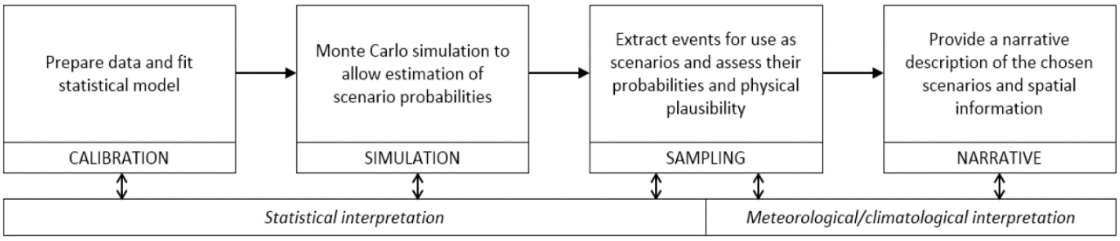


Figure 2.1 Scenario development process overview

To select scenarios for the NRA, a combination of objective analysis (the statistical model) and subjective criteria (the NRA requirements) were used. The requirements for the evidence base for the NRA, as set out in the project proposal and during discussions with the project board, were for:

- 4 fluvial flooding scenarios
- 2 surface water flooding scenarios
- 1 joint inland–coastal scenario

An overview of the process used to develop and select the required scenarios for the 2016 update to the NRA is shown in Figure 2.2.

Each selected scenario is accompanied by a narrative providing a scenario overview and meteorological context for the event, highlighting the conditions which may lead to flooding as presented in the scenario. The narrative also includes some flood history context, relating the scenarios to past historic flood events and providing a comparison of severity.

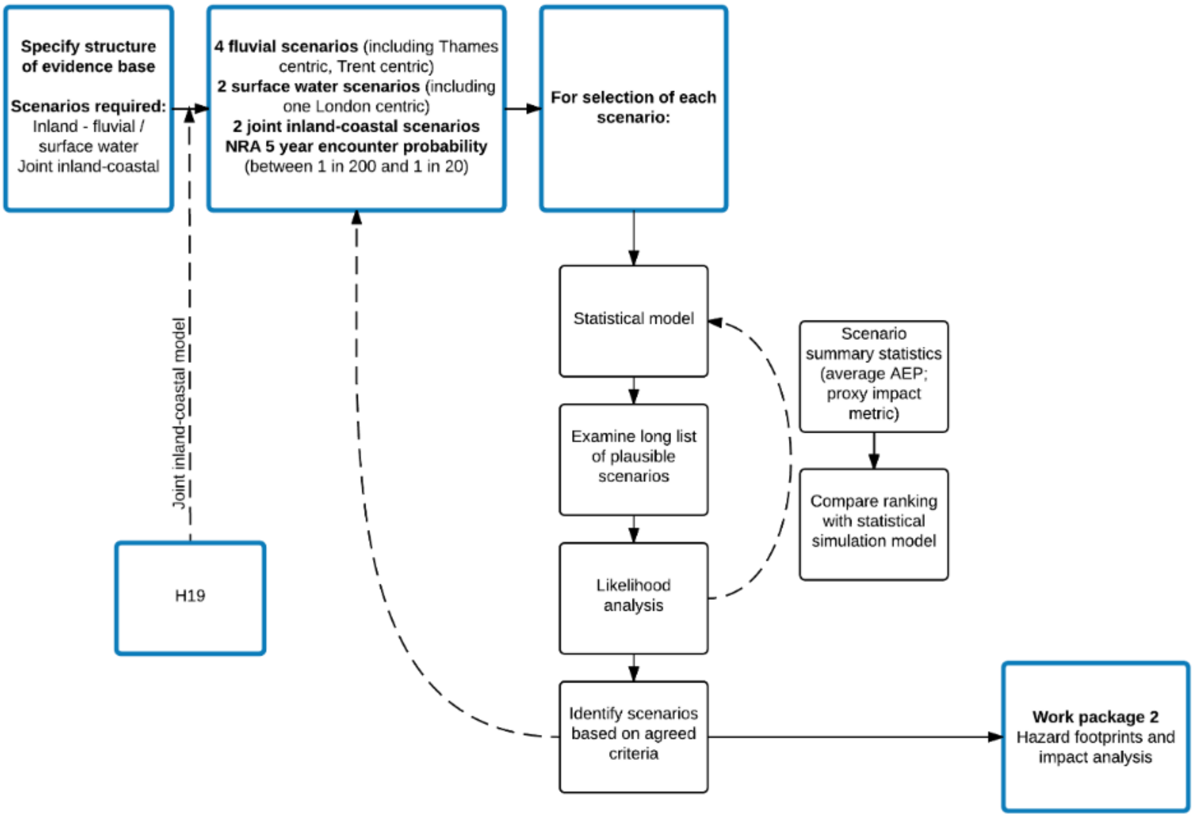


Figure 2.2 Scenario selection process

2.2.1 Assessment of event likelihoods

A range of scenarios were available for selection. Those with lower likelihoods were more extreme with more damaging impacts and those with higher likelihoods had smaller impacts. The new inland scenarios developed during this project use the same likelihood criteria as selected in the 2015 National Risk Register (that is probabilities of between 1 in 200 and 1 in 20 over the next 5 years) (Figure 2.3). The subset of scenarios was refined further using this range of likelihoods.

Figure 2.3 National Risk Register matrix

Source: Cabinet Office (2015a)

Each scenario's likelihood of occurrence over the next 5 years was used to enable selection of those that met the requirements for the NRA. Events generated by the statistical model are defined on an annual probability scale and needed translating to a likelihood of occurring over the next 5 years to determine which ones met the requirements of the NRA. Figure 2.4 shows how the annual probabilities were converted to 5-year likelihoods. Ultimately, scenarios that meet the requirements of the 5-year likelihood range of the NRA must have a 5-year probability of between 1 in 20 and 1 in 200.

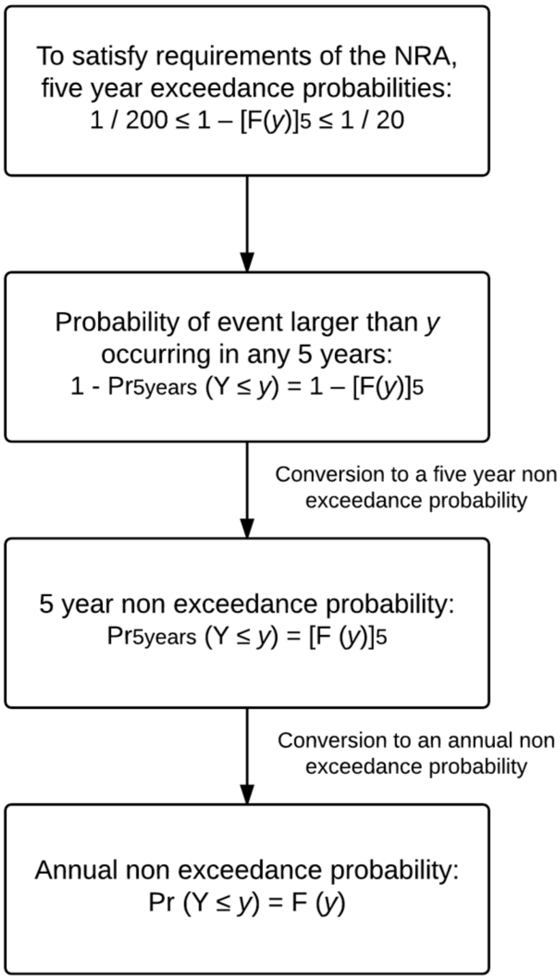


Figure 2.4 Flow chart showing process to determine required event annual exceedance probabilities for consideration of fluvial and surface water scenarios for the NRA

Constructing the full, unconditional distribution of flood impacts was not possible for the following 2 reasons.

- It would be very complex and very expensive.
- There are numerous alternative ways to quantify impact that would need to be considered and it was not clear which impact metric would be the most appropriate choice.

As a result, the statistical simulation was used as a guide to the event likelihood and combined with subjective interpretation to select the scenarios (as shown in Figure 2.2 – scenario summary statistics and ranking comparison). A number of summary statistics were calculated for each event as a proxy for the potential seriousness of the flooding. Several different summary statistics were tested and the project board chose to use the mean of the local flow extremes – standardised in terms of annual exceedance probability (AEP) – as the metric for scenario selection.

Fluvial and surface water events were ranked hydrologically by the mean. To refine the list of eligible scenarios, the approximate annual likelihood of each simulated event was estimated, using Equation 2.1:

Equation 2.1 Approximate annual likelihood = $k / (N + 1)$

where:

k = rank of a given event

N = number of years simulated (10,000 in this instance).

The most extreme simulated event has an AEP of 1 in 10,000. The rank 2 event has an AEP of 1 in 5,000 and so on for all simulated events.

The simulated events needed to satisfy the following criteria:

$$1{,}000 \geq (10{,}000 + 1) / k \geq 100$$

giving values of k between 10 and 100. The simulated events that can be considered for the NRA need to be between the 10th and 100th ranks, with rank 10 being the most extreme event of this selection and subsequent ranks decreasing in extremeness.

Potential fluvial scenarios within the required range for the NRA were also ranked using an exposure-based score. Property counts in Flood Warning Areas (FWAs) were used as a proxy for exposure to risk. For each simulated event, the number of properties in a FWA was multiplied by the event probability at the nearest gauge. The scores for each FWA were summed to create an event exposure-based score. The rank 1 event was the most severe, with the largest exposure-based score. A total of 15 fluvial events were within the required range of the NRA (ranks 10 to 100) using both the hydrological and exposure-based ranking methods, and were considered as potential scenarios for the NRA.

All fluvial and surface water events meeting the requirements for the NRA were examined to determine the hydrometeorological plausibility, drawing on the work by the University of Reading to identify meteorological conditions associated with widespread flooding in the UK (Allan 2015). Any scenarios that were not meteorologically plausible were removed from the process and were not considered for inclusion in the NRA scenarios.

2.2.2 Joint inland–coastal scenario selection

One joint inland–coastal event was required for the evidence base for the 2016 update to the NRA. The joint fluvial–coastal scenario was developed by applying the same statistical model to both the fluvial and coastal offshore datasets simultaneously. To ease the computational burden, a subset of fluvial gauges was used to carry out the multivariate analysis. Using the fitted multivariate extreme value model, 10,000 years' worth of joint coastal and fluvial events were synthesised.

The scenario to be simulated was then selected from these by considering events that were extreme in both fluvial and coastal aspects. Nearby sea level and fluvial gauges were paired and the highest exceedance probability for each pair identified. Each event was assigned the lowest (that is, the most extreme) exceedance probability of this pair at all locations. The events were then ranked according to this metric and 9 events were found to be associated with the relevant range of likelihoods based on the National Risk Register likelihood–impact matrix. These events were screened manually and one event that met the widespread flooding criteria was identified. Although this method does not provide a joint likelihood assessment, it provides a reasonable basis for shortlisting plausible joint events.

To fulfil the inland element of the joint inland–coastal scenario, an event was selected from the generated inland fluvial event set. To select the fluvial scenario, the squared distance was calculated to compare the AEPs of the 10 gauges in the joint scenario with all events in the fluvial event set. The 5 scenarios with the smallest squared distance were selected and the meteorological plausibility of each considered. One event was selected for the scenario and used for the hazard modelling (see Section 3).

2.3 Input data

Flow and rainfall data were obtained for gauges across England and Wales from the Environment Agency, the National River Flow Archive (NRFA) held by the Centre for Ecology and Hydrology (CEH), Natural Resources Wales and the Met Office. A summary of the data obtained is provided in Table 2.1.

To achieve a balance between having a sufficient sample size to fit a statistical model and retaining a good spatial coverage of gauges, expert judgement deemed that gauge records should have a minimum length of 20 years. The number of gauges meeting this requirement is also included in Table 2.1. Extensive quality checks were made to ensure that gauges were suitable for inclusion in the analysis. Further discussion of the quality checks performed and the issues encountered can be found in Appendix A.

Table 2.1 Data summary

	Total number of gauges	Number of gauges with ≥20 years' data	Percentage of gauges with ≥20 years' data
Daily mean flow (CEH NRFA)	485	485	100%
15-minute flow (Environment Agency and Natural Resources Wales)	1,347	906	67%
Hourly TBR (Environment Agency and Met Office)	1,426	455	32%
Storage rainfall gauge (Environment Agency)	1,866	1,314	70%

2.3.1 River flow

Flow data were obtained from 2 sources:

- daily mean flow data from the NRFA at CEH
- sub-daily flow data from the Environment Agency and Natural Resources Wales

Data from the NRFA had previously been obtained for a large number of gauges and quality checked as part of project SC060088 (Environment Agency 2011a). Updated data for these gauges were requested and used in the analysis for this project.

The required data input for the model was one flow value per gauge per day. The NRFA gauge records had a daily resolution and did not require any further processing. The sub-daily data from the Environment Agency and Natural Resources Wales had a temporal resolution of 15 minutes. Data for these gauges were processed, taking the maximum flow value per day to achieve this. The locations of 682 flow gauges used in the model are shown in Figure 2.5.

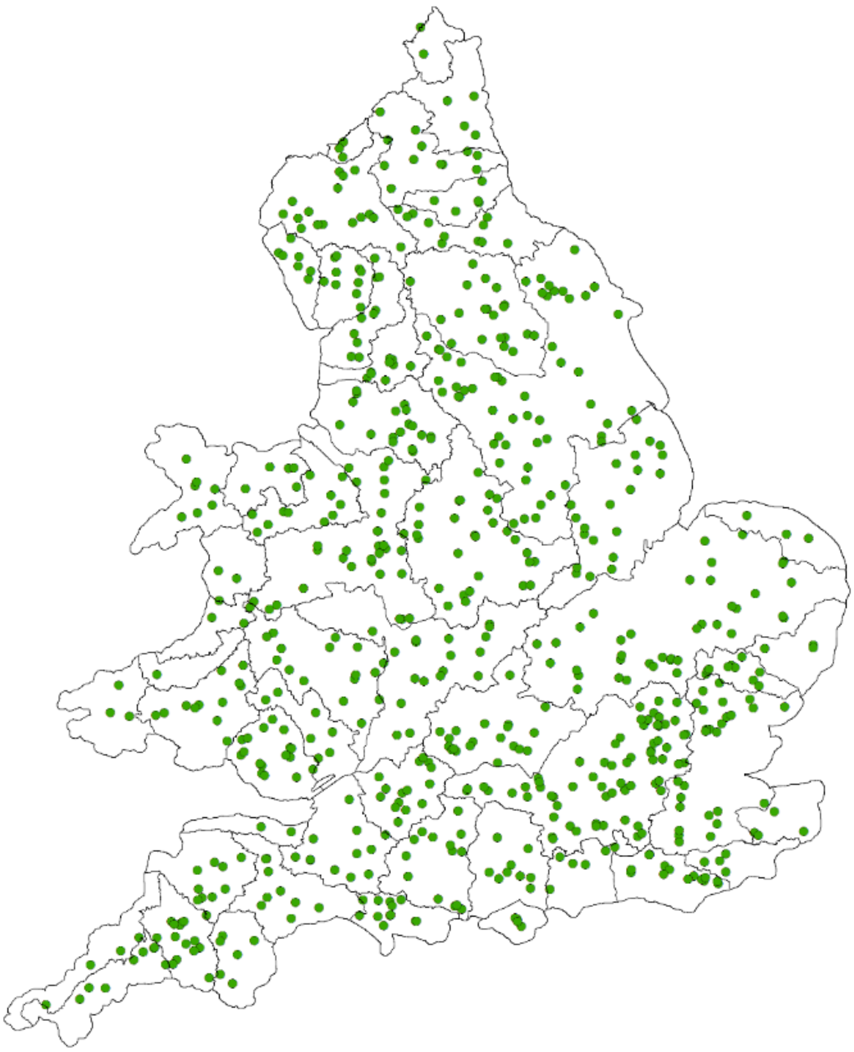


Figure 2.5 Locations of flow gauges in England and Wales used in the statistical joint probability model

2.3.2 Rainfall

Rainfall data were initially obtained from the Environment Agency for TBRs across England and Wales. The raw data received from the Environment Agency were recorded as 'time of tip' data. Initially, the data were acquired in hourly accumulations and needed processing to achieve these accumulations. Newcastle University had used the same TBR data for another project and had processed the data into hourly accumulations and completed quality checks for over 1,400 gauges across England and Wales. The checked data from Newcastle University were transferred to JBA for use in the joint probability modelling.

Any gauges with record lengths shorter than 20 years were immediately removed from the analysis. Records for TBRs tend to be much shorter than flow records and removing those with lengths less than 20 years considerably reduced the number of gauges available for inclusion in the analysis.

On inspection of the spatial coverage of gauges with the minimum data length, it became clear that there was a considerable gap in gauge coverage across south-west England. Although some gauges were located in this region, very few had the required 20 years of data needed. To address this issue, the Met Office was approached to supply data for 10 selected TBRs across south-west England that had a record length in excess of 20 years.

To capture potential surface water flooding events, 3-hourly accumulations were calculated for all gauges to be included in the event set. As with the flow data, one rainfall value per day was required as the input for the analysis and the maximum of the 8 accumulations across the day was used. The locations of 190 rainfall gauges used are shown in Figure 2.6.

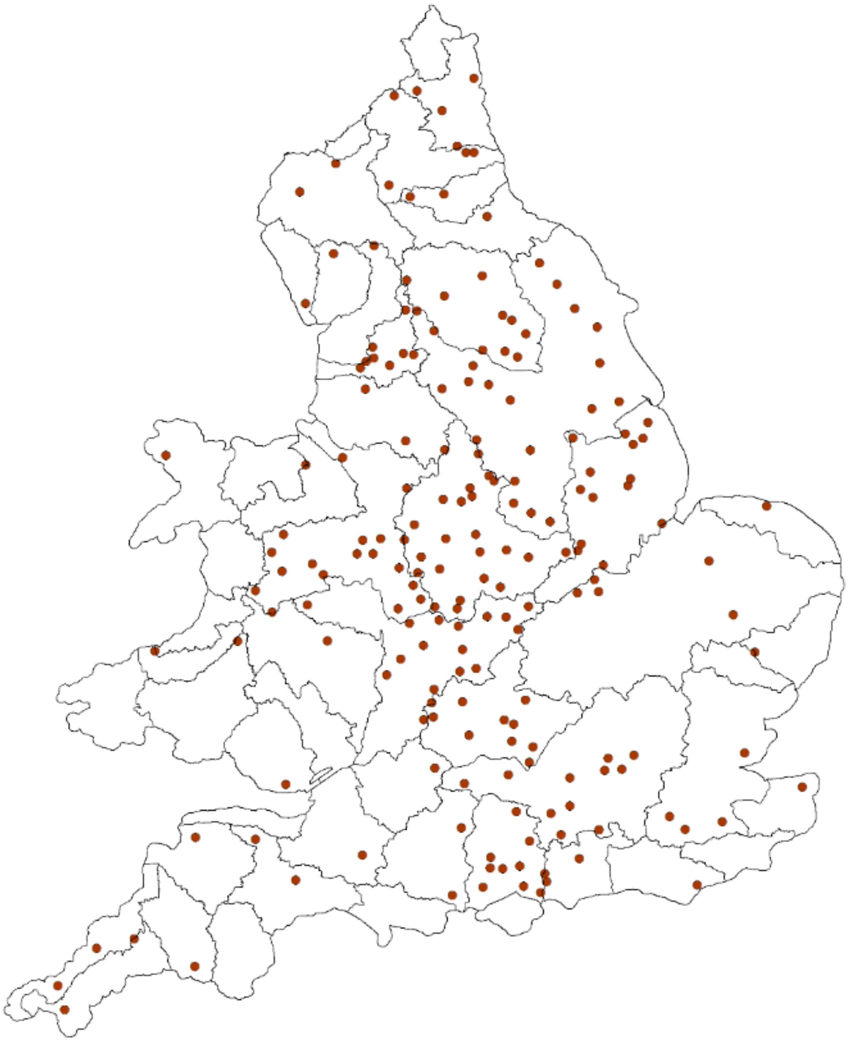


Figure 2.6 Locations of rainfall gauges used in the event set

Data for a number of storage rainfall gauges across England and Wales were also received. These records contained daily rainfall totals and, in some instances, monthly rainfall totals. As occurrences of surface water tend to be the result of short duration rainfall across a few hours, the storage rainfall gauge data were not suitable for inclusion in an event set looking at surface water flooding. The data were still utilised with calculated daily rainfall totals from the TBRs compared with nearby storage gauges as part of the quality checking process.

2.3.3 Coastal

The coastal analysis used methods and data established under the parallel coastal R&D project (Defra 2016). The relevant variables used in the coastal analysis include information on waves, winds, surges and tides. These data were initially analysed offshore. The datasets containing these variables were obtained from the 'A Class' National Tide and Sea Level Facility network of tide gauges managed by the Environment Agency and Natural Resources Wales, and an historic hindcast analysis of waves and winds derived by the Met Office using the WaveWatch III (WW III) model. The locations of the specific data used are shown in Figure 2.7.

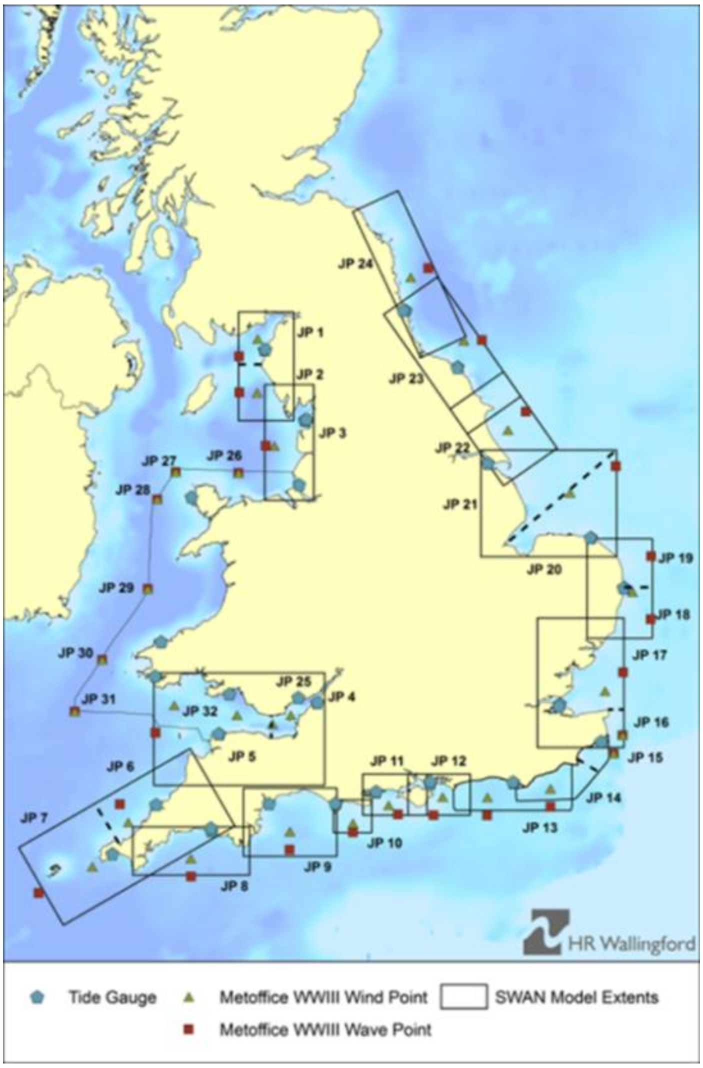


Figure 2.7 Locations of coastal waves, winds and surges and SWAN model grids

Source: Defra (2016)

2.4 Implementation

The Heffernan and Tawn (2004) methodology was implemented using an open source code package – texmex – in the R statistical programming language (CRAN 2013). The implementation of the method is discussed in more detail in Lamb et al. (2010) and Keef et al. (2013), and also in the report for project SC060088 (Environment Agency 2011a).

2.5 Example output

A large number of statistically plausible events were generated for fluvial, surface water and joint inland–coastal events. For each event, the output consists of an AEP at each individual flow or rainfall gauge. The generated events can be displayed on a plot such as that shown in Figure 2.8, where the AEP at each gauge is plotted, allowing the user to identify areas impacted by extreme AEPs. The event shown in Figure 2.8 affects much of the country with a band of low AEPs extending from south-west England through the Midlands. A large number of gauges have local AEPs of less than 1 in 500. Summary statistics for this event, including the 5-year encounter probability, are given in Table 2.2.

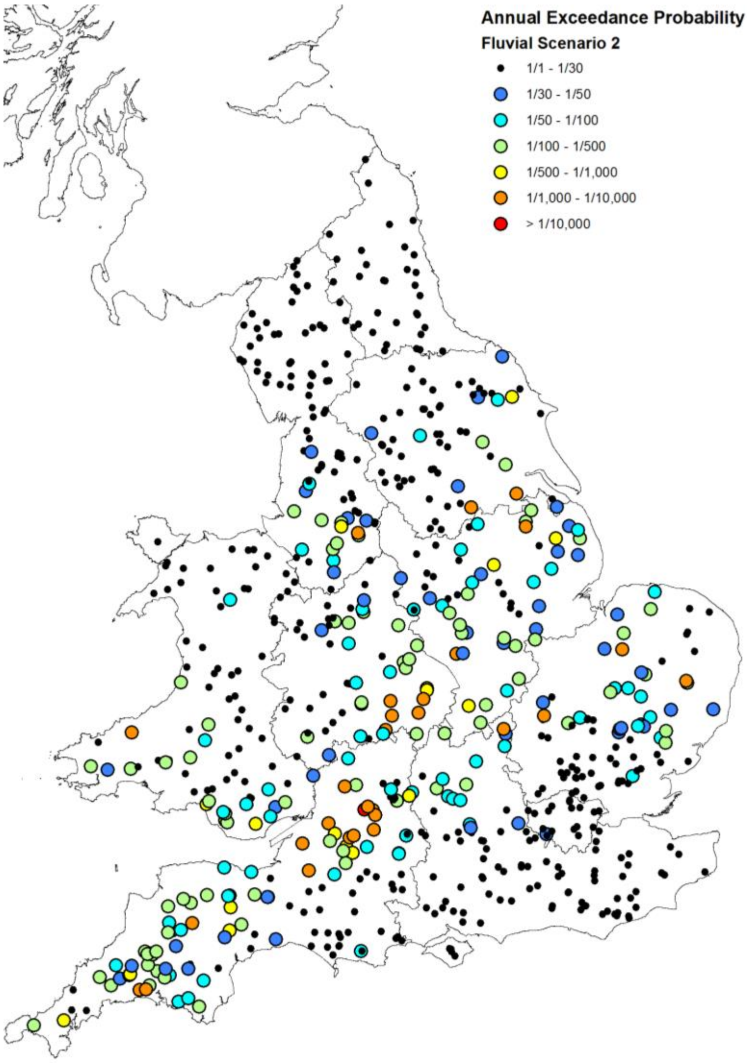


Figure 2.8 Example output – fluvial scenario

Table 2.2 Summary statistics for example output

5-year encounter probability	1 in 50
Minimum AEP	< 1 in 10,000
Mean AEP	1 in 242.69
Median AEP	1 in 10.27
Skewness	23.42

3 Hazard modelling

3.1 Method overview

To calculate the consequences of flooding associated with selected event scenarios, hazard footprints were developed using 2-dimensional (2D) hydrodynamic modelling. Probabilities evaluated at each river flow/rainfall gauge were interpolated to the required spatial coverage and combined with standard UK flow and rainfall estimation methods to provide suitable boundary conditions to the hydraulic models. The models were then run on a high throughput computing (HTC) cluster and the outputs combined to produce a single flood map for each event. These flood maps provide the basis for calculating economic damages and other impact metrics described in Section 4.

3.2 Fluvial

3.2.1 Method, input data and implementation

Step 1 – Event interpolation based on catchment centroids

The first stage in the hydraulic modelling approach developed here requires the probability estimated at each river flow gauge for a given event to be interpolated across the Environment Agency's Detailed River Network (DRN) dataset (Figure 3.1).

To perform the interpolation, the DRN is converted into distinct reaches and points created at regular 1km intervals. Points are also added at the start and end of reaches and 300m upstream and downstream of confluences. The centroid of the catchment upstream from each 'interpolation point' is then used as the starting point for a local search to identify the centroids of catchments upstream of each gauge location where probability has been estimated. Once the 5 closest gauges have been located, the event-specific probability is calculated at the interpolation point based on an inverse distance weighting of the probabilities estimated at those gauges. This process is described in more detail in Environment Agency (2011).

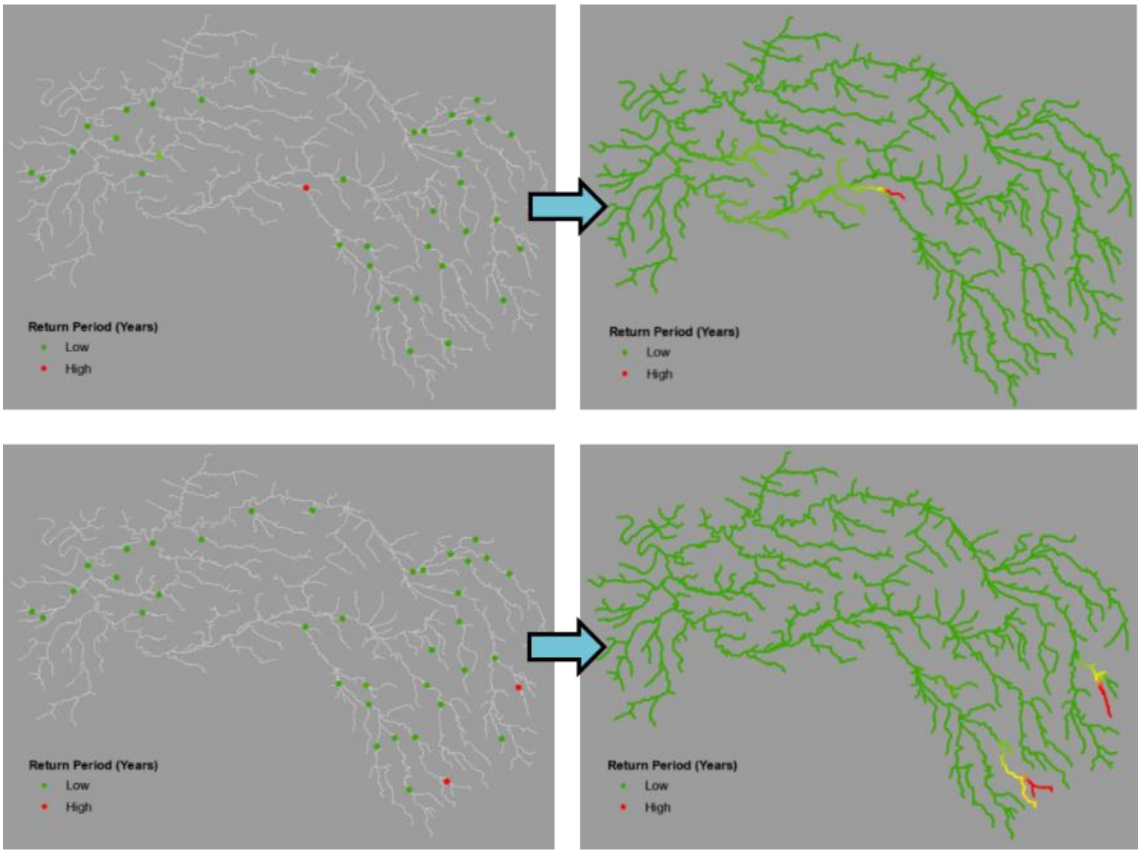


Figure 3.1 Interpolation of point probability values across the river network

Step 2 – Inflow calculation based on Flood Estimation Handbook flows with defence standard of protection or channel capacity adjustment

The next step is to convert the probability interpolated at each point into a boundary condition suitable for a hydraulic model. The original intention was to link the local probability estimate to the Environment Agency's National Fluvial Levels (NFLD) and Continuous Defence Line (CDL) datasets and calculate a per asset inflow based on the water level difference and estimated duration as per Modelling and Decision Support Framework 2 (MDSF2). However, initial results from the ongoing Environment Agency's State of the Nation risk modelling programme discouraged this approach in the time available and an alternative method based on flow was developed instead.

At each interpolation point along the DRN (approximately 125,000 in total), peak flow estimates and design hydrograph shapes were derived for the 1 in 2, 1 in 5, 1 in 10, 1 in 25, 1 in 50, 1 in 75, 1 in 100, 1 in 200 and 1 in 1,000 AEPs using the following automated approach and catchment descriptors extracted from version 3 of the Flood Estimation Handbook (FEH) CD-ROM:[1]

1. Estimation of the index flood QMED (median annual maximum flood) using the regression equation published in Environment Agency (2008).
2. Adjustment of QMED by automatic identification of a donor site, again using the procedures from Environment Agency (2008). The donor catchment with the closest centroid is chosen and the adjustment factor moderated so that it reduces for more distant donors.

[1] http://www.ceh.ac.uk/services/flood-estimation-handbook

3. Adjustment of QMED for urbanisation using the procedure given by Kjeldsen (2010).

4. Construction of a pooling group. A group is developed for each flow estimation point using the HiFlows-UK dataset. At the time of the analysis, the current version of the dataset was 3.3.4, released in August 2014.

5. Development of a pooled growth curve using the methods described in Environment Agency (2008) to weight results from gauges in the pooling group. A generalised logistic distribution is used to represent the growth curve. The growth curve is adjusted for urbanisation using the methods from Kjeldsen (2010).

6. Scaling of the growth curve by QMED to give the design flows for the return period(s) needed.

7. Development of a hydrograph shape using the Revitalised Flood Hydrograph (ReFH) method, applied with default options for catchment descriptors and storm event characteristics.

8. Scaling of the hydrograph so that its peak matches the peak flows estimated from the FEH statistical method.

A series of tools were developed to check the outputs from the automated process and to discard any anomalous results. However, it is important to appreciate that any automated application of hydrological methods, such as this, cannot be expected to result in design flows that are as reliable as those that benefit from local knowledge, expert judgement, additional sources of information and all the other components that go into detailed hydrological studies. An example of the hydrological information available at each interpolation point is shown in Figure 3.2.

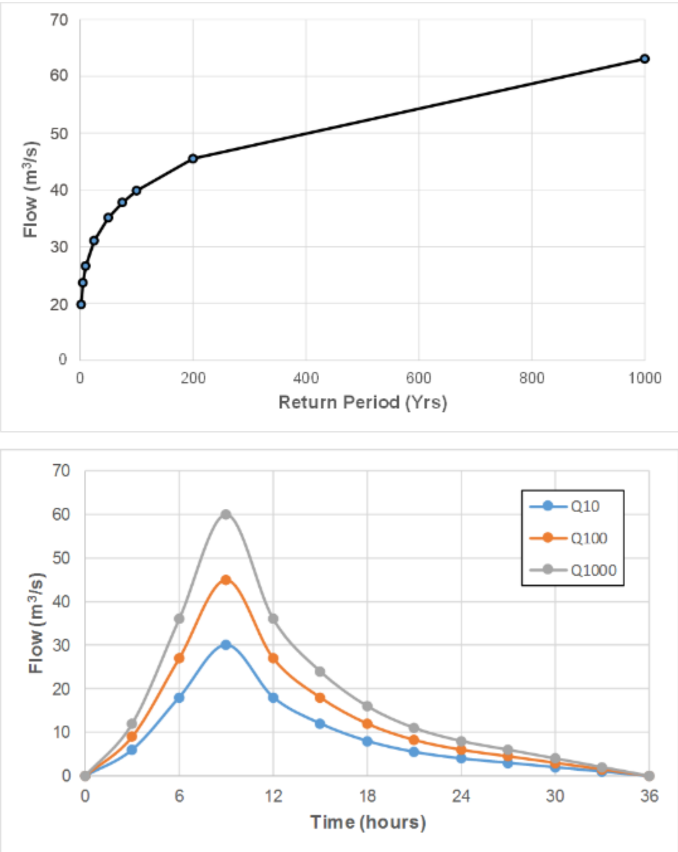


Figure 3.2 Peak flow estimates and hydrograph shapes derived at each interpolation point using FEH methods and data

Inflows to the hydraulic models are calculated on a per asset basis using spatial and attribute information contained in the CDL and National Flood Risk Assessment 2008 (NaFRA08) tramline datasets available across England and Wales respectively. Each asset is linked to the nearest interpolation point along the river network and the standard of projection (SoP) attributed to the asset is used to threshold the inflow hydrograph to determine a net inflow into the floodplain (Figure 3.3). In undefended areas, the SoP for each asset section is set to a 1 in 5 AEP to represent a nominal channel capacity as per the State of the Nation programme.

In the vast majority of situations, the interpolated probability at any given location is unlikely to correspond to one of the 9 AEPs for which peak flows and hydrograph shapes have been pre-determined. Therefore, simple linear weighting of the 2 nearest flows/shapes is used to produce a hydrograph (or SoP threshold) for the required probability.

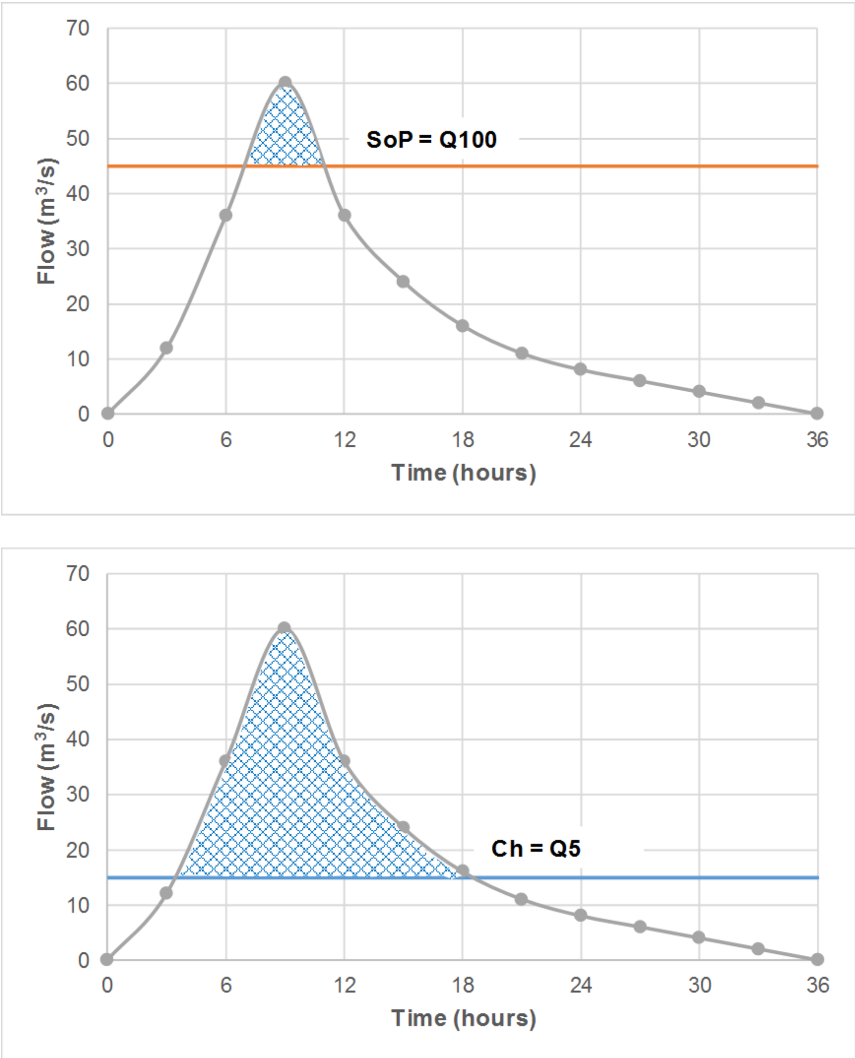


Figure 3.3 Thresholding of design hydrograph by SoP in defended areas (top) or channel capacity in undefended areas (bottom)

Notes: Net inflow into the floodplain is denoted by the blue hatched area in each case.

Step 3 – 2D hydraulic modelling and flood mapping

It is not currently possible to model an entire widespread flood scenario within a single hydraulic model. Therefore, the approach developed here has sought to reuse available data with complete coverage across England and Wales in a process that can be entirely automated. As such, the flood hazard footprint for each event is actually built from hundreds of thousands of individual hydraulic model simulations.

Hydraulically discrete Flood Areas, available from the State of the Nation programme in England and NaFRA08 in Wales, are used to define the extent of each hydraulic model (see Figure 3.4a). Modelling from the State of the Nation and NaFRA08 also provides a continuous description of individual asset sections along the channel–floodplain interface of each Flood Area. Here the term 'asset' describes both distinct sections of raised flood defence and top-of-bank in undefended areas.

As described above, a net inflow hydrograph for each event is derived at each asset which is subsequently adjusted by either an assumed channel capacity in undefended areas or the SoP in defended areas. This adjusted hydrograph is then applied along the individual asset section within the 2D hydraulic model and the resulting inundation spreads across the Flood Area according to local topography (Figure 3.4). The external perimeter of the hydraulic model is assigned a transmissive boundary condition to allow water to leave the simulation domain and avoid artificial over-deepening through 'glass walling'.

Simulations are run independently for all asset sections within the Flood Area where there is a positive net inflow (that is, $Q_{in} > Q_{SoP}$ or 1 in 5 AEP in defended and undefended areas respectively) and the results are mosaicked together using a maximum operator to produce a combined flood map for each Flood Area. Flood maps for each Flood Area are then combined to produce a single hazard footprint for each event; see the example in Figure 3.5 and Figure 3.6.

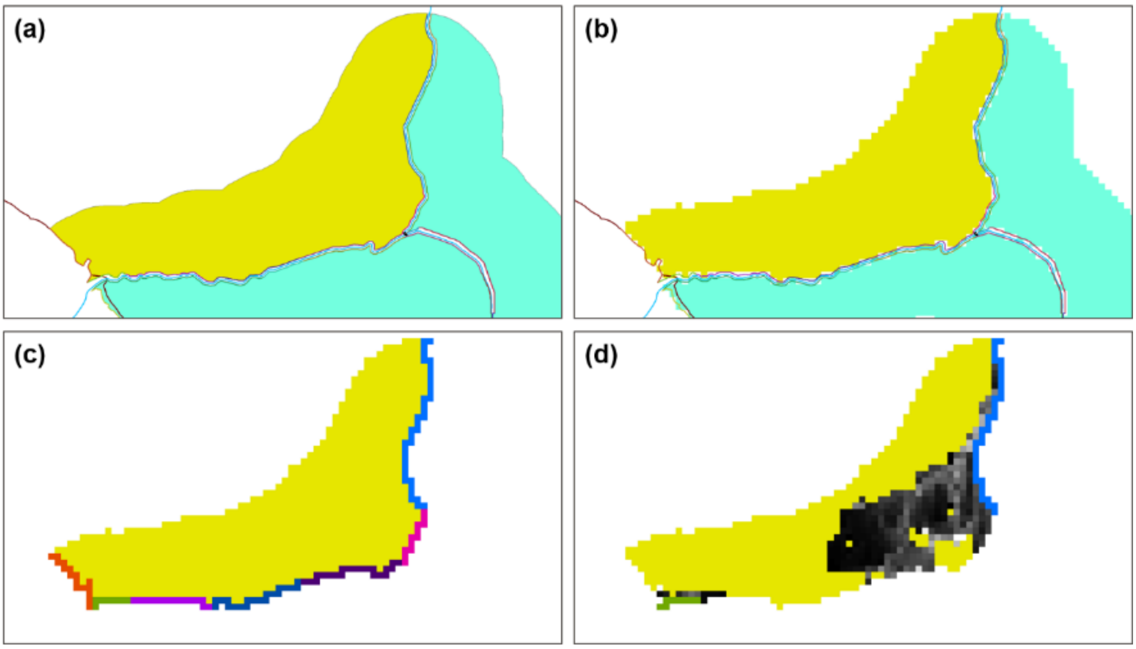


Figure 3.4 (a) Automated set-up of 2D hydraulic models using State of the Nation Flood Areas. Flood Areas (b) and individual asset objects (c) are rasterised to provide the simulation domain extent and inflow locations respectively. (d) Simulations are run on a per asset basis where there is a positive net inflow and the outputs combined using a maximum mosaic operator to produce a final flood map.

The flood modelling is carried out on a regular grid of 10m × 10m resolution parameterised using topographic information from the Environment Agency's Integrated Height Model 2014. Hydraulic roughness is described using a single Manning's *n* value of 0.1 and the model's output maps of maximum extent, depth, velocity and hazard rating. These footprints subsequently provide a much richer basis for conducting national flood impact assessments at an unprecedented spatial scale.

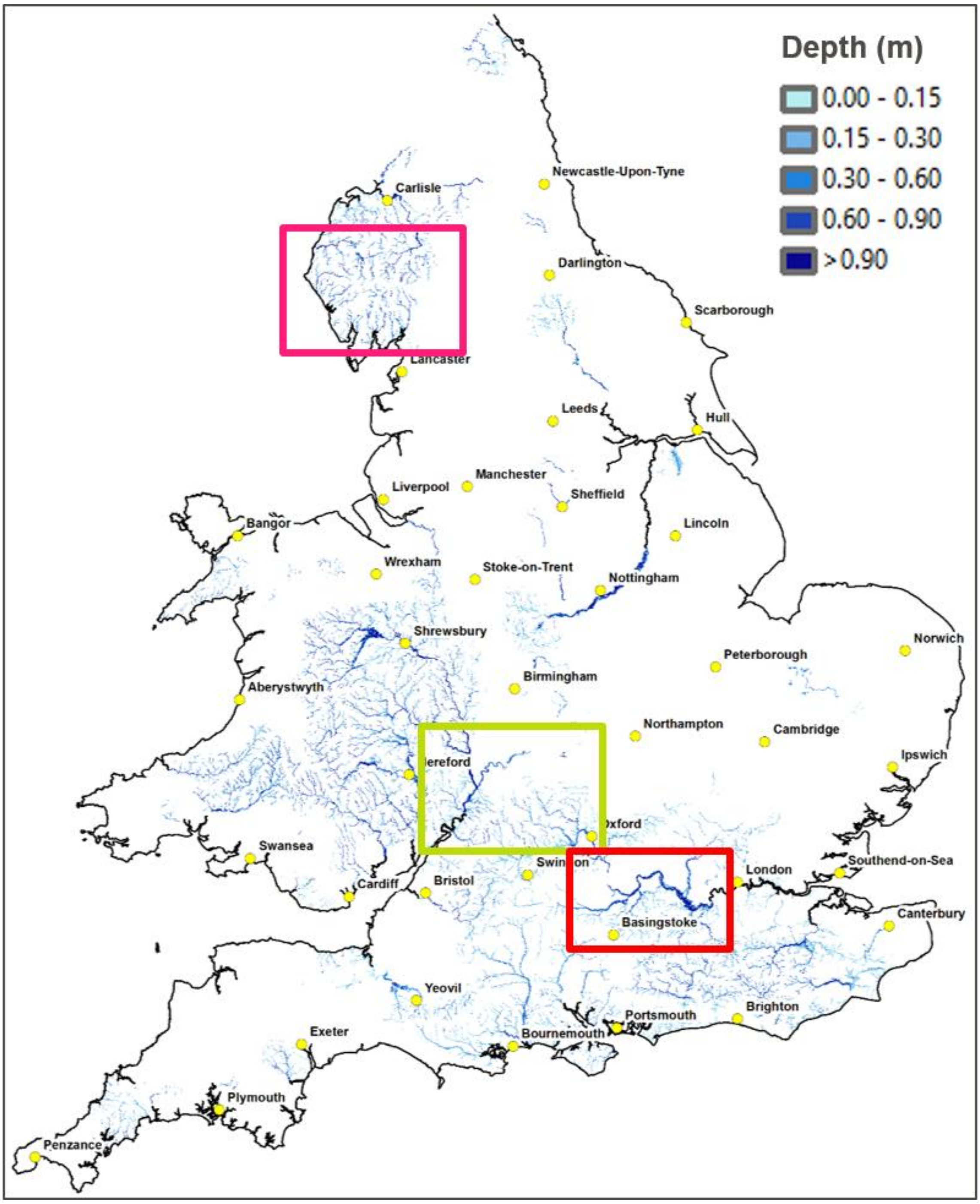


Figure 3.5 Widespread floodplain inundation associated with an extreme river flow scenario

Notes: More detailed subplots for each coloured rectangle are provided in Figure 3.6.

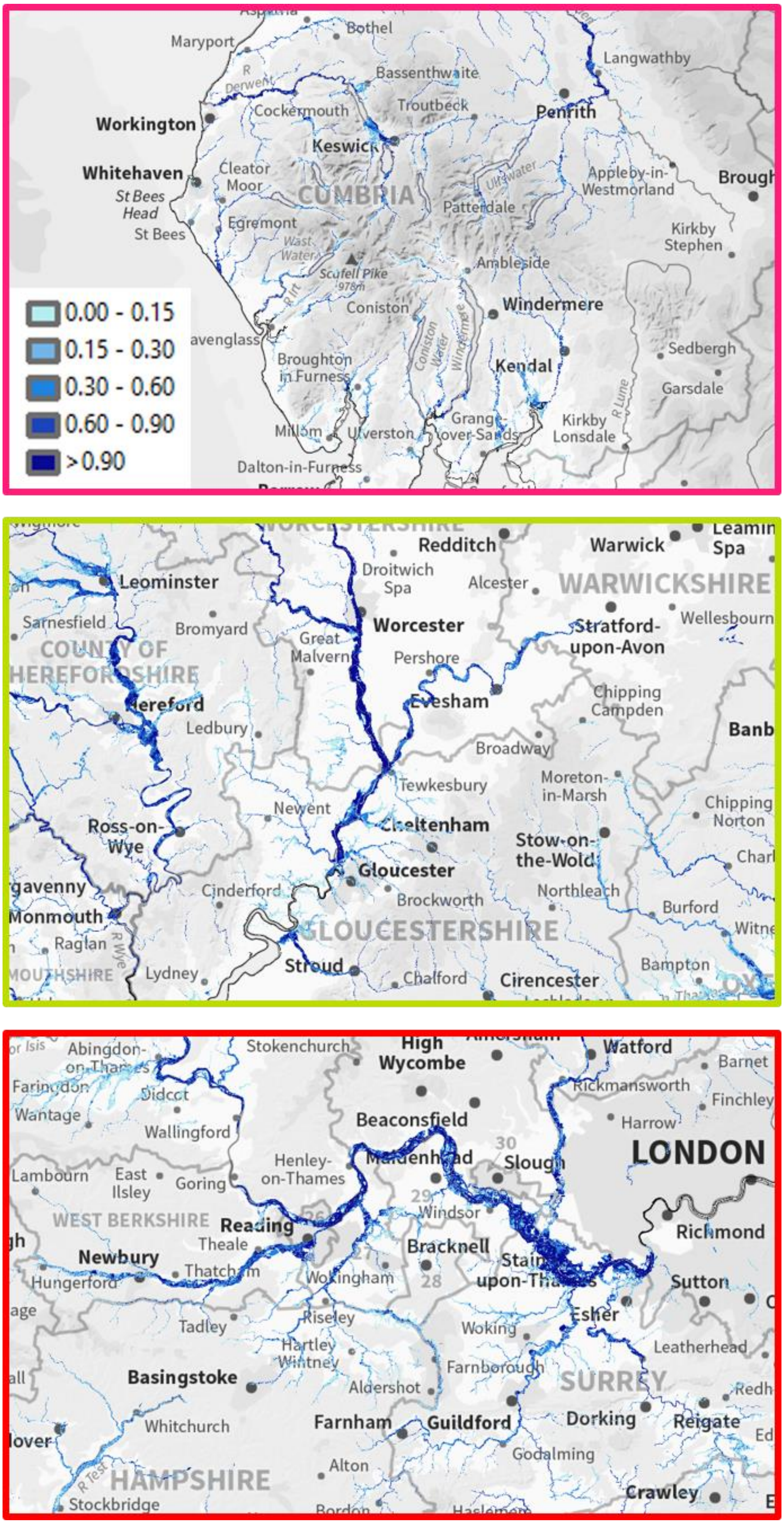


Figure 3.6 Detailed examples of widespread floodplain inundation associated with an extreme river flow scenario for areas identified in Figure 3.5

Step 4 – Breach modelling

For each event, the effect of breaching was considered at a small number of example locations and new modelling was produced for the failed asset sections and incorporated into the final hazard footprints.

The failed sections were identified from causes to effects by analysing the event-specific loading on each asset (>1 in 100 AEP), its SoP (>1 in 20 AEP) and condition grade (either 4 or 5). Defence height (that is, crest level – ground level) and RASP (Risk Assessment of flood and coastal defence for Strategic Planning) type information was also cross-checked for the list of potentially failed sections to minimise the chance of data errors affecting the selection process.

The selected subset was then intersected with the Office for National Statistics' Urban Settlement Area Boundaries dataset to identify defences where there were likely to be significant local impacts should breaching to occur.

Finally, each defence in the remaining subset was inspected manually within a geographical information system (GIS) alongside other contextual information to determine if it was a suitable choice for inclusion, for example, by considering local floodplain topography and the distribution and type of receptors in the immediate vicinity. After consultation with technical specialists from the Environment Agency's National Asset Management Team, it was decided to include between 5 and 10 breach locations per event scenario.

A very simple modification was applied to the original net inflow hydrograph derived in Step 2 at each asset determined to have breached based on the preceding logic (Figure 3.7). The defence is assumed to breach at the time of peak inflow and the SoP threshold is subsequently reduced over a 30-minute period from its actual value to the value assumed to represent channel capacity in undefended areas (1 in 5 AEP).

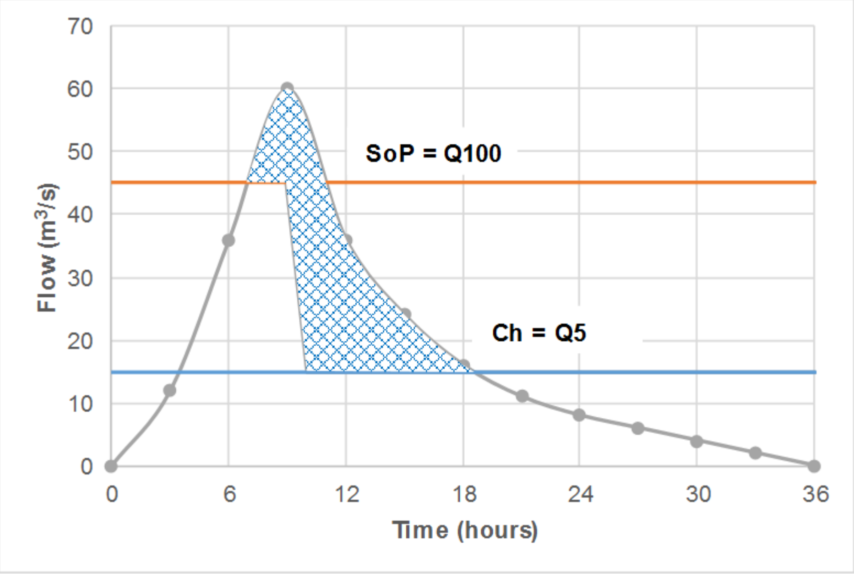


Figure 3.7 Adjustment of the net inflow hydrograph to account for breaching

Notes: Net inflow into the floodplain is denoted by the blue hatched area.

3.3 Surface water

3.3.1 Method, input data and implementation

Step 1 – Event interpolation using kriging

The approach used to produce hazard footprints for the selected surface water events is based on reusing the Updated Flood Map for Surface Water (uFMfSW) models and data. The hydraulic modelling that underpins the uFMfSW uses a total of 7,095 5km × 5km individual simulation domains (Figure 3.8b). Hence the first step requires estimation of the probability at each rainfall gauge for a given event to be interpolated across the model 'tiles'.

The point probability estimates shown in Figure 3.8a are interpolated to a 5km × 5km regular grid using kriging within ArcGIS. Interpolation was completed in the log domain. Models-of-fit and parameters used in the kriging analysis were determined for each event through analysis of the semi-variogram within ArcMap. The interpolation aimed to produce realistic event footprints by careful choice of semi-variogram parameters, in particular the range that controls the maximum distance over which an individual rainfall gauge influences the interpolated grid.

For consistency with the scenarios mapped in the uFMfSW, interpolated probabilities less than 1 in 30 AEP were set to zero and discontinued from the subsequent analysis and modelling.

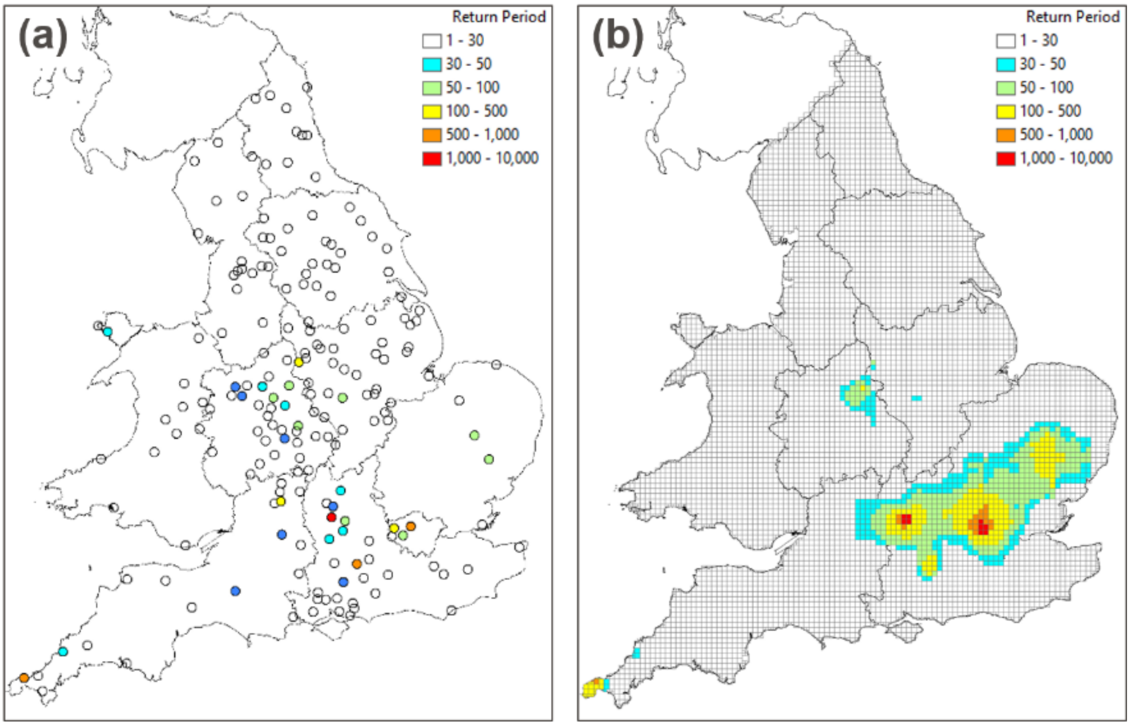


Figure 3.8 Kriging of point probability values (a) across the 7,095 5km × 5km uFMfSW modelling units (b)

Step 2 – Rainfall calculation using FEH depth–duration–frequency methods

As part of the uFMfSW project, a depth–duration–frequency (DDF) model was constructed for each 5km × 5km modelling unit using FEH methods and data. The interpolated probability (that is, frequency) and 3-hour storm duration were used as inputs to the tile-specific DDF model and a spatially varying total rainfall depth was calculated for each event scenario on a tile-by-tile basis (Figure 3.9).

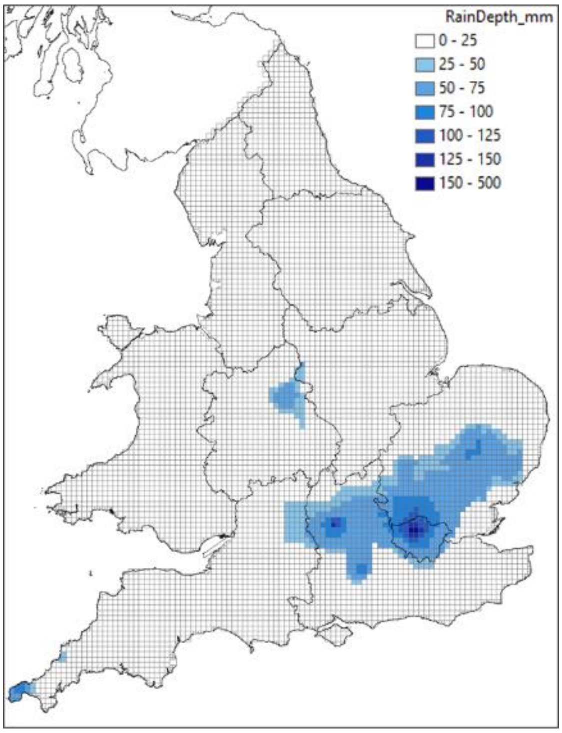


Figure 3.9 Total rainfall depths calculated using FEH DDF methods

Step 3 – 2D hydraulic modelling and flood mapping

The new rainfall inputs were applied to the 2D hydraulic models developed during the uFMfSW project (Environment Agency 2013). The models were run using the same loss assumptions, edited Digital Terrain Model (DTM) and depth-varying hydraulic roughness specification.

The models were run for 6 hours. Maps of maximum extent, depth, velocity and hazard rating were output and subsequently post-processed using the same criteria developed for the uFMfSW. Flood maps for each model tile are then combined to produce a single hazard footprint for each event; Figure 3.10 shows an example.

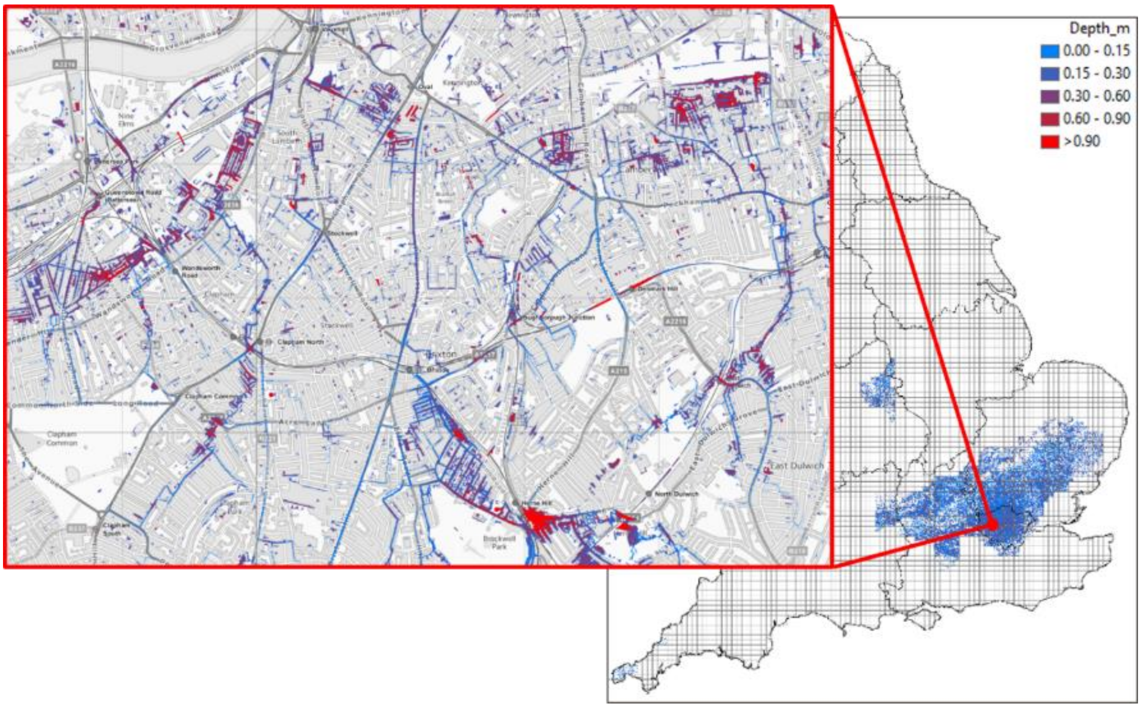


Figure 3.10 Widespread floodplain inundation associated with an extreme surface water scenario

Notes: Ordnance Survey data
© Crown copyright and database 2016

3.4 Joint fluvial–coastal

3.4.1 Method

The joint fluvial and coastal flood inundation scenario was modelled using the methods applied under the inland flooding NRA project (fluvial only, see Section 3.2) and parallel coastal flooding NRA project (Defra 2016).

The statistical dependence between the different flood sources was appropriately captured using the methods described above. However, time and resource constraints mean that the hydraulic interaction between the different sources was not explicitly simulated; this development is recommended for future consideration.

In the analysis here, the 2 sources were modelled separately and the flood depth results compared in areas where the flooding overlapped and the maximum depth was selected.

A detailed description of the coastal methods is described in the coastal methods report (Defra 2016) and only a summary is presented here. The most important stages in this analysis are:

- transforming the offshore conditions to the nearshore, taking account of processes such as wave refraction and shoaling
- assessing the potential for breaches to occur and identifying potential breach locations

- transforming the nearshore conditions into wave overtopping or overflow rates (that is, the rates of water flowing over or through the defences into the floodplain) to form the boundary conditions to the inundation modelling
- simulating the propagation of water across the floodplain using a flood inundation model

3.4.2 Input data

The SWAN wave transformation models used to propagate the offshore sea conditions to the nearshore were already available from previous Environment Agency and Natural Resources Wales projects. The locations of these models are shown in Figure 2.7.

There were over 20 breaches of coastal flood defences during the winter flooding of 2013 to 2014. It was therefore considered appropriate to implement breaches within the extreme flood scenarios. A range of data sources, including the number and defence types of breaches that occurred over the winter 2013 to 2014 were analysed. The findings of this analysis were discussed and agreed with the project partners and the breaches implemented within the flood model simulation.

Data on the location, type and geometry of coastal flood defences stored within the Environment Agency's Asset Information Management System (AIMS) database and Natural Resources Wales's Asset Management eXpert (AMX) system were used as the input data for the wave overtopping and overflow calculations.

There were some known issues in relation to the crest level of defences in some specific locations, particularly within the Environment Agency's AIMS database at that time. In these areas, crest level information was adjusted, in discussion with Environment Agency representatives, to be more appropriate based on information on extreme sea levels and knowledge of historical flooding and general SoPs.

Natural Resources Wales's AMX data were supplemented with additional information from its ongoing data collection programme to ensure the best available information was utilised.

Information on sea conditions output from the wave transformation modelling was used to provide the boundary conditions for the overtopping model. HR Wallingford's BAYONET model (Kingston et al. 2008) was used to undertake the flood defence wave overtopping calculations.

3.4.3 Implementation

For flood inundation simulation, it was necessary to transform the offshore wave conditions for each scenario to the nearshore, taking account of processes such as refraction, wave growth and breaking. The data were transformed to the nearshore to a series of points on approximately the -5m Ordnance Datum Newlyn (ODN) contour, which covered the entire coastline of England and Wales. This SWAN 2D output data were then transformed through the surfzone using the SWAN 1D model. The output then formed the input to the BAYONET wave overtopping model (Figure 3.11). The overtopping rates formed the basis of the boundary conditions for the flood inundation simulation.

The flood inundation simulations were carried out using the Caseg model. Caseg solves the shallow water equations to simulate the propagation of water over the floodplain. The model was constructed using data from the Environment Agency's

AIMS database, Natural Resources Wales's AMX system and topographical information from the Environment Agency and Natural Resources Wales's composite 2m LiDAR (light detection and ranging) dataset. Where relevant, breaches were introduced into the model assuming the initiation occurred at the peak of the hydrograph inflow. The output of the inundation model would be a time series of depth and velocity for each grid cell over the flooded area. This information was used in subsequent impact analysis.

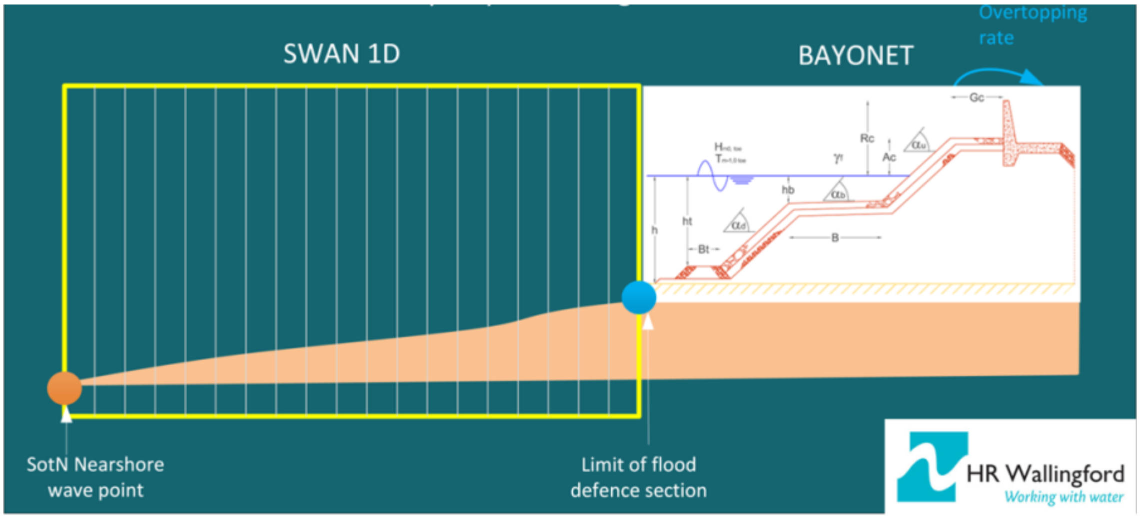


Figure 3.11 Conceptual profile diagram of the SWAN 1D and BAYONET overtopping models

Source: Defra (2016)

3.4.4 Example output

An example output from the Caseg flood inundation model is shown in Figure 3.12.

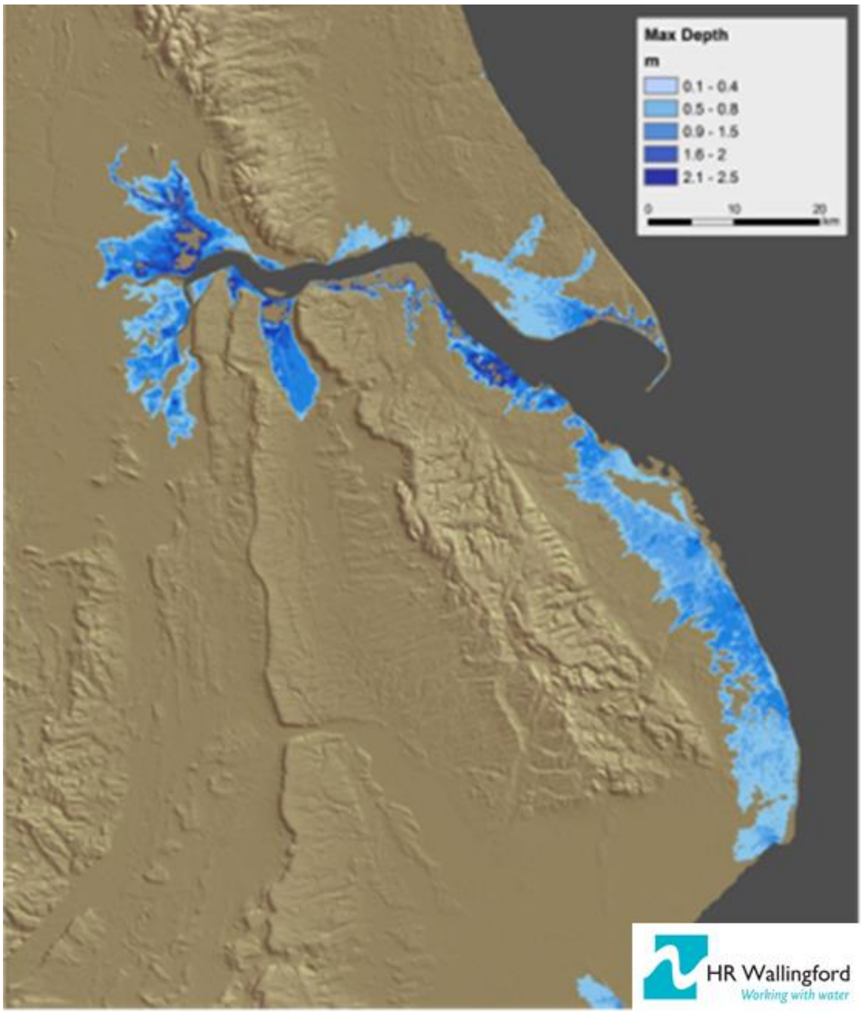


Figure 3.12 Example output from Caseg flood inundation model

3.5 Flood hazard rating

The inundation model outputs were used to provide depth and velocity data for each 50m cell across the duration of the flood. A flood hazard rating was calculated using the depth and velocity values at the time of the maximum hazard rating over the full flood scenario, with a depth-related debris coefficient. The hazard rating was then classified into categories corresponding to increasing hazard severity (Table 3.1). Hazard scores below 0.575 were removed in alignment with methods used for the uFMfSW (Environment Agency 2013).

Table 3.1 Hazard categories

Hazard rating	Degree of flood hazard	Description
0.575–0.75	Low	Caution 'Flood zone with shallow flowing water or deep standing water'
0.75–1.25	Moderate	Dangerous for some (that is, children) 'Danger: Flood zone with deep or fast flowing water'
1.25–2.00	Significant	Dangerous for most people 'Danger: Flood zone with deep fast flowing water'
>2.00	Extreme	Dangerous for all 'Extreme danger: Flood zone with deep fast flowing water'

Source: Environment Agency and HR Wallingford (2008)

4 Impact analysis

4.1 Method overview

The methodology for the impact assessment had 3 main components:

- collection and formatting of receptor datasets into a single standardised receptor database
- development of impact assessment metrics
- implementation of the impact assessment model including aggregation of results to local authority boundaries

The implementation of these steps is described below.

4.2 Input data

4.2.1 Collection of receptor datasets

Receptors are features or elements that are potentially exposed and vulnerable to the flood hazard. The receptors included in this impact assessment can be categorised into 5 groups:

- population
- property
- infrastructure
- transport
- agriculture

The best available information on these 5 categories was sourced from:

- government organisations including the Environment Agency, Department for Transport (DfT) and Health and Safety Executive (HSE)
- national data providers – Ordnance Survey (OS), Office for National Statistics (ONS)
- infrastructure asset owners

Direction was provided by previous flood risk assessment work (Environment Agency 2009a, Aldridge et al. 2011, Aldridge et al. 2015), as well as infrastructure-specific work on climate change (ITRC 2013, HR Wallingford 2014).

Population

The spatial data used for population receptors were largely derived from the National Population Database (NPD). The NPD is a GIS database providing spatially referenced estimates of population numbers for different population types and scenarios. The NPD was originally created for HSE by Staffordshire University in 2004 and has since been adapted and improved by the Health and Safety Laboratory (HSL), which continues to develop and maintain it (Smith et al. 2005, Smith and Fairburn 2008). A list of the populations included in this assessment is given in Table 4.1 Several population layers were created outside the NPD and are described below.

The census catchment of vulnerability information in Table 4.1 details the size of census boundary used to calculate the proportion of people more vulnerable to flooding for a given population theme. The different sized catchments reflect the fact that the population within different types of property (schools, workplaces, caravan parks and so on) are likely to be drawn from different sized spatial catchments depending on the population theme.

Census data are collected into hierarchical administrative boundaries to preserve anonymity and to provide appropriate data for different applications. The smallest spatial unit available is the Output Area (OA), which represents an average of 100 residents (ONS 2016). Lower Super Output Areas (LSOAs) represent an aggregation of OAs and contain an average of 1,500 residents. Middle Super Output Areas (MSOAs) represent an aggregation of LSOAs and contain an average of 7,200 residents. Local authorities represent an aggregation of MSOAs and are the largest spatial units used in this analysis.

Where population analyses are based solely on households, the most local level census information such as OA is an appropriate choice. LSOAs were used to represent wider spread daytime residential populations.

Evidence from the National Travel Survey (DfT 2014a) suggests that the average commute to an educational establishment in 2013 was 3 miles, which is equivalent to a commuter travelling across an average sized MSOA. In the same survey, the average commute to work was 8.8km. This roughly equates to a commuter travelling halfway across an average sized local authority. Local authorities were also used to represent larger, regional catchments, including stadium populations.

The workforce layer was derived from the NPD and the National Receptors Dataset (NRD) – described in the Property section – including the NPD's temporal scenarios, which model daytime, night-time and weekend employment levels.

The location and population of camping/caravan sites and other leisure accommodation were produced specifically for this project. Locations were derived from OS AddressBase Premium, Valuation Office Agency (VOA) and Camping and Caravan Club data. Campsite populations were derived from Camping and Caravan Club data and online campsite directories. Other leisure accommodation populations were derived from bed space information contained in VOA data.

Table 4.1 Impacted populations aggregated from flood impact data into geographical units

Population theme	Breakdown	Data used	Census catchment of vulnerability
Residential	Night-time	NPD	OA
	Daytime (term time)	NPD	LSOA
	Daytime (non-term time)	NPD	OA
Sensitive	Schools	NPD	MSOA
	Colleges	NPD	MSOA
	Care homes	NPD	100% vulnerability
	Childcare facilities	NPD	MSOA
	Hospitals	NPD	100% vulnerability
	Prisons	NPD	LAD
Working population	Weekday workers	NPD/NRD	LAD
	Saturday workers	NPD/NRD	LAD
	Sunday workers	NPD/NRD	LAD
Leisure	Caravan/camping sites (peak/low season)	AddressBase Premium / VOA	National average
	Other tourist accommodation (peak/low season)	AddressBase Premium / VOA	National average
	Stadiums (capacity)	NPD	LAD

Notes: LAD = local authority district

Property

The NRD property point data formed the basis of the receptor database for this work (Table 4.2). The NRD was created by the Environment Agency and Natural Resources Wales for flood risk assessment (Environment Agency 2011b). The NRD, which is based on OS datasets, aims to locate and attribute all properties in England and Wales that are addressable or have a floor level footprint larger than 25m^2. Attributes include:

- residential type
- non-residential usage categories
- building footprint size
- indicators for the lowest floor of the property
- unique reference identification codes such as the Unique Property Reference Number (UPRN) (Geoplace 2016)

This provides the basic information required for estimation for property impact analysis. The NRD property point dataset was filtered to remove points that did not represent buildings (for example, advertising hoardings and telephone boxes) and properties recorded as being above ground floor. Listed buildings are not explicitly included in the

NRD, but are considered important sites in case of a flood. Therefore, the locations of listed buildings were acquired from Historic England and Cadw.

Table 4.2 Impacted property types and source datasets

Property type	Source
Residential (detached, semi-detached, terraced, flats)	NRD property point dataset
Shop/store	
Vehicle services	
Retail services	
Office	
Distribution/logistics	
Leisure	
Sport	
Public building	
Industry	
Miscellaneous	
Unclassified	
Listed buildings	Listed buildings in Wales GIS point dataset (Cadw) Listed Buildings in England GIS point dataset (Historic England)

Infrastructure

Infrastructure types and datasets used are listed in Table 4.3. The majority of the infrastructure sites are located in properties and were therefore represented as points. Roads and railways were represented as lines. Individual infrastructure types were grouped into broad infrastructure categories as detailed in Table 4.3. Infrastructure categories are listed below:

- **Emergency services** are the emergency response providers. This includes police, ambulance, fire and coastguards. These features are important as the effectiveness of their response to the consequences of flooding may be adversely affected by the flood hazard itself.
- **Key sites** are identified as core public sites that either provide essential services or might create significant societal problems if disrupted by flooding. As such, there is a priority for the sites to be open and accessible. Key sites include hospitals, schools, doctor's surgeries, care homes and prisons.
- **Utilities** provide important services for water or provision of energy. Major outages of power or water are already present in the NRA as separate risks in their own right, but these are still significant features in flood impact assessment. They include water treatment works, sewage pumping stations, power stations and electrical installations.

- **Potentially hazardous** sites are locations that have the potential to cause further harm if disrupted by flooding. This may be through diffusion of waste or pollutants into the environment or through emission of dangerous substances into the atmosphere. Such instances could have serious consequences for danger to life and the environment. These sites include major hazard sites and industrial sites that produce radioactive or waste materials that require specific licences and management.
- **Transport** infrastructure includes the road and railway network as well as transport hubs such as bus and train stations. Disruption of the transport network could have serious short-term consequences during a flood when evacuation routes or emergency service routes require diversion. In the longer term, impacts on the transport infrastructure may result in increased traffic and longer journey times with consequences for services, society, costs relating to lost working hours and other indirect business costs.

Road networks were populated with estimates of vehicle and passenger numbers. Average daily flow data from the DfT provided information on the type, number and average speed of different types of vehicle passing along each node-to-node segment of major road in a 24-hour period. The length of the road segment in kilometres was multiplied by the number of vehicles/passengers/lorries on that segment to produce metrics for passenger km, vehicle km and lorry km (ITRC 2013). Larger values indicate busier and more important road segments. The vehicle km provides a measure of how busy a road is, the passenger km populates road segments with people and the lorry km provides a proxy measure for commercial traffic.

Table 4.3 Infrastructure types under consideration and data sources

Infrastructure category	Infrastructure type	Data source
Emergency services	Fire stations Ambulance stations Police stations Coastguard facilities	OS AddressBase Premium / VOA
Key sites	Doctor's surgeries and health centres	Care Quality Commission GP practice data
	Hospitals Care homes Schools Prisons	NPD
Transport	Bus stations	NPD
	Train stations	NPD / National Rail station data
	Roads (including primary/trunk roads)	NRD Roads
	Railway (km)	NRD Railways
	Ports	DfT Transport Statistics PORT0101
Utilities	Electrical substations	OS AddressBase Premium / National Grid

Infrastructure category	Infrastructure type	Data source
	Large electrical substations (>100m^2) Power stations	DECC DUKES 5.10 dataset
	Nuclear sites	HSE library
	Waste water treatment works Sewage pumping stations Water treatment works (clean water)	Environment Agency/Natural Resources Wales consented discharge to controlled waters
	Petrol stations	OS AddressBase Premium / VOA
Potentially hazardous	Major hazard sites	HSE library
	Waste and recycling facilities	Environment Agency (Environmental Permitting Regulations – waste sites)
	Industrial Installations (covered by IPPC directive)	Environment Agency (Environmental Permitting Regulations – industrial sites)
	Sites handling radioactive substances (RAS authorities and registrations)	Environment Agency (Radioactive Substances Register 2011)

Notes: DUKES = Digest of UK Energy Statistics; IPPC = Integrated Prevention and Pollution Control; RAS = radioactive substances

Core infrastructure components

In the previous coastal flooding assessment (Aldridge et al. 2015), HSL was asked by stakeholders to incorporate the influence that core infrastructure components may have on the entire infrastructure network. To highlight the importance of these core sites, agreed infrastructure types were filtered by site size or capacity to identify more significant sites in the infrastructure networks. These were as follows.

- **Railway stations**. The NPD railway stations layer was enriched with station category data from National Rail. Categories A and B represent national hubs and national interchanges, and were selected to represent major railway stations.
- **Electrical substations**. The base substation layer derived from OS AddressBase Premium was joined with National Grid data, which provides data on the largest substations in the national network (Supergrid and bulk substations). These substations transfer energy cross-country to smaller, local substations. The largest Supergrid substations (400kV) were chosen to represent significant substations in the network.
- **Power stations**. Although all energy generation sites are important, this project considered sites that generate above 1,000MW to be large sites.

This followed work conducted for the Committee on Climate Change (HR Wallingford 2014) and was completed using the DUKES database.[2]

- **Waste water treatment works (WWTW)**. Following the work for the Committee for Climate Change (HR Wallingford 2014), the Environment Agency/Natural Resources Wales controlled discharge to consented waters dataset was used to subset large WWTW sites, taking sites that process more than 30,000 cumecs (cubic metres per second as a unit of flow of water) dry weather flow as a threshold.
- **Major hazard sites**. UK major hazard sites are regulated by HSE, the Environment Agency, Scottish Environmental Protection Agency (SEPA) and Natural Resources Wales under the European Seveso directives. The UK implements these directives through the Control of Major Accident Hazards (COMAH) regulations, which includes categorisation of sites by the type and volume of hazardous substances stored, and the methods of storage. Major hazard sites identified as 'top tier' under the Control of Major Accident Hazards (COMAH) Regulations were selected to represent large-scale sites.
- **Roads**. Trunk roads and motorways are routes of strategic importance in the road transport network. For this research, classification data in the NRD roads layer were used to identify trunk roads and motorways as important routes.

Agriculture

Flooding of agricultural land can cause damages with severe consequences for both arable and pasture farming (Penning-Rowsell et al. 2013). Agricultural land data were taken from the Agricultural Land Classification (ALC) included in the 2010 version of the NRD. The layer itself was created in 1988 and remains the most recent version available. The ALC covers England and Wales, and separates the entire landscape into 5 grades of agricultural land (from 1 highest value to 5 lowest value), urban and non-agricultural land. For this research, only graded agricultural land was used. To highlight damage to the highest quality land, grades 1 and 2 were also included separately as an additional impact metric.

Reporting areas

Stakeholder discussion identified that reporting flood hazard impacts at administrative boundaries will provide summary information that is easier to disseminate and more relevant to local planners and decision-makers.

The update to the East Coast flooding impact analysis (Aldridge et al. 2015) made use of Local Resilience Forum (LRF) areas and their composite local authority boundaries. As a requirement of the Civil Contingencies Act 2004, LRFs are multi-agency partnerships created to plan and prepare for localised incidents and catastrophic emergencies (Cabinet Office 2013). LRFs are composed of front line Category 1 responders; this includes local emergency services, local authorities, the National Health Service, the Environment Agency and others. LRF boundaries align with local police force areas for easier management of local emergencies.

[2] https://www.gov.uk/government/collections/digest-of-uk-energy-statistics-dukes

Use of LRFs can potentially promote co-operation between neighbouring LRFs when flood impacts cross boundaries. A list of LRF areas and their component local authorities is provided in Appendix C of the Results Report.

4.3 Development of impact assessment metrics

4.3.1 Population

Following previous studies (Aldridge et al. 2011, Aldridge et al. 2015), the population impacts of the flood scenarios were based on the Flood Risks to People (FRTP) methodology, implemented as outlined in the FRTP Phase II guidance document (HR Wallingford et al. 2006) and supplementary note (Environment Agency and HR Wallingford 2008). As well as information on flood hazard intensity (depth, velocity, debris), FRTP was created in a UK context and takes into account receptor-specific factors including personal and physical vulnerability to flooding, and vulnerability associated with local influences. An additional requirement for estimating evacuation was also addressed.

In this analysis, population impacts were presented as counts of:

- people within the flood area
- people who are more vulnerable to the flood hazard (calculated as a proportion of the total impacted population)
- people who are at risk of injury
- potential fatalities[3]
- people requiring evacuation, including those requiring assistance, or identified as needing a priority evacuation response (within 24 hours)

Table 4.4 presents the impact metrics selected for each of the population impact calculations. These metrics were calculated for each of the populations listed. In all cases, people are considered to be impacted by flooding if they are located inside the flood extent at locations where flood hazard reached 0.575 or higher. This is in line with outcomes of Defra capacity building workshops for the uFMfSW and based on agreement with project partners.

Table 4.4 Population impact metrics

Impact name	Impact metric
Impacted population	Count of all population that are located within flood extent (minimum flood hazard rating = 0.575, minimum depth = 0.005m)
Impacted vulnerable population	Count of impacted population identified as more vulnerable to flooding
Injuries	Count of injuries sustained Based on FRTP and applied using Equation 4.3 (see Section 4.3.4)

[3] The injuries and fatalities metrics presented in this research are better considered as the extent to which contributing physical factors combine to present a danger to life or of injury (see below).

Fatalities	Count of fatalities sustained Based on FRTP and applied using Equation 4.4 (see Section 4.3.4)
Evacuation and priority evacuation	Count of people requiring priority evacuation Based on statistics from the winter 2013 to 2014 flood review (Environment Agency 2016) and a count of vulnerable people impacted

The number of people that might be more vulnerable to flooding is an important statistic, which can help with prioritising flood management action. Vulnerability is defined as the characteristics and circumstances of a community, system or asset that make it susceptible to the damaging effects of a hazard (UNISDR 2009).

This research followed the approach for vulnerability described in the FRTP methodology (HR Wallingford 2006, Environment Agency and HR Wallingford 2008). Application of the FRTP methodology requires measurements for 2 types of vulnerability:

- people vulnerability (see Section 4.3.2)
- area vulnerability (see Section 4.3.3)

4.3.2 People vulnerability

The FRTP methodology defines people vulnerability as the ability of those affected to respond effectively to flooding. Two population groups are considered to be vulnerable to flooding, based on physical attributes:

- people suffering from limiting long-term illness
- people aged 75 or over

These 2 populations were calculated using the 2011 Census tables listed in Table 4.5. The variables considered were:

- age
- long-term limiting illness
- economic activity

These variables provide information on whether an individual should be deemed vulnerable and whether they are likely to be at home or at another location during the day, for assessment of daytime population scenarios.

The geographical units reflect the size of the catchment that the population is drawn from (see discussion in Section 4.2.1). Populations in hospitals and care homes were attributed 100% vulnerability based on the assumption that all patients/residents would be elderly or ill, and therefore less mobile and more vulnerable to the physical effects of flooding. The full breakdown of the specific vulnerability and the calculations for each population type are detailed in Appendix B.

Table 4.5 Census data used for people vulnerability calculation

Census table code	Census table name	Geographical units
KS102EW	Age structure	OA, LSOA, local authority
QS601EW	Economic activity	OA, LSOA

L3302EW	Long-term health problem or disability by general health by sex and age	OA, LSOA, local authority
LC3101EWLS	Long-term health problem or disability by sex and age	LSOA
LC6302EW	Economic activity by hours worked by long-term health problem or disability	LSOA

4.3.3 Area vulnerability

FRTP describes area vulnerability as the characteristics of an area that affect the chance of people in the floodplain being exposed to the hazard. The area vulnerability is composed of 3 components:

- scope and effectiveness of Environment Agency flood warnings
- speed of flood onset
- nature of area with regard to the physical characteristics of individual receptor locations

In FRTP, these 3 components are each assigned scores between 1 and 3. These scores are summed together using Equation 4.1 to provide a score (AV) of between 3 and 9, where 3 indicates the areas least vulnerable to flooding and 9 represents areas most vulnerable to flood impacts:

Equation 4.1

$$\text{AV} = \text{flood warning score} + \text{speed of onset score} + \text{nature of area score}$$

Flood warning score

The first component, the flood warning score, uses 3 Environment Agency flood warning measures that are based on 3 key targets in relation to flood warning:

- P1 – percentage of Warning Coverage Target met (percentage of at risk properties covered by flood warning system) (target 80%)
- P2 – percentage of Warning Time Target met (target 100%)
- P3 – percentage of Effective Action Target Met (percentage of people taking effective action) (target 66%)

The P3 target is based on the finding of the Public Flood Survey of the 2013 to 2014 flood warnings (795 post flood interviews) when 66% of people who received a warning took action.

The flood warning score is calculated using Equation 4.2:

Equation 4.2

$$\text{flood warning score} = 3 - (\text{P1} \times (\text{P2} + \text{P3}))$$

Equation 4.2 produces scores from 1 to 3, where a value of 3 indicates a weak/no flood warning and 1 indicates a good flood warning and action system.

Speed of onset score

The second component is the speed of onset score. A value of 1 indicates an onset of several hours, while 3 indicates flash flooding occurring in minutes.

For this report, the speed of onset score was based on the source of flooding. For coastal and fluvial flooding, a value of 2 was used indicating a flood onset of approximately one hour. For surface water flooding, a value of 3 was used indicating a fast onset with little time for preparation of response.

Nature of area score

The final component is the nature of area score. This component describes the vulnerability of a receptor in terms of the physical attributes of its location. For example, multi-storey buildings are considered to be less vulnerable because residents are more likely to live in higher storeys, while bungalows and campsites are considered more vulnerable.

This report follows the update to the East Coast flooding impact analysis report (Aldridge et al. 2015) and the FRTP methodology as detailed in Table 4.6. All other populations were given a nature of area score of 2.

Table 4.6 Nature of area vulnerability modelling descriptions

Population type	How modelled	Nature of area score
Populations near busy roads	Primary routes and trunk roads were extracted from the NRD roads layer (based on OS Integrated Transport Network data). Population locations within 40m of these were allocated as high risk.	3
Multi-storey apartments	Residential buildings within the area of interest with more than 10 households present were looked at with OS MasterMap data to gauge whether they might be described as multi-storey. Those that fit were classified as low risk.	1
Campsites	Population locations within campsites were determined as high risk.	3
Single storey schools	Schools for young children or for those requiring special care were considered to have a high likelihood of being single storey, and so were classified as more vulnerable. Classifications taken from the NPD and NRD for infant, junior, primary and special schools were used to set the nature of area risk for schools meeting this description. Although this is not always the case, it is a reasonable assumption.	3
Road populations	Those in cars classified as high risk.	3

4.3.4 Injuries and fatalities

Following the FRTP methodology, the number of injuries and fatalities sustained in a flood are calculated as functions of the vulnerable population. The number of injuries is calculated using the number of vulnerable people at a location, the area vulnerability and the flood hazard rating at that location (Equation 4.3).

Equation 4.3

$$Ninj = 2 \times Nz \times HR\left(\frac{AV}{100}\right) \times PV$$

where:

Ninj is the number of injuries

Nz is the number of people at risk

HR is the hazard rating

AV is the area vulnerability

PV is the people vulnerability (proportion of vulnerable people).

The number of fatalities (*Nf*) is calculated as detailed in Equation 4.4:

Equation 4.4

$$\mathrm{Nf} = 2 * \mathrm{Ninj}\left(\frac{\mathrm{HR}}{100}\right)$$

Population impact metrics were calculated at the level of individual receptor point and then aggregated to reporting level.

Injury and fatality estimates should be treated with caution. The multi-dimensional nature of the impacts of flooding on people presents a high level of uncertainty. Complicating factors relate to the nature of the hazard and the behaviour of the receptor. For example, the nature of the flooding within the cell and the effect of features in urban areas could alter localised flood depths and velocities, while the assumed location and behaviour of people within and around a flooded property could change how they are affected. Until these factors can be better understood and accounted for, the metrics used here were considered as the extent to which contributing physical factors combine to present a danger to life or of injury.

Evacuees

The number of affected residents (including those residing in hospitals, prisons and care homes) provided a baseline of the population who may require evacuation. Those requiring assistance with evacuation can be estimated based on the vulnerable population.

Priority evacuees represent a further subset of the vulnerable population, who may not be easily identified using the NPD or Census tables. These are people who require the most urgent evacuation assistance. Priority evacuees were identified under the assumption that they would have an immediate health requirement that presents a risk to well-being if care is not available at short notice. People in care homes and hospitals would be in this group.

In addition, the UK Homecare Association (UKHCA) estimates that approximately 512,000 people received state-funded domiciliary care (that is, care in the home) in England and Wales in 2013 to 2014, with a further 228,000 people receiving care that is privately funded in the UK (Holmes 2016).

Adjusting the UK private estimate based on the state-funded figure produces a total of 690,000 receiving some form of domiciliary care in England and Wales. This represents 1.2% of the total population according to ONS mid-year population estimates. This percentage was used in addition to the numbers of residents in care homes and hospitals to estimate priority evacuees.

4.3.5 Property

In all scenarios, the property point datasets were overlaid with the flood hazard extent dataset. However, different approaches were adopted to derive standardised impact results for different flooding sources.

For coastal and fluvial flooding, a property was considered flooded if the property point was located within the flood hazard extent using a base flood depth threshold of 0.00m.[4] This aligns with the NaFRA approach.

The method adopted by the Environment Agency for counting properties flooded due to surface water flooding is different from that used for NaFRA. This is due to differences in the nature of the flooding (Environment Agency 2013) and the method of modelling inundation due to surface water flooding (Section 3.3), which is not directly compatible with the intersection of property points and flood hazard used for NaFRA. The surface water flooding property count method was developed specifically for the uFMfSW by Horritt Consulting (2013) and requires an assessment of flooding for individual building footprints. This is a computationally intensive task which was performed with the release of the uFMfSW dataset, but is beyond the scope of this project. Instead, the NaFRA method was used to estimate initial counts of flooded properties. Property counts were then scaled up by a factor of 1.8 based on a comparison of counts of flooded properties generated by the NaFRA method and the specific surface water flooding method for England based on the uFMfSW dataset.

Property damage was estimated by considering depth-related impacts for individual properties and aggregating to reporting areas. Damage calculations for different property uses and types (Table 4.2) were based on flood depth information and damage curve calculations published in the Multi-Coloured Manual (MCM) (Penning-Rowsell et al. 2013).

The MCM is an established and comprehensive framework for assessing the economic impacts of flooding. It offers flood damage information for a range of different flood types (salt/fresh water, short, medium, long durations) and includes a breakdown of cost components including domestic clean-up, household inventory damage and building fabric damage. These values are provided in the MCM in the form of depth–damage curves for different types of property at a range of flood depths, with the component costs summarised to produce total damage and total damage per square metre for a given depth. The following depth–damage curves were used for each type of flooding:

- Coastal – long duration, salt water
- Fluvial – long duration, fluvial water
- Surface water flooding – short duration, fluvial water

MCM damage calculations require information on:

- building use and footprint – which can be found in the NRD
- speed of flood onset
- nature of the flood water (fresh water and salt water)

Different curves were applied for different property uses and types. For each affected property, a value of damage per m^2 was calculated from the depth–damage curves based on the flood depth and property type. This was then multiplied by the footprint of

[4] This was implemented as a depth threshold of 0.005m to eliminate artefacts in the flood hazard data which modelled very low depths over large areas.

the building (m^2) to provide an estimate of the damage in pounds. An example is shown in Figure 4.1, which demonstrates the damage calculation for a $300m^2$ retail property flooded to depth of 1.5m.

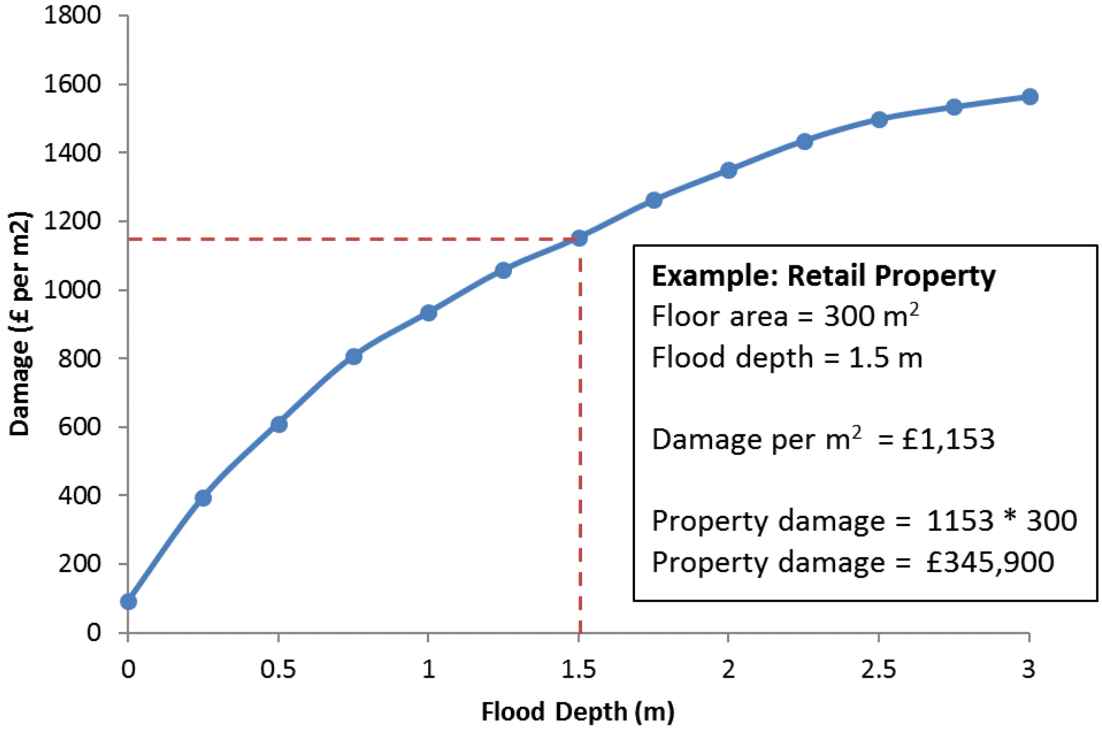


Figure 4.1 Example property damage calculation for a $300m^2$ retail property flooded to a depth of 1.5m

Flood warnings have been shown to reduce the damage to property due to the opportunity presented to take action, primarily by protecting or moving personal possessions, stock or moveable equipment (Penning-Rowsell et al. 2013). The MCM estimates the reduction in costs as a proportion of the total damages. For residential properties, the cost in damages can be calculated for a warning of less than 8 hours and a warning of over 8 hours. For non-residential properties, the damages can be calculated for a warning time of 4 hours. This can provide useful best and worst case scenarios.

Where specific building type information was not recorded, the curves for the 'average' category were used for residential properties and the 'unclassified' category for non-residential properties. To model the damage to properties below entrance thresholds (doorsteps) accurately, the modelled flood depth was reduced by 0.25m when calculating property damage following the NaFRA approach.

For surface water flooding, an adjustment of property damages was required to account for the damages to properties added by the uplift factor for surface water flooding flooded property counts. These additional properties represent flooding below the threshold of the building. To estimate the value of these damages effectively, representative damage figures for properties with flooding below the 0.25m threshold were applied based on mean average damages from one of the fluvial scenarios. This was done independently for each property classification.

4.3.6 Infrastructure

Disrupted infrastructure metrics were based on the potential exposure to different levels of flood hazard ratings (Table 3.1). As for population, flood hazard ratings less than 0.575 were ignored. Metrics were produced as counts of flooded sites and the percentage of that infrastructure type flooded in the local authority. These metrics provide information on the absolute magnitude of impact and indicate the pressure on local resilience and contingency. This was completed for all infrastructure assets listed in Table 4.3.

To evaluate the disruption of key sites (which are typically buildings requiring access by the public), additional metrics were calculated to count the number of key sites inundated to a depth greater than 0.2m. This depth corresponds to Environment Agency advice on sandbag usage (Environment Agency 2009b), assumed here as a minimum level of protection that might be expected at these sites.

Road and rail networks are considered to be affected if they are inundated to a depth of 0.15m or greater, based on typical vehicle ground clearance[5] as an indicator for roads becoming impassable. Disrupted journeys are evaluated based on impacts on the mean average total kilometres travelled each day by vehicles, passengers and lorries – as indicators of domestic and commercial traffic. Additional metrics for trunk roads, including motorways, are reported as a subset of the full transport network.

4.3.7 Agriculture

The MCM provides a damage cost per hectare for each of the 5 agricultural land grades. Grade 3 land is given 2 values depending on the proportions of livestock and arable crops grown on the land. Consequently, an average of the 2 values was taken (Table 4.7) for this analysis.

Table 4.7 Cost of agricultural land by grade

ALC class	Indicative land use	Flood cost (£ per ha)	
1	Intensive arable (100%)	1,320	
2	Intensive arable (60%) Extensive arable (35%) Horticulture (5%)	1,000	
3a	Extensive arable (70%) Intensive arable (30%)	600	Mean = 470
3b	Extensive arable (50%) Intensive grass (50%)	340	
4	Intensive grass (100%)	180	
5	Extensive grass (100%)	100	

Notes: From 1 highest quality to 5 lowest quality
Adapted from Penning-Rowsell et al. (2013)

[5] http://www.autoevolution.com

The method of measuring short-term and long-term impact to agriculture follows the NaFRA methods by calculating:

- area impacted (above 0.00m of flooding)
- cost in damages where flood depths exceed 0.5m

The impact and damage to Grade 1 and 2 agricultural land were included as a separate impact metric. Estimates do not specifically account for damage costs associated with salinity.

4.3.8 Wider economic impacts

The wider economic costs use a selection of the impact metrics described above as inputs to economic calculations. Economic costs to tourism and the environment have not been included. Table 4.8 presents the calculation for each of the economic metrics, which ultimately are summed together for the entire hazard.

Table 4.8 Calculations for wider economic cost impacts

Wider economic cost	Calculation	Notes
Fatalities and casualties	(Worst case of night-time or daytime fatalities × £1,836,054) + (worst case of night-time or daytime injuries × £80,690)	Values relate to the cost of a fatality and the cost of an average injury based on DfT estimates using a 'willingness-to-pay' approach (DfT 2014b, HSE 2011).
Lost assets	Total of property damage (with warning)	Using MCM approach as described above.
Lost working hours – employment impacts	Daytime working population × £11.61 × flood duration (hours)	£11.61 is the median hourly pay (ONS 2014). Flood duration is assumed to be 15 hours (2 days) based on discussion within project team.
Lost working hours – transport impacts (commuting and business trips)	[Total journey time (hours) per local authority / percentage of flooded transport network] × £11.61	Total journey time is calculated using DfT statistics on average journey times and multiplying up to the local authority population.
Shelter – short term	Night-time impacted population × flood duration × £35	Flood duration is assumed to be 15 hours (2 days). £35 is average cost of short-term accommodation per person per night (Department of Communities and Local Government, judgement-based figure based on expert consultation).
Shelter – long term	Impacted residential properties × 0.46 × £10,345	0.46 is the proportion of impacted properties likely to require extensive repair work. £10,345 is the average per property cost for this relocation – from Environment Agency

Wider economic cost	Calculation	Notes
		review of the 2013 to 2014 flood impacts (Environment Agency 2016).

Notes: DCLG = Department for Communities and Local Government
Source: Cabinet Office

4.4 Implementation/application

The assessment of flood impacts was separated into 5 steps as demonstrated in Figure 4.2.

- **Step 1** intersects flood hazard data supplied by HR Wallingford with the receptor database. It identifies the receptors potentially at risk from the flood and assigns flood attributes (depth, velocity and hazard rating).
- **Step 2** uses the attributed receptor information along with auxiliary lookup tables to calculate specific impact information for each receptor using the methods described in Section 4.3 as summaries of danger to life, economic damage to property, disruption of infrastructure and agriculture impacts.
- **Step 3** aggregates individual impacted receptors into regional summaries by local authority and LRFs.
- **Step 4** collects appropriate aggregate impact metrics to calculate wider economic costs.
- **Step 5** compiles the impacts and provides an overview of results for each scenario.

For the joint inland–coastal scenario, Steps 1 and 2 were applied separately for the different sources of flooding. The separation was based on the dominant source of flooding, determined by the highest flood depth. The results were combined prior to Step 3.

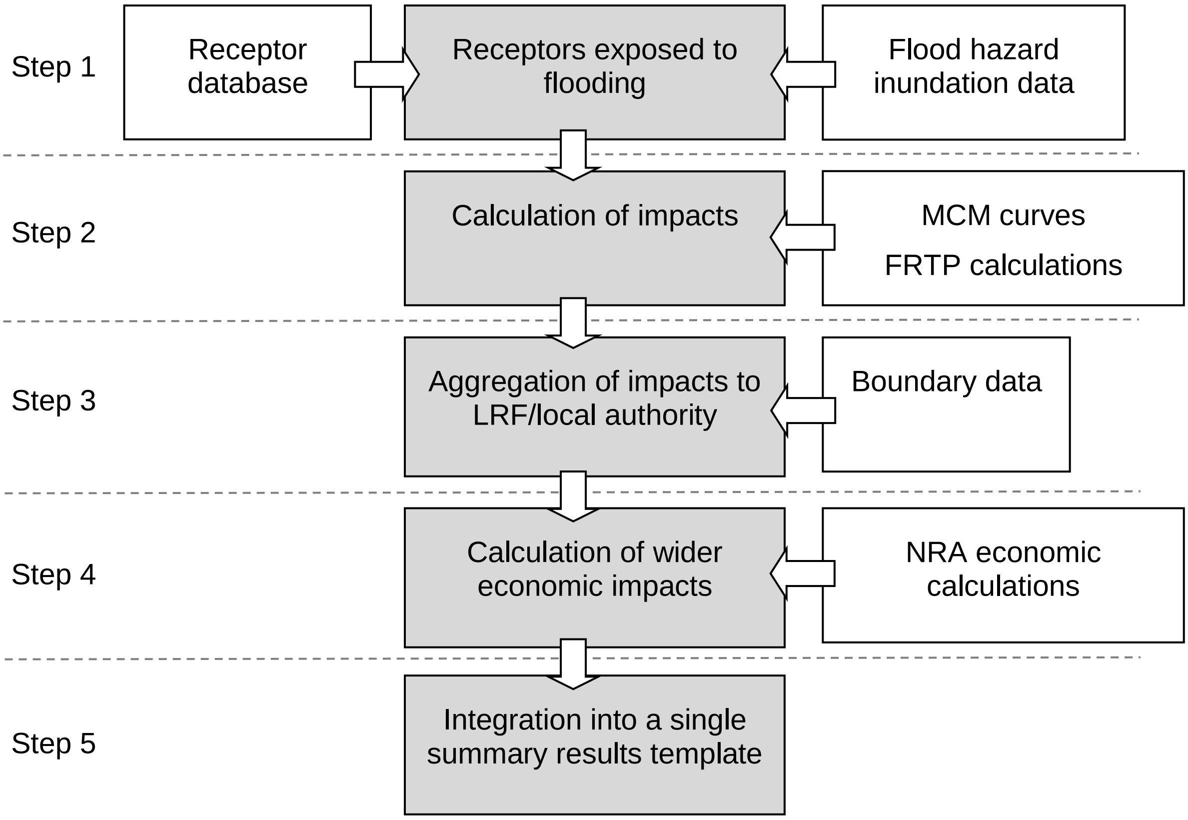


Figure 4.2 Impact assessment implementation

The process was largely automated using the statistical software package R. R is open source statistical software capable of efficiently managing and manipulating multiple large datasets.

4.4.1 Quality assurance

Quality assurance was completed throughout the impact assessment task.

1. Each receptor dataset was quality checked against established secondary sources or OS base mapping. This involved manual verification of a random sample of points to check that the coverage, location, function and other attributes of the data were correct.
2. Automated impact assessment metrics, including MCM methods were checked against manual methods to confirm that the processes were correct.
3. Visual checks and spot checks were performed on the final outputs to:
    - capture extreme values and assess their sensitivity
    - ensure that calculations had been processed correctly
    - ensure that the aggregation of results into local authority boundaries was completed correctly

5 Conclusions

This report set out the process and methodology of generating and selecting reasonable worst case scenarios to inform the evidence base for the 2016 NRA inland flooding risks.

Historical flow and rainfall data were carefully collated and a sophisticated statistical joint probability model used to generate a large number of statistically plausible reasonable worst case events for fluvial, surface water and joint inland–coastal flooding.

The joint probability method developed by Lamb et al. (2010) and Keef et al. (2013) was used to generate widespread, spatially coherent extreme events based on statistical modelling of observed data. This application of data-driven modelling improves the science underpinning the assessment of impacts, distribution and likelihoods for scenarios informing the evidence base for the NRA. The choice of scenarios that are considered plausible is guided by the improved method and each scenario has an associated narrative based on physical meteorological reasoning. The resulting scenarios are extreme and stretching compared with recent historical precedents, yet are also considered plausible and illustrative of potential extreme flooding at regional and national scale.

Hazard footprints were created using 2D hydrodynamic modelling to calculate the consequence of flooding for each scenario. Using the hazard footprints, receptor data were collated and impacts analysis performed to generate impact assessment metrics which provide an indication of the flood severity presented in each scenario.

The methodology described this report enabled the development of an improved evidence base for the 2016 NRA, as the underlying methods deliver more realistic spatial patterns of extreme yet plausible flooding across England and Wales. The impact assessment used is an improved approach and provides a more comprehensive overview of the potential impacts of widespread flooding from different sources.

References

ALDRIDGE, T., FORDER, K. AND BALMFORTH, H., 2011. *East Coast flooding impact analysis.* Health and Safety Laboratory Reference MSU/2011/43. Unpublished report.

ALDRIDGE, T., GUNAWAN, O. AND RASTALL, P., 2015. *East Coast flooding impact analysis update*, Summary report, Health and Safety Laboratory Reference MSU/2015/03. Unpublished report.

ALLAN, R.P., 2015. *Scientific knowledge of meteorological drivers of widespread flooding*, Report to JBA and the Environment Agency. Unpublished report.

CABINET OFFICE, 2013. *National Risk Register of Civil Emergencies, 2013 edition*. London: Cabinet Office.

CABINET OFFICE, 2015a. *National Risk Register of Civil Emergencies, 2015 edition*. London: Cabinet Office.

CRAN, 2013. CRAN – Package texmex. Available from: https://cran.r-project.org/web/packages/texmex/index.html [Accessed 16 August 2016].

DEFRA, 2016. *National Risk Assessment: Extreme coastal flooding scenarios. Methodology report*. London: Department for Environment, Food and Rural Affairs.

DfT, 2014a. *National Travel Survey: England 2013*. Statistical release. London: Department for Transport.

DfT, 2014b. *TAG UNIT A4.1 social impact appraisal*. London: Department for Transport.

ENVIRONMENT AGENCY, 2008. *Improving the FEH statistical procedures for flood frequency estimation*. Science Report SC050050. Bristol: Environment Agency.

ENVIRONMENT AGENCY, 2009a. *Flooding in England: a national assessment of flood risk.* Bristol: Environment Agency.

ENVIRONMENT AGENCY, 2009b. *Sandbags: how to use them properly for flood protection.* Bristol: Environment Agency.

ENVIRONMENT AGENCY, 2011a. *The risk of widespread flooding – capturing spatial patterns in flood risk from rivers and coasts*. *Spatial Coherence of Flood Risk – Technical Methodology Report.* SC060088/R1. Bristol: Environment Agency.

ENVIRONMENT AGENCY, 2011b. *National Receptor Dataset.* Unpublished.

ENVIRONMENT AGENCY, 2013. *The updated Flood Map for Surface Water (uFMfSW) property points dataset*. Report version 1.0. Bristol: Environment Agency.

ENVIRONMENT AGENCY, 2015. *Spatial joint probability for FCRM and National Risk Assessment – user requirements summary report.* SC140002/D1. Bristol: Environment Agency.

ENVIRONMENT AGENCY, 2016. *The costs and impacts of the winter 2013 to 2014 floods.* Report SC140025/R1. Bristol: Environment Agency.

ENVIRONMENT AGENCY AND HR WALLINGFORD, 2008. *Supplementary note on flood hazard ratings and thresholds for development planning and control purpose – Clarification of the Table 13.1 of FD2320/TR2 and Figure 3.2 of FD2321/TR1*. Bristol: Environment Agency.

GEOPLACE, 2016. *The power of the UPRN* [online]. Available from: https://www.geoplace.co.uk/addresses/uprn [Accessed 4 November 2016].

HEFFERNAN, J.E. AND TAWN, J.A., 2004. A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66 (3), 497-546.

HOLMES, J., 2016. *An overview of the domiciliary care market in the UK*. Wallington, Surrey: UK Homecare Association Ltd.

HORRIT CONSULTING, 2013. *Recalibration of FFC surface water flooding decision support tool with UKV data*. Flood Forecasting Centre Report WHR1328 v2.0.

HR WALLINGFORD, MIDDLESEX UNIVERSITY FLOOD HAZARD RESEARCH CENTRE and RISK & POLICY ANALYSTS LTD, 2006. *Flood Risks to People Phase 2: the risks to people methodology*. Defra and Environment Agency Report FD2321/TR1. London: Department for Environment, Food and Rural Affairs.

HR WALLINGFORD, 2014. *Indicators to assess the resilience of infrastructure in England to the projected impacts of climate change*. Reference MCR5195-RT003-R01-00. Report prepared for the Adaption Sub-Committee of the Committee on Climate Change. Wallingford: HR Wallingford.

HSE, 2011. *The costs to Britain of workplace injuries and work-related ill health in 2006/07 – workplace fatalities and self reports*, HSE Research Report RR897. Bootle: Health and Safety Executive.

ITRC, 2013. *ITRC Second assessment of national infrastructure pilot results report.* 4th ITRC Stakeholder Engagement Workshop, July 2013. Oxford: UK Infrastructure Transitions Research Consortium.

KEEF, C., TAWN, J.A. AND LAMB, R., 2013. Estimating the probability of widespread flood events. *Environmetrics*, 24 (1), 13-21.

KINGSTON, G., ROBINSON, D. and GOULDBY, B. 2008. Reliable prediction of wave overtopping volumes using Bayesian neural networks. In *Flood Risk Management: Research and Practice*, Proceedings of FLOODrisk 2008 (Oxford, 30 September to 2 October 2008), edited by P. Samuels, S. Huntington, W. Allsop and J. Harrop, pp. 561-565. Leiden, The Netherlands: CRC Press and Balkema.

KJELDSEN, T., 2010. Modelling the impact of urbanization on flood frequency relationships in the UK. *Hydrology Research*, 41 (5), 391-405.

LAMB, R., KEEF, C., TAWN, J., LAEGER, S., MEADOWCROFT, I., SURENDRAN, S., DUNNING, P. AND BATSTONE, C., 2010. A new method to assess the risk of local and widespread flooding on rivers and coast. *Journal of Flood Risk Management*, 3 (4), 323-336.

MET OFFICE, 2011. *Flooding – Summer 2007* [online]. Exeter: Met Office. Available from http://www.metoffice.gov.uk/about-us/who/how/case-studies/summer-2007 [Accessed 25 January 2017].

MET OFFICE, 2015. UK climate – Extremes [online]. Exeter: Met Office. Available from: http://www.metoffice.gov.uk/public/weather/climate-extremes/#?tab=climateExtremes [Accessed 9 December 2015].

MET OFFICE, 2016. *Past weather events* [online]. Available from: http://www.metoffice.gov.uk/climate/uk/interesting [Accessed 10 February 2016].

PENNING-ROWSELL, E., PRIEST, S., PARKER, D., MORRIS, J., TURNSTALL, S., VIAVATTENE, C., CHATTERTON, J., AND OWEN, D., 2013. *Flood and Coastal*

Erosion Risk Management: A manual for economic appraisal. Abingdon, Oxfordshire, Routledge.

ONS, 2014. *ONS Annual survey of hours and earnings: 2014 provisional results*. London: Office for National Statistics.

ONS, 2016. *Output Area (OA)* [online]. London: Office for National Statistics. Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/output-area--oas-/index.html [Accessed 2 June 2016].

SMITH, G., ARNOT, C., FAIRBURN, J. AND WALKER, G., 2005. *A national population data base for major accident hazard modelling*. HSE Research Report RR297. Bootle, Lancashire: Health and Safety Executive.

SMITH, G. AND FAIRBURN J., 2008. *Updating and improving the National Population Database to National Population Database 2*. HSE Research Report RR678. Bootle, Lancashire: Health and Safety Executive.

UNISDR, 2009. *2009 UNISDR terminology on disaster risk reduction*. Geneva: United Nations International Strategy for Disaster Reduction.

List of abbreviations

AEP	annual exceedance probability
AIMS	Asset Information Management System [Environment Agency]
ALC	Agricultural Land Classification
AMX	Asset Management eXpert [Natural Resources Wales]
CEH	Centre for Ecology and Hydrology
CDL	Continuous Defence Line [Environment Agency dataset]
Defra	Department for Environment, Food and Rural Affairs
DfT	Department for Transport
DDF	depth–duration–frequency
DRN	Detailed River Network [Environment Agency]
FEH	Flood Estimation Handbook
FFC	Flood Forecasting Centre
FRCM	flood and coastal risk management
FRTP	Flood Risks to People
FWA	Flood Warning Area
GIS	geographical information system
HTC	high throughput computing
HSE	Health and Safety Executive
HSL	Health and Safety Laboratory
JP	Joint Probability
LAD	local authority district
LRF	Local Resilience Forum
LSOA	Lower Super Output Area
MSOA	Middle Super Output Area
MDSF2	Modelling and Decision Support Framework 2
MCM	Multi-Coloured Manual
NaFRA	National Flood Risk Assessment
NFLD	National Fluvial Levels Dataset [Environment Agency]
NPD	National Population Database
NRD	National Receptor Dataset
NRA	National Risk Assessment
NRFA	National River Flow Archive

ONS	Office for National Statistics
OA	Output Area
OS	Ordnance Survey
QMED	index flood (median annual maximum flood)
ReFH	Revitalised Flood Hydrograph
SoP	standard of protection
TBR	tipping bucket raingauge
uFMfSW	Updated Flood Map for Surface Water
UPRN	Unique Property Reference Number
VOA	Valuation Office Agency

Appendix A: Data

A.1 Quality checks

A number of quality checks were performed on the flow and rainfall data received to ensure that they were suitable for inclusion in the joint probability model. A minimum of 20 years' data at each gauge was required for the gauge to be included in the analysis. Any gauges that had less than 20 years of data, based on the start and end dates of the records, were removed.

Time series plots were created for all remaining gauges to allow for further checks on data completeness and quality. Any gauges that obviously had less than 20 years of data were removed, often due to large gaps in the record. One such example is shown in Figure A.1. A maximum acceptable gap in the data at each gauge needed to be specified and was defined as one year, or 365 data points. Where such a gap was observed, data either before or after the gap was removed if there were still 20 years of data available. Where the removal of some data would result in a total record length of less than 20 years, the gauge was removed from the analysis entirely.

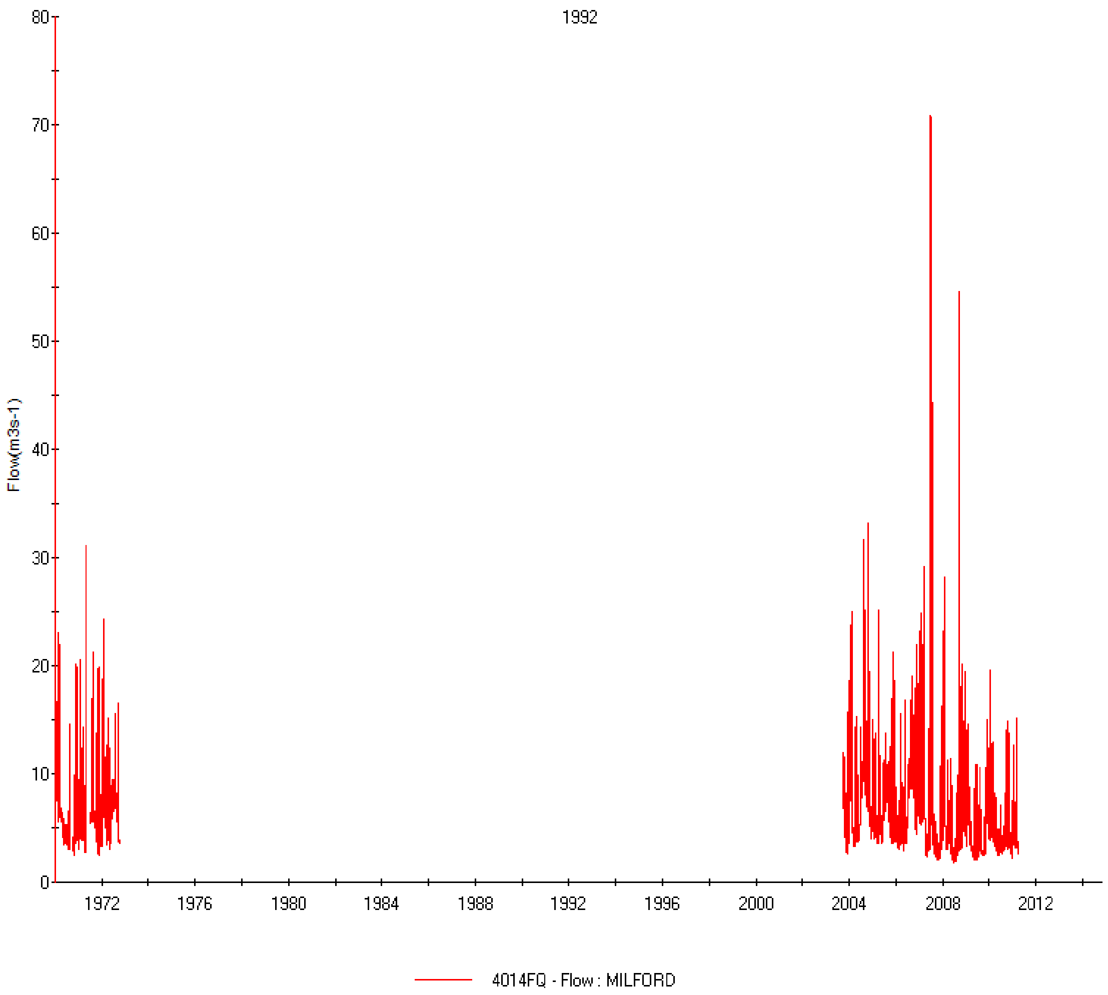
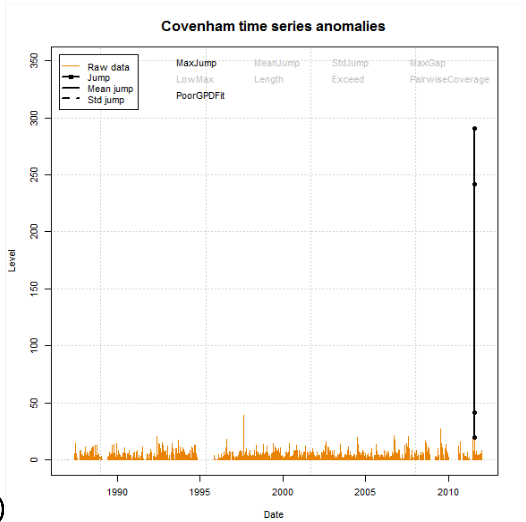


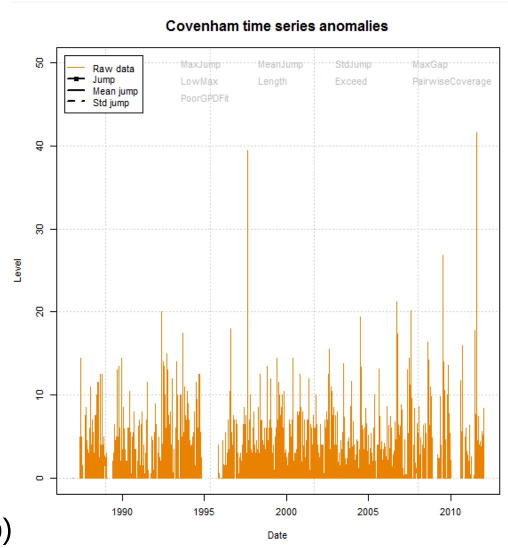
Figure A.1 Time series for flow gauge 4014, Milford

The plotting of time series also allowed for a sense check of the values within the data. Where any suspicious values were identified, the dates on which the values occurred were investigated. Where extreme values could be verified as being truly representative of an event, the values were retained. Where no explanation could be found for an extreme value, the data point was removed from the record.

Hourly TBR records were processed into 3-hourly accumulations. The highest 3-hourly rainfall recorded in the UK was 178mm (Met Office 2015). The 3-hourly accumulations at all gauges were scrutinised and any accumulations in excess of 178mm were removed. One such example is shown in Figure A.2, where several instances of 3-hourly accumulations in excess of 200mm were observed and subsequently removed from the record.



(a)



(b)

Figure A.2 Time series plot for Covenham TBR with suspicious values towards the end of the record (a) and with suspicious values removed (b)

Periods of continuous values were also identified on inspection of the time series plots. Where these were deemed to be an error, the continuous values were removed. This was a particular issue for the rainfall gauges, especially where there were periods of continuous values greater than 0mm. Where this was observed, it suggested that there was an issue with the TBR at the time of recording and the value recorded was not the true value of the rainfall occurring.

Using the refined list of gauges after initial data checks had been completed, more rigorous quality checks were used to assess the records for trends, non-stationarity, step changes and truncated data. Instances of truncated data were identified by visual inspection of the time series plots. Figure A.3 provides an example of the data checking plots for a record with truncated values at the beginning of the record (in the top plot) and then the record after the truncated values were removed (bottom plot) with an acceptable marginal Generalised Pareto Distribution (GPD) fit being obtained.

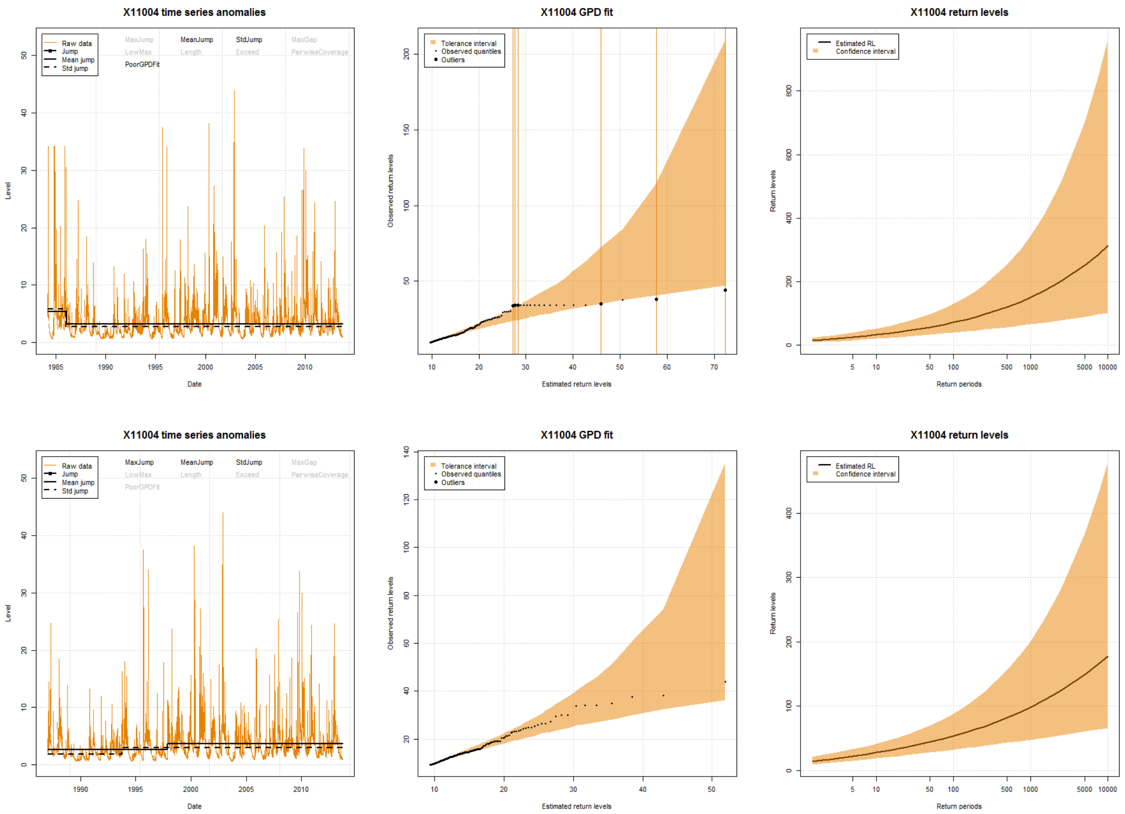


Figure A.3 Example of truncated data at the beginning of a gauge record

In some locations, both daily mean and 15-minute flow data were available. Some overlap was expected between the NRFA daily mean and Environment Agency 15-minute flow records. The locations of all gauges were checked to ensure that only one record was retained at each location. Where 2 or more records were available at a location, the longest most complete record was retained. In the majority of cases, this was the daily mean flow record as these tend to be longer than the 15-minute records. Where a 15-minute record contained information that was not evident in the daily mean record (for example, a peak during a known flood event), the 15-minute record was retained.

A.2 Issues with data

This section highlights important issues with the data which considerably reduced the final number of gauges suitable for inclusion in the joint probability model.

For all data received from the Environment Agency, the expected start and end dates were taken from the WISKI catalogue which provides metadata for the gauges held in the WISKI database. Originally these dates were used to identify records that were at least 20 years in length and would potentially be suitable for inclusion in the event set. On inspection of the data files received from the Environment Agency, it became clear that the start and end dates in the metadata of the WISKI catalogue did not truly represent the data available at the gauges. Having identified this, time was taken to interrogate each data file and determine the true start and end dates for each gauge.

In several instances, more than one data file was received for a gauge. This tended to occur where data had been stored in 2 separate systems with the first file containing data from before 2003 and the second file containing data from 2003 onwards (the time of migration to WISKI). In a number of cases, the values within the 2 files varied considerably and did not appear to be records for the same gauge (Figure A.4). Where this issue occurred, efforts were made to determine which record contained the true values for the gauge and the incorrect part of the record was removed. Where the remaining record was less than 20 years in length, the gauge had to be removed from the analysis.

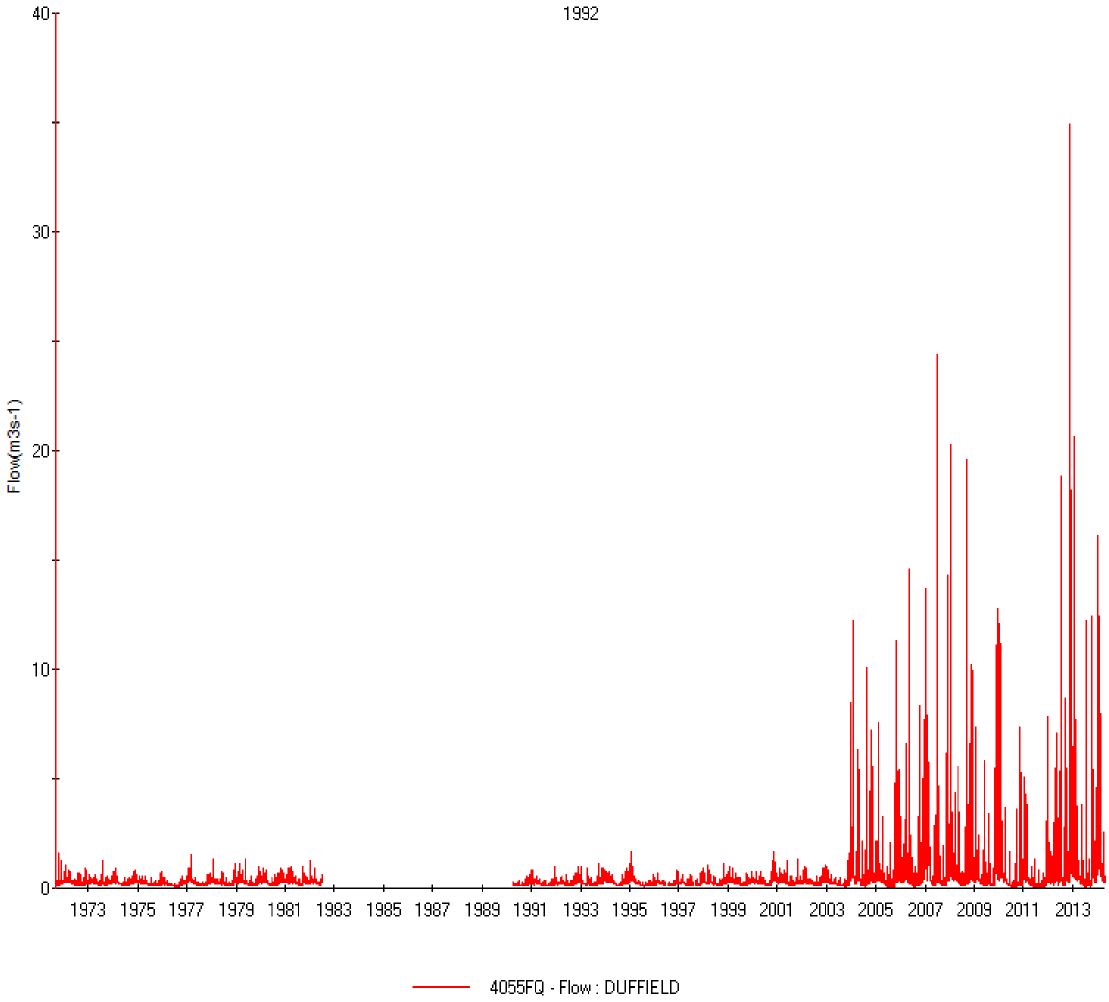


Figure A.4 Time series for flow gauge 4055, Duffield

In the Environment Agency gauge records, gaps in data were often observed around 2003. These gaps tend to be as a result of the migration from the previous data storage system to the WISKI system. Where gaps were identified, Environment Agency staff performing the data extractions were able to check whether the missing data existed and supplied additional files containing the missing data where possible.

Some gauge records contained suspicious values – with some instances of negative values and a number of records containing exceptionally high values. Gauges that contained negative values of flow (for example, Figure A.5) were removed from the analysis as the values could not be used.

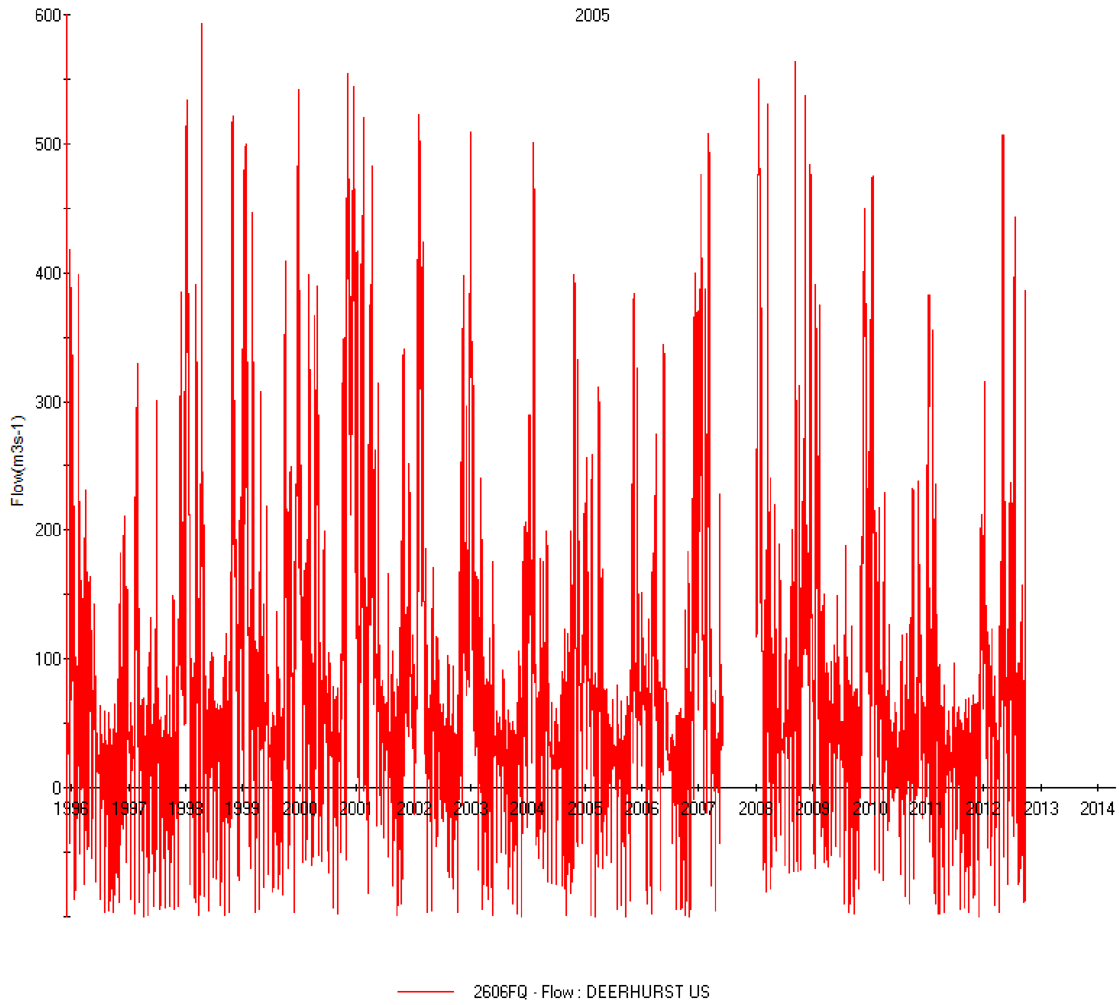
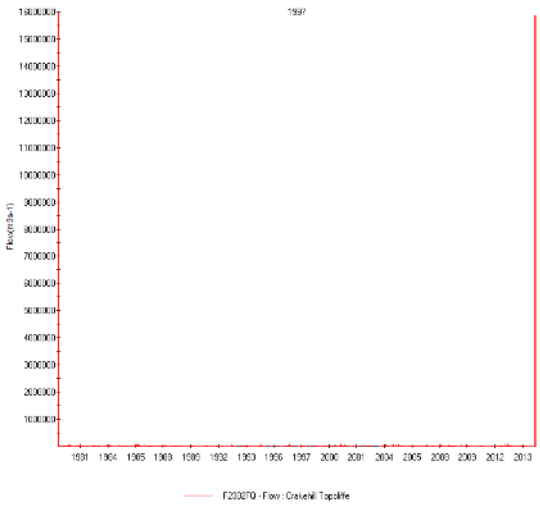


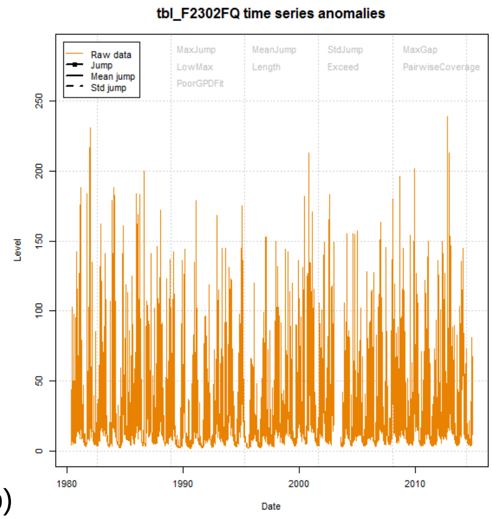
Figure A.5 Time series plot for flow gauge 2060, Deerhurst US

Where records contained values that were suspiciously high, these values were investigated to determine whether they could be true values during an event. Where the values could be justified, often for events such as the summer 2007 or autumn 2000 floods, no further action was taken. Where the extreme values could not be verified, they were removed from the record.

Some gauge records contained values that were so extreme that they could not be considered as a potentially real flow value. Where these values occurred just a few times within a gauge record, the values were removed and the remainder of the record retained. One such example is shown in Figure A.6, where an unrealistic flow value of 16,000,000 m^3s^{-1} was included in the received data file. In any instances where extreme values occurred repeatedly, the gauge was removed completely from the analysis.



(a)



(b)

Figure A.6 Time series plot for flow gauge F2302, Crakehill Topcliffe with unrealistic value at the end of the record (a) and with unrealistic value removed (b)

For some gauges, the data received were at differing temporal resolutions over the course of the record. A variety of resolutions including hourly, 2-hourly, 3-hourly and daily were identified. These variations tended to occur at the start of gauge records. There were also some instances of records containing data at times other than the 15-minute intervals. For gauges where this occurred, an additional entry was included at 08:59:59 followed immediately by an entry at 09:00:00. Only the data entries at 09:00:00 were included in the analysis.

On inspection of the time series plots, it became clear that the data for some gauges had been truncated at a specific value throughout the record. One such example is shown in Figure A.7. Where truncated values were present in a record, the values were either removed (as shown in Figure A.3) and the remaining record still used in the analysis, or the entire record was removed as none of the data were suitable for use (Figure A.7).

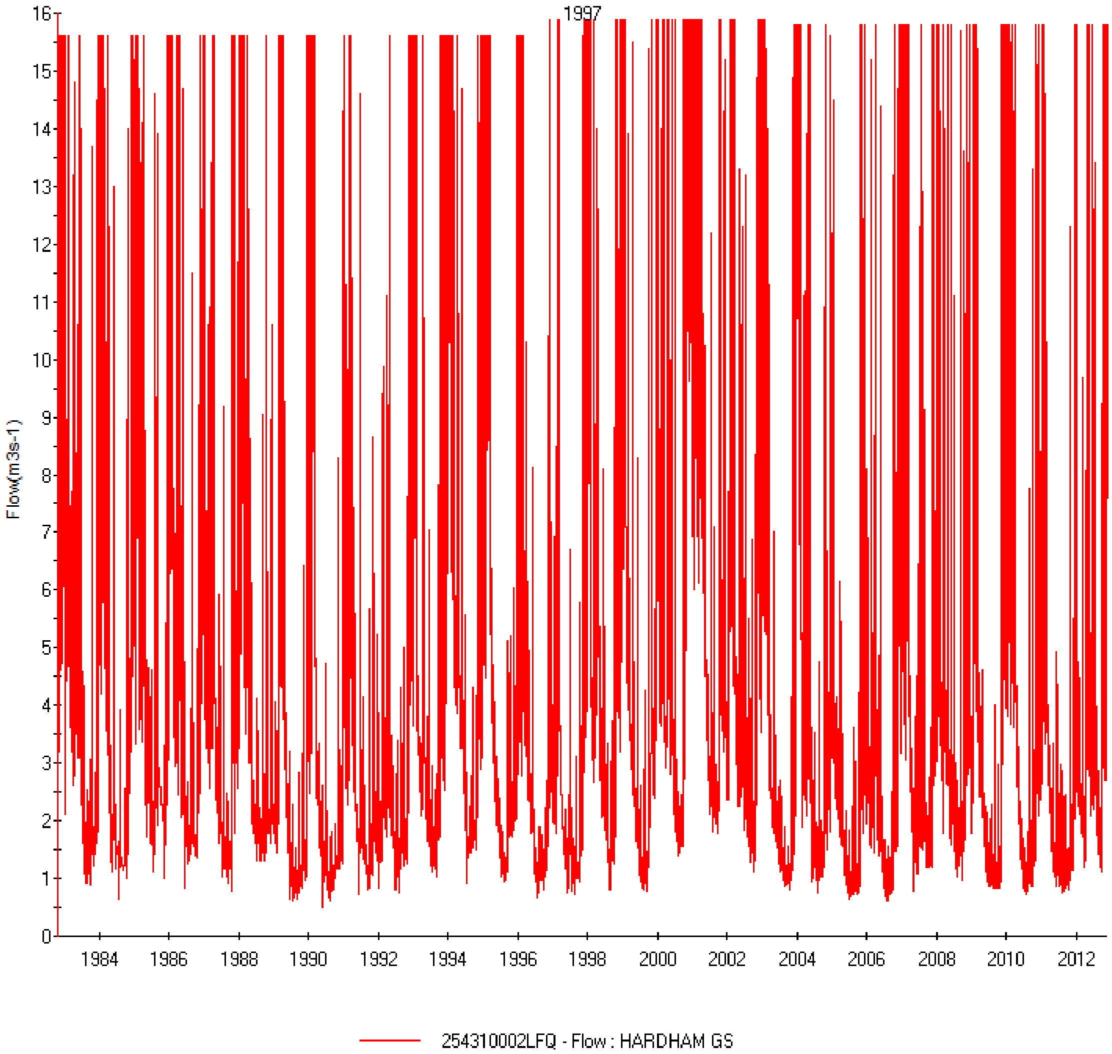


Figure A.7 Truncated flow values for flow gauge 254310002L, Hardham GS

Appendix B: 2011 Census calculations for population vulnerability

1. Residential Night-time (OAs) =

Numerator:

- Total population suffering from a limiting long-term illness aged 0–74 (LC3302EW)
- Population aged 75+ (KS102EW)

Denominator:

- Total population (KS102EW)

2. Residential Daytime term time (LSOAs) =

Numerator:

- Total population suffering from a limiting long-term illness aged 0–4 (LC1301EWLS)
- Total population aged 16–74 not at work due to limiting long-term illness (QS601EW)
- Population aged 75+ (KS102EW)

Denominator:

- Total population aged 0–4 (KS102EW)
- Total population aged 16–74 at home (QS601EW)*
- Population aged 75+ (KS102EW)

3. Residential Daytime non-term time (OAs) =

Numerator:

- Total population suffering from a limiting long-term illness aged 0–15 (LC3302EW)
- Total population aged 16–74 not at work due to limiting long-term illness (QS601EW)
- Population aged 75+ (KS102EW)

Denominator:

- Total population aged 0–15 (KS102EW)
- Total population aged 16–74 at home (QS601EW)*
- Population aged 75+ (KS102EW)

* Total population aged 16–74 at home is calculated from the following categories in QS601EW:

- Economically Active - Unemployed
- Economically Active - Full-Time Student
- Economically Inactive - Retired
- Economically Inactive - Student
- Economically Inactive - Looking After Home/Family
- Economically Inactive - Permanently Sick/Disabled
- Economically Inactive - Other

4. Sensitive **Schools and childcare** (MSOAs) =

- Total population suffering from a limiting long term illness aged 0-15 *(LC3302EW)*

divided by

- Total population aged 0-15 *(LC3302EW)*

5. Sensitive **Hospitals and care homes** =

All people in these populations are considered to be vulnerable.

6. **Places of work** (LADs) =

- Total population suffering from a limiting long term but at work *(LC6302EW)*

divided by

- Total population at work *(QS601EW)**

7. **Roads, stadia, prisons (sensitive), and transport** (LADs) =

- Total population suffering from a limiting long term illness aged 0-74 *(LC3302EW)*
- Population aged 75+ *(KS102EW)*

divided by

- Total population *(KS102EW)**

8. **Leisure Caravans and camping, Other accommodation** (National average: England and Wales) =

National proportion of vulnerable people = 0.20**

* Total population aged 16–74 at work is calculated from the following categories in QS601EW:

- Economically Active - Employee
- Economically Active – Self-employed with employees
- Economically Active – Self-employed without employees

** Calculated from 2011 Census statistics as the overall proportion of the total England and Welsh population aged over 75 or long-term ill.

Figure B.1 2011 Census calculations for population vulnerability

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

Then call us on
03708 506 506 (Monday to Friday, 8am to 6pm)

email
enquiries@environment-agency.gov.uk

or visit our website
www.gov.uk/environment-agency

incident hotline **0800 807060** (24 hours)
floodline **0345 988 1188 / 0845 988 1188** (24 hours)

Find out about call charges (www.gov.uk/call-charges)

Environment first: Are you viewing this on screen? Please consider the environment and only print if absolutely recessary. If you are reading a paper copy, please don't forget to reuse and recycle if possible.