University of
Reading

A multi-scale framework for flood risk analysis at spatially distributed locations

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

A multi-scale framework for flood risk analysis at spatially distributed locations

Dr. Linda J. Speight

School of Civil Engineering and Geoscience, Cassie Building, Newcastle University. Newcastle upon Tyne, NE1 7RU, UK

Corresponding author: l.j.speight@gmail.com

Professor Jim W. Hall FREng

Environmental Change Institute, Oxford University Centre for the Environment, South Parks Road, Oxford, OX1 3QY, UK

Professor Chris G. Kilsby

School of Civil Engineering and Geoscience, Cassie Building, Newcastle University. Newcastle upon Tyne, NE1 7RU, UK

A multi-scale framework for flood risk analysis at spatially distributed locations

Abstract

This paper presents a multi-scale framework for flood risk analysis from fluvial and coastal sources at broad (including national) scales. The framework combines an extreme value spatial model of fluvial and coastal flood hazards using the Heffernan and Tawn (2004) conditional dependence model, with a new Markov approach to representing the spatial variability of flood defences. The nested multi-scale structure enables spatial and temporal dependence at a national scale to be combined with detailed local analysis of inundation and damage. By explicitly considering each stage of the process, potential uncertainties in the risk estimate are identified and can be communicated to end users to encourage informed decision making. The framework is demonstrated by application to an insurance portfolio of static caravan sites across the UK worth over £2bn. In the case study the largest uncertainties are shown to derive from the spatial structure used in the statistical model and limited data on flood defences and receptor vulnerability.

Key words

coastal; dependence; fluvial; flood risk; multi-scale; systems modelling

Introduction

Analysis of flood risk integrates consideration of (i) the probability of flood hazards occurring ('source's) (ii) the processes by which flood waters propagate to the places where they can cause harm ('pathways') and (iii) the exposure and vulnerability of people and assets that may be harmed by flooding ('receptors'). Systems based modelling of flood risk using a source-pathway-receptor framework is well established (Hall *et al.* 2003, Gouldby *et al.* 2008, Apel *et al.* 2006). Such a framework is useful as it allows consideration of all driving factors in flood risk.

Insurance industry catastrophe (Cat) models (Grossi and Kunreuther 2005) now provide essential tools for the pricing of insurance. They include quantification of sources, pathways and receptors but rather than focussing directly on flood damages they include an additional insurance losses module. They are national or international scale models which implicitly include the spatial and temporal structures in the driving source variable through the use of large ensembles of synthetic weather inputs. As the prominence and capabilities of Cat models increases there is concern within the industry that insurers are relying too heavily on Cat models rather than their own judgement (Gray reported by Lloyd's 2006; Clarke reported by Gusman 2008). This concern is reflected in the new Solvency II legislation which requires insurances companies to be able to justify any decisions and assumptions they make when pricing insurance (European Parliament Council 2009). To do this an improved understanding of systems risk models by end users is required.

Developments in research into individual components of the source-pathway-receptor framework have advanced considerably since the concept was originally outlined by Hall *et al.* in 2003. Recent interest has focused on the spatial and temporal dependences in extreme rainfall, river flow and

coastal events in the UK and potential correlations between different driving sources (Coles and Tawn 2005, Dixon and Tawn 1997, Hawkes et al. 2002, Keef et al. 2009a, Keef *et al*. 2009b, Keef et al. 2013, Svensson and Jones 2002, Svensson and Jones 2004). Proper treatment of dependence is essential for accurate estimation of flood risk. Consideration of spatial dependence is necessary for broad-scale assessments, including the probability of aggregate losses over a large spatial domain, such as occurred in England during the 2007 floods (Pitt 2008). Risk analysis at broad (including national scales) is necessary to plan the resources that might be necessary to respond to widespread emergencies. It is also essential for insurers to seek to diversify their portfolios in space.

For the pathways component the EU FLOODsite project has made major advances in the small scale understanding of flood defence reliability and breaching processes (Allsop *et al*. 2007). At the same time there has been improvements in the representation of flood defence reliability in systems risk models (Dawson *et at*. 2005, Apel *et al*. 2006, 2009, Vorogushyn *et al*. 2009 and 2010, Buijs *et al*. 2009) although these improvements still lag behind the developing local scale understanding.

Flood hazard mapping has also improved considerably. Since the first national flood risk assessments in England and Wales (Hall *et al.* 2003) agencies across Europe are now publishing second or third generation flood maps in line with EU legislation (De Moel *et al*. 2009) and continual developments in computational power and hardware, parallel processing and cloud computing are enabling faster and more efficient inundation modelling (for example Glenis *et al*. 2013, Lamb *et al*. 2009, Neal *et al*. 2010). Work on understanding flood damage however has not progressed as much as other areas. In the UK the Multi-Coloured Manual (Penning-Roswell *et al*. 2005, updated in 2013) remains the main source of flood damage information despite the limited database and necessary assumptions used in the calculations. Elsewhere Merz *et al.* (2004, 2009, 2010) have reviewed potential uncertainty in depth-damage data and alternative datasets do exist (for example HOWAS in Germany and the Dundee Tables in the UK as well as insurance databases) although access to these is limited and the value of the additional information varies by the receptor of interest.

The inclusion of these advances into full systems models is limited and where it does occur is often focused on particular areas of interest and does not explicitly address all of the key driving factors influencing risk with relative weight. For example recent approaches by Lamb *et al*. (2010) and Wyncoll and Gouldby (2013) make considerable headway in incorporating the spatial and temporal dependence in the source variables but do not provide similar detail for the hazard impact assessment. An integrated framework that considers all contributing factors and the spatial and temporal dependencies between them is therefore required.

Hall *et al.* (2003) outlined a hierarchy of assessment methods for different scales of analysis. While this approach allows for increasing levels of detail to be included at smaller scales it does not support cascading of information from one level to the next. For example the most detailed level recommends continuous simulation of hydraulic loads (Gouldby *et al.* 2008), since this is difficult at a national scale, there is no consideration of wider scale dependencies in the detailed analysis. Keef *et al*. (2009a) produced a national scale assessment of spatial and temporal correlations in flood risk. Their assessment focused on correlations between extreme river events at gauged locations and used an interpolation method to provide data between the gauging stations. In recognition of the computational load of a national scale model the study outlined a relatively simplistic methodology by which the output from the statistical model could be used to assess the impact of extreme river levels on receptors (Lamb *et al.* 2010).

This paper outlines a framework for risk analysis that incorporates recent advances in statistical dependence modelling combined with an assessment of consequences. It draws on the above ideas to develop a novel nested multi-site approach. In this way more detailed analysis can be carried out at the sites of interest including consideration of defence failure and detailed inundation modelling while still maintaining a national structure and incorporating the broad scale spatial correlations between sites. A modular approach is adopted which allows for changes in one component of the system without having to change all components, or for adding additional components. It also means that each stage of the process can be validated independently. This is important when taking a whole system approach as validation of the end to end system model can be difficult. The framework explicitly couples large scale spatial dependencies in extreme fluvial and coastal events with local scale dependencies in defence crest height, defence reliability, floodplain inundation and damage. The framework is illustrated using a case study of flood risk to an insurance portfolio of static caravans (also known as mobile homes or trailers) located across the UK. The framework is based on a nested structure. This allows for large scale spatial dependence structures to be included, in this case between extreme fluvial and coastal events in the UK, at a high level as well as detailed modelling of inundation and damage at sites of interest.

A key challenge with systems risk models is the cascade of uncertainty through the system. Each component makes a contribution to the uncertainty in the calculation of overall risk. Previous studies (for example Dawson *et al*. 2008 and Hall *et al*. 2011) have attempted to identify sources of uncertainty within the risk model and it is generally assumed that uncertainty in the extreme values statistics of the sources component is much larger than other factors (Apel *et al*. 2004). However there is an increasing need for transparent systems risk modelling that acknowledges the uncertainties within the system and communicates this uncertainty to the end user. A qualitative assessment of the relative importance of each component to the overall uncertainty is made in this paper with the aim of helping to increase understanding of systems risk models to help enable informed decision making for flood risk management and insurance pricing. This process also aims to identify future research to improve the most critical areas.

Outline of the proposed framework

The hazard variables ('sources') considered in the system model are river flow, tide, storm surge and waves. The 'pathways' include rivers, flood defences and floodplains and the receptors in this example are static caravan sites. The consequence modelled is flood damage. The dependence structure within, and between, these components applies at different scales. Representing this as a nested structure in the modelling framework, as show in Figure 1, enables clear identification of appropriate levels of modularisation within a systems model and illustrates how dependence structures are cascaded from the large scale down to site level.
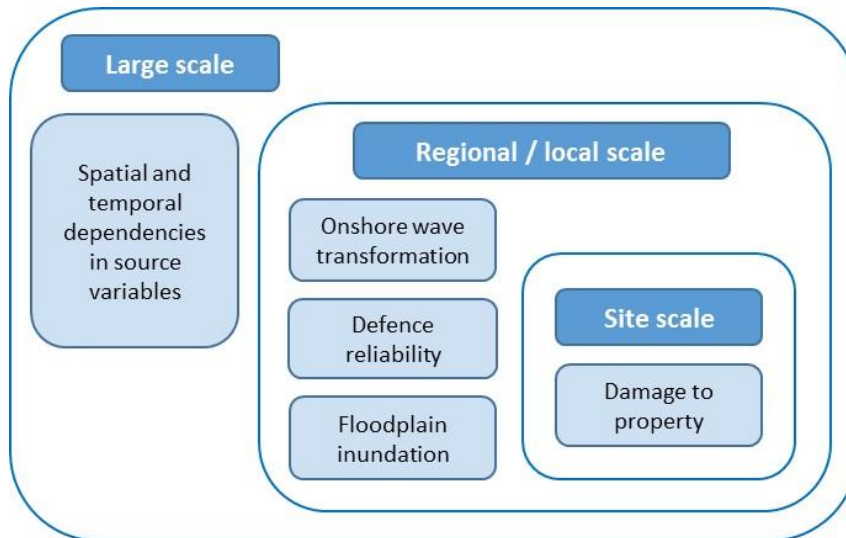
Figure 1 Nested framework structure of systems components

Sources input

Two sources, and the dependence structure between them, are considered: river flows and total sea level. For each event, i, (where an event is defined when one or more gauges in the network is above the 99% quantile) daily mean river flows and skew surge are simulated at all the gauges in the network using the Heffernan and Tawn (2004) conditional dependence model. The model separates the marginal and dependence characteristics and models them separately through the use of a Copula. Within the conditional dependence model the marginal characteristics are modelled using a standard Generalised Pareto Distribution (GPD). For events where one or more gauges in the network is extreme, the model is used to simulate values at all gauges at lags of up to five days from the event peak. As the model is primarily concerned with modelling extreme events, the distribution below the threshold is assumed to follow an empirical distribution.

Previous applications of the Heffernan and Tawn model (Keef *et al.* 2009a 2009b, 2011, 2012; Lamb *et al.* 2010; Wyncoll and Gouldby 2013) have identified that a number of decisions and assumptions are required when applying the model to multi-scale flood extremes, particularly in selecting the marginal distribution, the extreme event threshold and the definition of an event at individual and multiple sites. Following testing of the assumptions for suitability to this application, the implementation of the model in this paper largely followed that of Keef *et al.* (2011). One notable exception is that the lag time used define multi-site events is increased from the three days proposed by Keef *et al.* (2011) to five days to better represent the spatial dependence between upstream sites and the larger catchments of the Severn and Thames included in this study. The same five day lag is also used when de-clustering the gauged data at individual sites.

The total sea level is assumed to comprise of a tide, surge and wave component. The tide component is modelled deterministically using the full tidal range at the sites of interest. The surge component is modelled using the conditional dependence model fitted to the skew surge (the difference between the maximum observed sea level and the maximum predicted tide from the nearest tidal cycle) from each event. The swell waves from the nearest wave rider gauge are also modelled using a pairwise version of the conditional dependence model conditional on skew surge at the nearest tidal gauge.

Using a similar approach to Gouldby *et al*. (2014), the conditional dependence model is fitted to offshore wave height. To enable transformation to onshore wave heights the CRESS v10 (Netherlands Ministry of Public Works, IHE-Delft *et al*. 2010) calculation package is used with wave period simulated based on a linear regression on wave height, and wave direction simulated from the conditional probability of wave direction given wave height. The simulated vector of fluvial and coastal sources across the network, X_i, represents a large scale spatial event.

The local dependencies are then addressed at each site or risk cluster, *s*. Using a national scale statistical model for local scale analysis is difficult due to the juxtaposition of requiring long, concurrent datasets on a national network (Keef *et al*. 2009a) and the suitability of gauge locations and using daily mean flow data (CEH 1999) for site scale flood risk assessment. The daily mean river flow near the caravan site of interest is estimated by applying a standard ungauged site data transfer method as recommended by the Flood Estimation Handbook (CEH 1999) to the observed daily mean flow at the nearest gauge to the site (denoted q_i). The daily mean flow is then converted into a peak flow and hydrograph Q_i, using a form of the Sangal (1983) method modified for use with UK data (Speight 2013), such that;

$$Q_i = f(q_i, k_s) \qquad \text{Eq.1}$$

where k_s is the ratio of daily mean flow to flood peak specified for each site. The temporal structure of the inflows at different sites is maintained through the definition of event length specified when fitting the conditional dependence model at relevant time lags as detailed by Keef *et al*. (2011).

The combined coastal component at the site of interest is labelled z_i. This is transformed to represent inshore wave heights at the site of interest such that,

$$Z_i = f(z_i, h_s) \qquad \text{Eq.2}$$

where h_s represents the local bathymetry between the wave rider and shoreline of interest. A representative surge profile is sampled using a version of Ward's algorithm (as per Apel *et al*. 2004, 2006).

Further details of the practicalities of using a complex conditional dependence model in an applied systems model for fluvial and coastal flood risk are discussed in Speight, Hall and Kilsby (in progress).

Consideration of pathways

Pathways are the routes by which water reaches the sites of interest. The main focus is on the reliability of flood defences as this is likely to cause the greatest variability in flood risk. The framework proposed is modular, so in the example application two types of inundation model were tested to illustrate the ability to use model components suitable for the local requirements; the fluvial sites were represented using a 1D ISIS model (CH2MHILL 2014) and the coastal sites were modelled using a 2D shallow water flow model (Liang 2010).

Defence system

In a flood defence system with *n* sections, d_1 to d_n, characterised by their construction type and standard of protection, any one of the defences can fail in one or more locations resulting in

inundation of the floodplain. Based on analysis of mean crest height for each section (as often used to describe flood defence profiles) and surveyed crest height along the section length of coastal defences in the Environment Agency East Anglia region it was found that the mean crest height exclude important low points and variation in crest height along the defence length. In most cases the difference was between 0.1m and 0.5m. This is a maximum error of 8% in crest height which presents a significant uncertainty for flood risk modelling. To account for this important variation in a systems model where detailed survey data may not be available, a methodology to generate realistic spatially varying crest heights using a Markov model is proposed such that the degree of variation depends on the defence type and condition;

$$c_{j,i} = f \, (defence \; type, defence \; condition) \hspace{3cm} \text{Eq.3}$$

where $c_{i,j}$ is the simulated crest height and resistance to load of any given defence, j, for each event, *i*. Three defence types are represented; walls, embankments and sand dunes. For each defence type, a misalignment type (consistent, varying, block changes or low/high points) is sampled based on the observed occurrence rate for each defence type from the survey data compared to mean crest height for each section. Then for each simulation a realisation of the crest height is generated based on a Markov chain process, whereby at each simulation point (taken as every 20m) there is a strong probability of the simulated crest height being the same as the previous crest height and a lower probability that it will move to a new state. The probabilities are based on the distribution properties of the observed data and are different for each misalignment type. The combined vectors of crest height for the whole system is referred to as C_i. A similar approach could be taken to represent variability in the defence strength along its length although this was not tested as detailed data on defence strength is not easily available.

Water level and overtopping

The water level is referred to as L_i through the modelled reach or shoreline, or $l_{j,i}$ at a particular defence section. Coastal water level is determined by inshore transformation of the simulated sea state from Eq.2. The fluvial water level is simulated using a hydraulic model for the specified flows and defence crest heights;

$$L_i = f(C_i, Q_i \, Z_i) \hspace{3cm} \text{Eq.4}$$

Overtopping of defences is considered deterministically based on the modelled water level conditional on the simulation of the defence crest height. The probability of overtopping, $P(OT_i | C_i)$ depends on the probability of the extreme event. Due to the simulated variation in crest height (C_i) this may not be equivalent to the standard of protection of the defence. Defence overtopping occurs when the water level at the defence is greater than the crest height. This can be defined throughout the model OT_i, or for individual defences, $ot_{j,i}$ or points within an individual defence, $ot_{j,y,i}$, where *y* is the chainage along the defence section from 1 to *t*. Therefore;

$$OT_i = f(C_i, L_i) \hspace{3cm} \text{Eq.5}$$

Defence failure

A defence is said to have failed if it has breached in one or more locations. Overtopping is considered directly via the modelled water level so failure only relates to breaching following overtopping. The failure of defence d_j is labelled as event D_j. The non-failure of defence d_j is labelled as $\bar{D}_j$.

The conditional probability of failure for any given event, $P(D_{j,i}|\alpha_i)$ is modelled as a function of the water level, amount of overtopping, and defence reliability (where α_i is a matrix of these loading variables for event i) such that for any given defence;

$$P(D_{j,i}|\alpha_i) = f(l_{j,i}, ot_{j,i}, r_{j,i})$$ Eq.6

This can be extended to consider any point along the defence;

$$P(D_{j,y,i}|\alpha_i) = f(l_{j,y,i}, ot_{j,y,i}, r_{j,y,i})$$ Eq.7

Assuming that breaches are independent given the loading conditions, the probability of breaches at locations $y=1$ and $y=2$ is;

$$P(D_{j,1,i}|\alpha_i) \; P(D_{j,2,i}|\alpha_i) \; P(\bar{D}_{j,3,\ldots,t,i}|\alpha_i)$$ Eq.8

where t is the total number of locations along the defence system considered.

The probability of defence failure for any given load is defined using a fragility curve, giving the conditional failure probability $P(D_j|\alpha_j)$, in this case the loading variables are defined at defence j. The unconditional failure probability, $P(D_j)$ is given by;

$$P(D_j) = \int_0^\infty p(\alpha_j)P(D_j|\alpha_j)d\alpha$$ Eq.9

Where $p(\alpha_j)$ is the probability density function of the loading variables, α, at defence j.

Sample probabilities of breach scenarios

A central assumption of most previous work on flood defence reliability (e.g. Hall *et al.* 2003) is that all defence sections are loaded at the same time. In practice this may not be the case due to the potential reduction in water level following an upstream breach, particularly in areas with significant floodplain storage (Apel *et al.* 2009).

An iterative sampling procedure is proposed to incorporate the dependence on previous upstream breaches for fluvial loading. Firstly the hydraulic model is run assuming no failures to give $L_{i,\overline{D_1,\ldots,D_n}}$. Successive sequences of one or more defence breaches, D_{ss}, are sampled based on the conditional failure probabilities at each point in the defence $P(D_{j,y,i}|L_{i,\overline{D_1,\ldots,D_n}})$ until $P(D_{ss}|L_{i,\overline{D_1,\ldots,D_n}}) \to 0$. The hydraulic model is run for each sequence with a standard breach size and growth rate to provide water levels at each section including the impact of potential reduction in water level from upstream breaches. The conditional probability of each failure sequence $P(D_{ss}|L_i)$, is the product of the failure probability at each breach point $P(D_{j,y,i}|L_i)$, for the modelled water level $l_{j,y,i,D_{ss}}$. One of the failure sequences, $D_{ss,i}$, is sampled from the probability distribution $p(D_{ss}|L_i)$ for each event. Each breach in $D_{ss,i}$ is assigned a maximum width given by;

$$W_{ss,i} = f(L_i, defence\ type, floodplain\ size\ and\ shape, timing\ of\ breach) \qquad \text{Eq.10}$$

Note that the loading variable used in the iterative sampling process is water level, defence crest height and reliability are assumed to remain constant within each sampled event. Discussion of how the defence type, floodplain size and shape, and timing of the breach in relation to the flood peak influence the breach size are given by Morris *et al.* (2009a; 2009b) and Muir-Wood and Bateman (2005).

For coastal loading there is unlikely to be a reduction in total water level and the fluvial clusters selected for further analysis in this research did not contain significant storage behind levees. The above steps are therefore included for completeness if the methodology were to be applied in different areas in the future; however they have not been tested.

By not considering the sequential failure probabilities there is no need for iterative model runs and the probability of each failure sequence $P(D_{ss}|L_j)$, is the product of the failure probability at each breach point $P(D_{j,y,l}|L_j)$, for the modelled water level $L_{i,\overline{D_1,...,D_n}}$.

Damage conditional on event

For each event, X_i, for the specified defence state variables, C_i, $D_{ss,i}$, and $W_{ss,i}$, a hydraulic model is run to calculate water level in the channel and flows onto the floodplain. A raster based floodplain inundation model is used to calculate flood depths across the floodplain. Using receptor specific depth damage curves these floodplain depths are converted to loss estimates at each site using a simple financial model, and summed together to give the loss across the portfolio conditional on the event. Loss conditional on event is denoted φ_i. The process is illustrated in Figure 2.
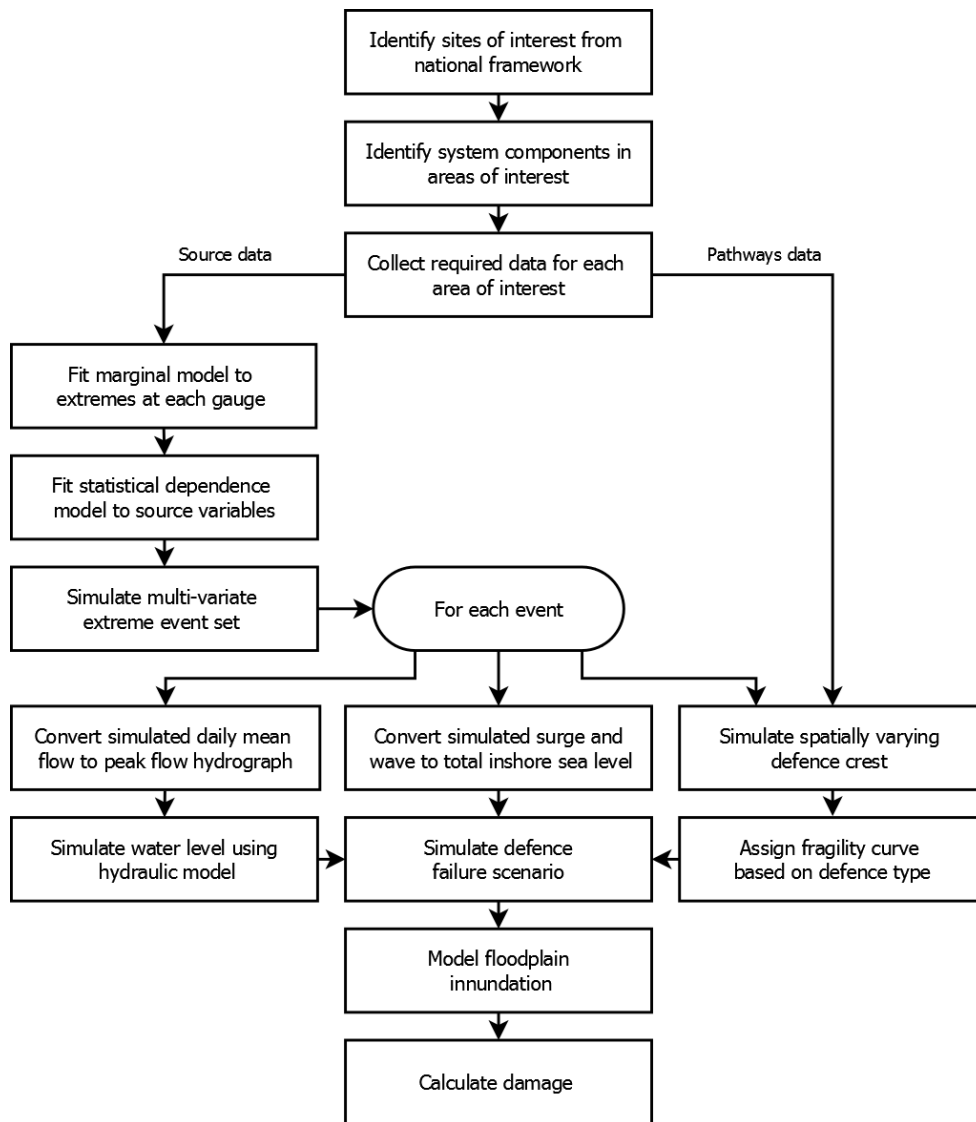
Figure 2 Overview of the risk assessment framework

Calculating risk

Risk is defined, as is standard practice, as the probability of an event multiplied by the consequence (damage) integrated over all possible events;

$$R = \int p(\alpha)\, V(\alpha)\, d\alpha \qquad\qquad \text{Eq.11}$$

where $p(\alpha)$ is the joint probability function and $V(\alpha)$ is the damage function of the loading variables (α) simulated through Eq.1 to Eq.10 which now includes flow through any resulting breaches as well as water level, crest height, overtopping, and defence reliability.

Assuming a large enough sample size of extreme events simulated from the conditional dependence model $p(\alpha)$ can be established empirically such that:

$$p(\alpha_i) \approx \frac{1}{\sigma} \sum_{s=1}^{k} \# \{X_s > x_{i,s} \mid X_1 > x_{i,1}, \dots, X_{s-1} > x_{i,s-1}\}$$

where σ is the sample size, s is the closest gauge to the site or risk cluster, k is the total number of gauges in the system, x is the simulated value at each gauge for event i, and X is a vector of all simulated values at each gauge. #{.} is a count of the number of times the value at gauge s for all simulated events is larger than the event, $x_{i,s}$, conditional on meeting the rejection sampling r. This ensures the large scale dependence structures in the data set are maintained and events are not double counted. The event probability is established from the probability at the gauging stations, $P(X_i)$. $V(\alpha)$ is the damage function which is a sum of damage sustained at all sites of interest for the event. This is equivalent to φ_i.

One way of evaluating this integral is by Monte Carlo integration, based upon η Monte Carlo samples from $p(\alpha)$, in which case:

$$R \approx \frac{1}{\eta} \sum_{i=1}^{\eta} V_i(\alpha)$$

The proposed framework also considered the effect of flood defence systems in modifying the probability of flooding. In the simplest instance, for a system of n flood defence sections, there are $t= 1,\dots,2^n$ possible system states, each of which has a probability of failure which is conditional upon the loading variables, written $P(D_{ss,t}|\alpha)$. In this case Eq. 11 is modified such that risk is calculated as:

$$R = \int \sum_{t=1}^{2^n} p(\alpha) V(\alpha_{D_{ss,t}}) \, P(D_{ss,t} \mid \alpha) \; d\alpha$$

and the Monte Carlo estimate is:

$$R \approx \frac{1}{\eta} \sum_{i=1}^{\eta} \sum_{t=1}^{2^n} V_i(\alpha_{D_{ss,t}}) \, P(D_{ss,t}|\alpha)$$

Application

The framework was developed using risk to an insurance portfolio covering the majority of the static caravans in the UK, with a total value in excess of £2bn. The caravans are located across the country and therefore require analysis over large spatial scales. The caravan sites reflect characteristics of many other spatially distributed receptors across the UK and hence provide a useful case study. For example within the national portfolio there are clusters of known high risk, areas affected by fluvial, coastal and combined sources of flooding and areas particularly vulnerable to defence failure.

The first stage was to map the distribution of insured units and to screen the sites most at risk using existing hazard maps (in this case the Environment Agency (2014) extreme fluvial and coastal flood maps for England and Wales). Insurance cover is based on the premise that risk can be shared across multiple customers however if customers are grouped into particular areas of the country or if multiple customers might be exposed to extreme events at the same time then insurance companies may find themselves with a large claims bill. Four high risk clusters were identified based on five criteria:

- The value of assets at risk

- Achieving a balance of fluvial and coastal sites

- Accessibility of data

- Good spatial distribution of sites

- Links between risk processes across the cluster

The clusters selected were the Lincolnshire coast which had the largest concentration of caravan sites at risk and a high vulnerability to coastal flooding; the North Wales coast which has the second highest concentration of sites and is affected by both coastal and fluvial flooding from flashy mountain catchments; an area in the Midland near Worcester which has the highest concentration of inland sites and is at risk from the slow responding River Severn; and a small cluster of high value sites near Hurley on the River Thames. These four areas are used in the more detailed local analysis stage within the nested model structure.

Required data

A major consideration of a full system based model is the large amount of data required. The multisite approach restricts this to specific areas, however the range of data needed is still considerable. Table 1 lists the main data components used in this application and identified potential data driven limitations.

Table 1 Data used in the application of the risk assessment framework

Model component	Data required	Source used	Data format	Known limitations
Source (Fluvial)	Concurrent flow data for river reaches of interest	Centre for Ecology and Hydrology (CEH) National Rivers Flow Archive (NRFA)	Daily mean flow (DMF) at each gauging station	Limited number of extremes in record. Gauges not necessarily located near caravan sites. DMF does not provide flood peak.
Source (coastal)	Concurrent still water level including tidal and surge components for coastlines of interest	British Oceanic Data Centre (BODC) UK Tide Gauge Network	15 minute to hourly predicted and observed sea level	Limited number of extremes in record. Gauges not necessarily located near caravan sites.

Model component	Data required	Source used	Data format	Known limitations
Source (coastal)	Concurrent significant wave heights for coastlines of interest	Centre for Environment, Fisheries and Aquacultural Science (Cefas) Wavenet network	Wave height, period and direction at up to 30 minute resolution	Wave data recorded at offshore buoys, requires transformation to near shore
Pathway	River channel dimensions	Environment Agency LiDAR	Raster tiles at 2m resolution	Interpolation required below water level.
Pathway	Flood defence location, type and height	Environment Agency National Flood and Coastal Defence Database (NFCDD)	GIS files of defence location and other available data	Poor spatial resolution of data. Limited data for defence reliability analysis.
Pathway	Floodplain geometry	Environment Agency LiDAR	Raster tiles at 2m resolution	
Receptor	Location and value of spatially distributed receptors	Exposure database of caravan sites	Postcode locations of total insured value at each site	Number / value of individual units at any given site is not known. Postcode location may not be suitable.
Receptor	Damage functions for receptors	Historic claims record Cat model EP curves Multi-Coloured-Manual (MCM) curves for caravans Expert knowledge	Depth-damage curve	Commercial sensitivity of Cat model damage functions. Limited caravan specific data. Generalisation of curves.

Summary results

Whilst a full discussion of the results from the case study developed in Speight (2013) is not provided here, some of the key outputs from each component can be used to highlight the strengths, and further development areas, of the multi-scale flood risk assessment framework.

Firstly, despite adopting the rejection sampling methodology recommended by Keef et al (2013) to represent the distribution of times individual gauges are the most extreme in the network and using proportion of sites rather than total number of sites (Keef *et al*. 2009a), the assessment of spatial dependencies was shown to be an artefact of the network structure. The development of a gauged network consisting of suitable fluvial and tidal sites with long concurrent data sets close to the caravan sites of interest led to a network of 12 core gauges with 31 years of data being used for the analysis.

Using the summary spatial dependency measure, N(p), used by Keef *et al*. (2009a)for the expected proportion of gauges over a given threshold (u_p) given that conditioning gauge is above the threshold. Where X and Y are standardised daily mean flow and skew surge from the simulated event set, *j* is an index for the conditioning gauge and #{j ∈ Δ) is a count of the number of gauges in the network (Δ) where these condition are true (noting that Δ excludes the conditioning site which will always be above the given threshold);

$$N(p) = \frac{E\left(\#\{j \in \Delta: Y_j > u_p\}|\, X_j > u_p\right)}{\#\{j \in \Delta\}}$$

Eq.16

shows that areas where the gauged density is higher, such as in North Wales, display stronger spatial dependence (Figure 3).

Since coastal gauges were shown to be more correlated with other coastal gauges, having fewer coastal gauges in the network also made the dependence between coastal gauges appear lower. Using subsets of Y, by separating coastal and fluvial gauges or using the percentage of gauges within each cluster rather than the absolute value helped overcome this problem. However when developing a nested multi-site model the influence of the network structure on the result should be considered.
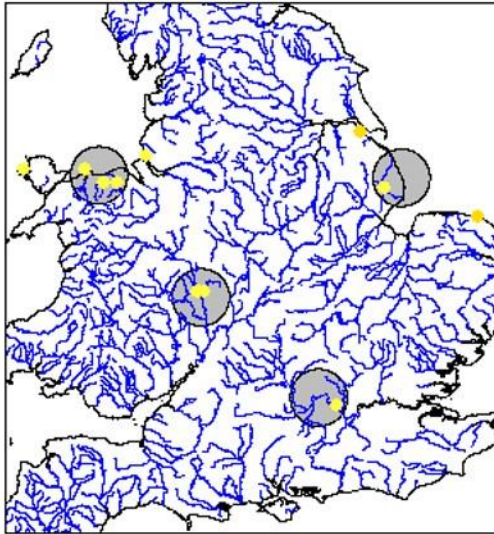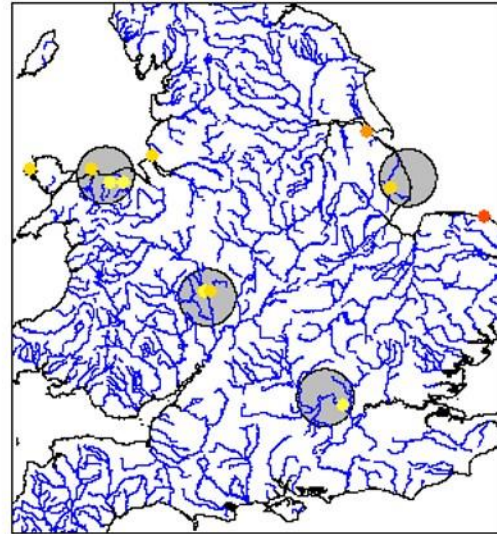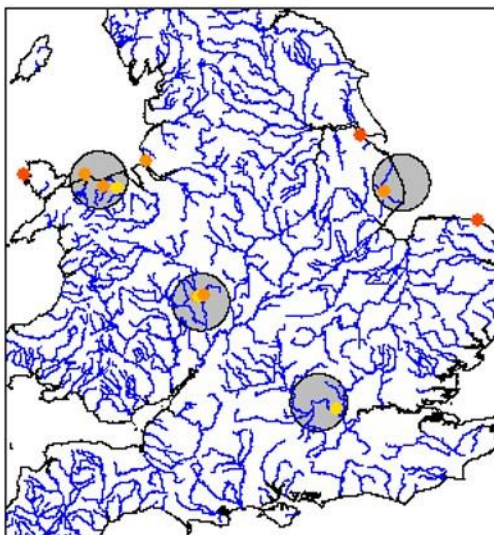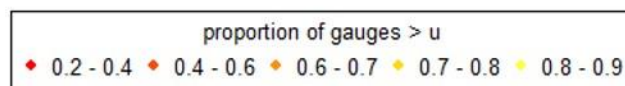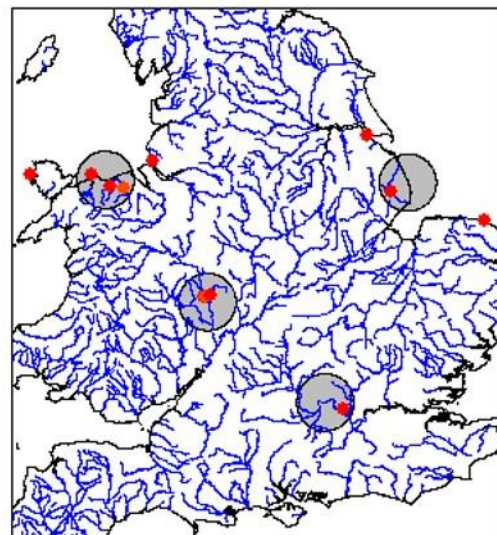
Figure 3 Spatial conditional dependency maps: expected proportion of gauges in Y > u | X > Q99

The grey circles are the risk clusters. Each dot represents a fluvial or coastal gauge. The thresholds (u) are the 0.75, 0.9, 0.95 and 0.99 quantiles of Y calculated from the simulated data set.

Flood defences are considered as the most significant pathway component in the systems model. One of the major uncertainties when modelling this component is available data on crest height. Using survey data for coastal defences in Lincolnshire a methodology was developed which used a Markov

model to represent spatially plausible variations in crest height compared to the long section mean which is often used in system risk models. To check the suitability of the methodology 1000 simulations of the 59km of Lincolnshire crest height used by Speight (2013) were made (to improve clarity, the first 500 are shown in Figure 4). The simulated mean and standard deviation for each defence type are given in Table 2 which shows a reasonable fit to the observed values with only the mean variation of simulated dune and embankment crests showing a difference of 5% compared to the observed data. This method provides a simple and effective means of incorporating plausible variation in crest height for coastal flood defences. It is particularly useful where only limited crest height data are available, the data quality is known to be poor for example if taken from LiDAR data, or the data is of a large spatial resolution which is often the case in systems risk modelling.
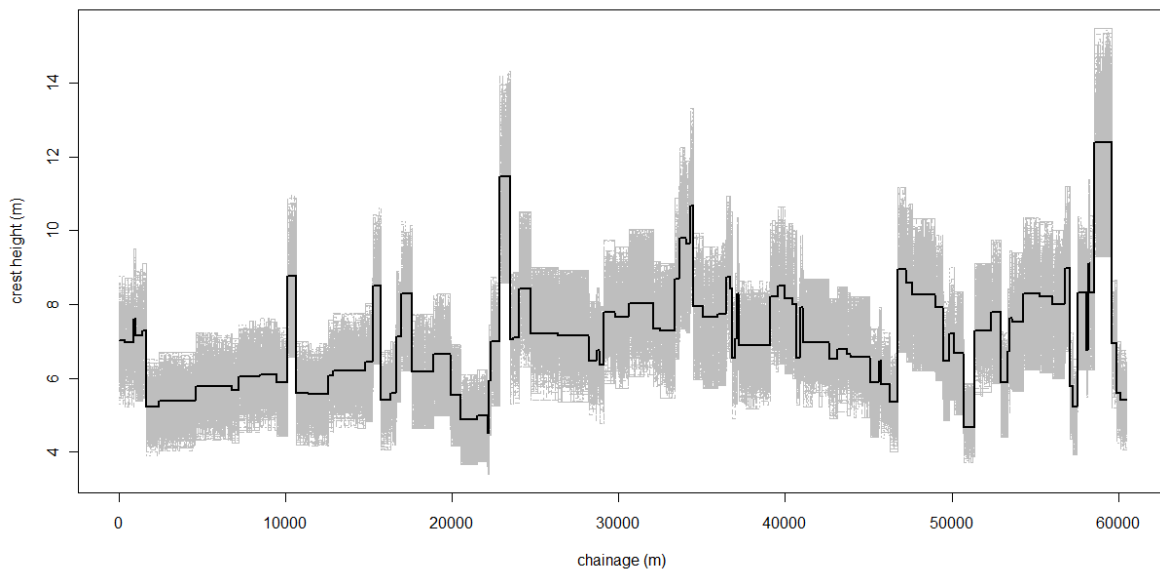

Figure 4 500 plausible simulations of varying crest height for the Lincolnshire coast compared to the long section mean

Table 2 Distribution parameters of the observed magnitude of crest height variation compared to the long section mean compared to 1000 plausible simulations for coastal defences in Lincolnshire

Defence type	Observed variation (proportion of crest height)		Simulated variation (proportion of crest height)	
Wall	Mean: 0.04	Std Dev: 0.10	Mean: 0.04	Std Dev: 0.09
Embankment	Mean: 0.00	Std Dev: 0.09	Mean: -0.05	Std Dev: 0.09
Dune	Mean: -0.05	Std Dev: 0.11	Mean: -0.00	Std Dev: 0.08

At the site scale the case study clearly identified the need to understand details about the receptor vulnerability to flooding. For caravans in particular this represents a significant problem as there is very limited data available to construct depth-damage curves and basic data about their exact location, their individual value and the floor level within the insurance portfolio is not known. To account for this the sensitivity of the damage calculation to the shape of the depth-damage curve was

tested, for example Figure 5 illustrates the effect of raising the floor level from 0m to 1m. The nested modular structure means that sensitivity tests such as this can be used to develop understanding of the receptor location, value and vulnerability at the site scale within a larger scale framework.
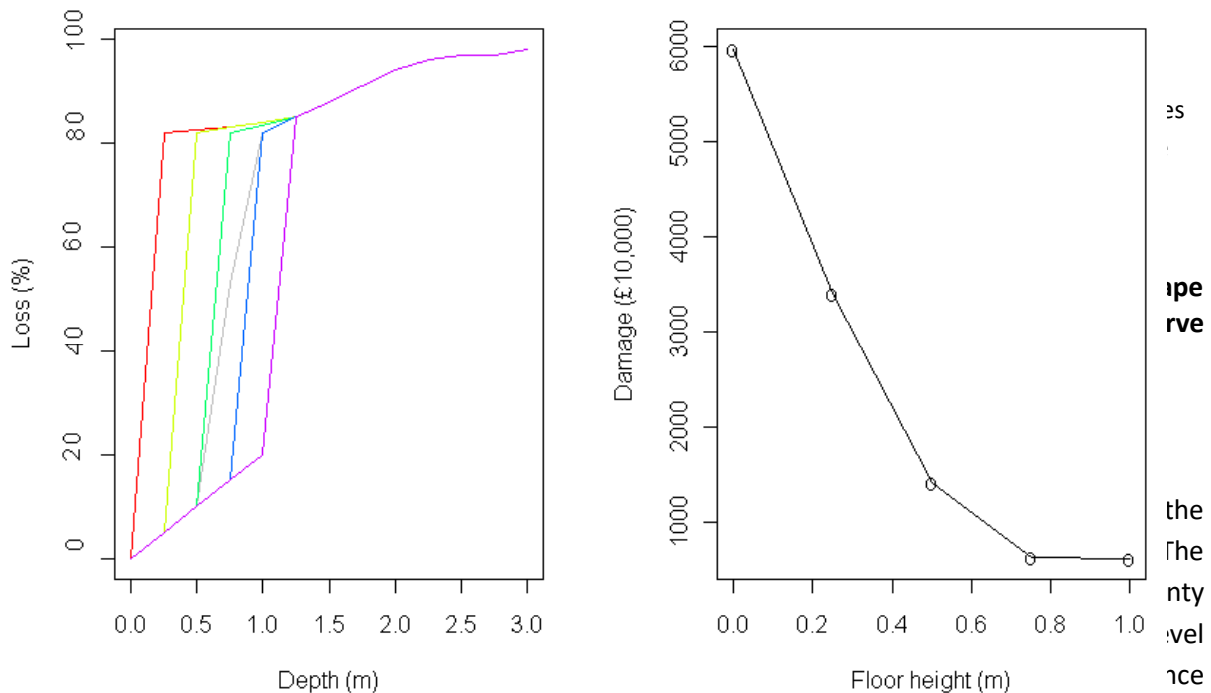
Probability (AEP) (1 in 75 year) event at individual locations as this is the threshold at which they are typically obliged to provide insurance cover. However an assessment of the system uncertainty at a 1.3% (AEP) event was deemed largely irrelevant due to the dominance of the flood defences in controlling whether damage occurs. This means that the uncertainty at a 1.3% AEP event is approximately zero as coastal defences are very unlikely to breach for events below a 0.5% AEP event. Once the event magnitude increases above the design standard of the flood defences there is an abrupt increase in uncertainty due to the difference between zero damages if defences do not breach and very high damages if they do. In light of these difficulties a qualitative review of the contribution to uncertainty from each component was made based on three criteria: whether the uncertainty is explicitly considered in the systems risk method for example through sensitivity analysis or methods to account for data uncertainty such as the defence crest height simulation approach, whether the contribution to the uncertainty in the individual component is significant and whether the contribution to the system uncertainty in significant. The reasoning behind this assessment is detailed in Table 3.

Discussion and Conclusions

The framework outlined in this paper illustrates the development of a full systems based model that is achievable at a national scale by focusing on particular areas of interest using a novel nested framework. The cross cutting analysis has drawn together state of the art methods to explicitly consider each risk component. More detailed presentation of the component elements will be the subject of forthcoming papers. The result is a transparent and flexible systems risk framework that

incorporates the critical spatial and temporal dependencies in flood risk at a range of scales through a conditional dependence model.

A key focus of the research was to demonstrate the use of a robust statistical model of extremes to provide a solid foundation for process based understanding and decision making. This is challenging due to the different requirements of statisticians and modellers and the end users making decisions. These challenges can be split into three areas; data availability at required locations, data quality, and communication of assumptions and results. Making the best use of available data and state of the art methods whilst being explicit about the sources of uncertainty is essential for well informed decision making and to identify where further research would have the greatest impact to reduce overall uncertainty.

The framework presented in this paper is similar to that used in Cat models. It is proposed that insurance companies could use this type of model alongside Cat models to help understand the particular issues and vulnerabilities within their portfolios, thereby helping to mitigate against the relative lack of information about modelling assumptions within Cat models and limited ability for end users to modify the model parameters. For example, if the risk model shows that the portfolio is particularly vulnerable to defence breaching and it is known that the Cat model in question is weak in its consideration of breaching, then this knowledge could be used to inform the portfolio management.

The development of a fully risk-based model of this type has wider benefit outside the insurance industry, particularly for applications involving spatial dependence at multiple scales as, for example, identifying the risk across a network of electricity sub-stations and identifying which sites within the network are the most vulnerable. The model structure is flexible; the conditional dependence model can be linked to other extreme weather events and the pathway and receptor components can be adapted for different end users and to incorporate new and developing analysis methods.

In developing and testing the framework a number of areas were identified which would help improve modelling capability and data availability so that systems based models such as this can be used with a consistent, and practical, level of uncertainty. Primarily this is based around data availability for example during post event data collection and on flood defence construction. Where it is acknowledged that there are limited data this should be explicitly included in the modelling methodology for example by incorporating a physically realistic spatially varying crest height into the defence failure methodology. There is also a need to consider how recent improvements in the understanding or individual components of risk can be included in a full systems model. Much work recently has focused on the representation of spatial and temporal correlations in sources. Equal focus is required to investigate how the recent improvements in the understanding of flood defence failure mechanisms can be included in systems risk models. Advances in computational ability mean more complex methods could be used in a risk based framework. However, questions still need to be answered about the suitable scale of application and the availability of detailed data to support increased complexity. The biggest challenge for the future however is how to assess the uncertainty in systems risk models and communicating this to end users to enable informed decision machining. The move towards increasing openness between Cat modellers, academia and end users (for example the Oasis (2014) framework) offers a promising potential to develop useable systems risk models with realistic uncertainty bounds and an appropriate weighting of effort in representing all stages of the system.

Word count: 6016

 References

Allsop, W., Kortenhaus, A., and Morris, M. (2007) Failure mechanisms for flood defence structures, FLOODsite. Report Number: T04-06-01.

Apel, H., Thieken, A. H., Merz, B., and Bloschl, G. (2004) Flood risk assessment and associated uncertainty. *Natural hazards and earth systems*, 4: 295-308.

Apel, H., Thieken, A. H., Merz, B., and Bloschl, G. (2006) A probabilistic modelling system for assessing flood risks. *Natural Hazards*, 38: 79-100.

Apel, H., Merz, B., and Thieken, A.H (2009) Influence of dike breaches on flood frequency estimation. *Computers and Geoscience*, 35(5): 907-923.

Buijs, F.A., Hall, J.W., Sayers, P.B. and Van Gelder, P.H.A.J.M. (2009) Time-dependent reliability analysis of flood defences. *Reliability Engineering & System Safety*, 94(12): 1942-1953.

CEH (1999) *Flood Estimation Handbook*, Centre for Ecology and Hydrology, Wallingford.

CH2MHILL (2014) ISIS technical summary [online]. http://www.isisuser.com/isis/documents/other/ISIS_Technical_Summary.pdf [accessed 1 November 2014]

Coles, S. and Tawn, J. (2005b) Bayesian modelling of extreme surges on the UK east coast. *Philosophical Transactions of the Royal Society A - Mathematical Physical and Engineering Sciences*, 363(1831): 1387-1406.

Dawson, R., Hall, J., Sayers, P., Bates, P. and Rosu, C. (2005) Sampling-based flood risk analysis for fluvial dike systems. *Stochastic Environmental Research and Risk Assessment*, 19(6): 388-402.

Dawson R.J., Speight L., Hall, J.W., Djordjevic, .S, Savic, D. and Leandro, J. (2008) Attribution of flood risk in urban areas. *Journal of Hydroinformatics*, 10(4): 275-288

De Moel, H.,Van Alphen,J. and Aerts, J. (2009) Flood maps in Europe - methods availability and use. *Natural Hazards Earth Systems Sciences*, 9: 289-301

Dixon, M. J. and J. Tawn (1997) Estimates of extreme sea conditions, Final Report: Spatial analyses for the UK Coast. Proudman Oceanographic Laboratory, Internal Document 112.

Environment Agency (2014) Flood map for planning [online], http://apps.environment-agency.gov.uk/wiyby/37837.aspx [accessed 1 November 2014]

European Parliament Council (2009). Directive 2009/138/EC on the taking-up and pursuit of the business of insurance and reinsurance (Solvency II).

Glenis, V., McGough, A., Kutija, V., Kilsby, C. and Woodman, S. (2013) Flood modelling for cities using cloud computing, *Journal of Cloud Computing: Advances Systems and Applications*, 2(7)

Gouldby, B., Méndez, F.J., Guanche, Y., Rueda, A. and Mínguez, R. (2014) A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis, *Coastal Engineering*, 88: 15-26

Gouldby, B., Sayers ,P., Mulet-Marti, J., Hassab, M. and Benwell, D. (2008) A methodology for regional-scale flood risk assessment. *Proceedings of the Institution of Civil Engineers-Water Management*, 161(3): 169-182.

Grossi, P. and H. Kunreuther, Eds. (2005) Catastrophe modelling: a new approach to managing risk. Huebner International Series on Risk, Insurance, and Economic Security. New York, Springer Science Business Media, Inc.

Gusman, P. (2008) Insurers overly focused on Cat models, says expert [online]. NU Online News Series. http://www.propertyandcasualtyinsurancenews.com/cms/nupc/Breaking%20News/2008/04/17-CATRELIANCE-pg. [accessed 26 February 2009]

Hawkes, P. J., Gouldby, B. R., Tawn, J.A. and Owen, M.W. (2002) The joint probability of waves and water levels in coastal engineering design. *Journal of Hydraulic Research*, 40(3): 241-251.

Hall, J. W., Dawson, R. Sayers, P., Rosc, C., Chatterton, J. and Deakin, R. (2003) "A methodology for national-scale flood risk assessment." Proceedings of the Institution of Civil Engineers-Water and Maritime Engineering 156: 235-247.

Hall, J.W., Manning, L.J. and Hankin, R.K.S. (2011) Bayesian calibration of a flood inundation model using spatial data. *Water Resources Research*, 47(5).

Heffernan, J. E. and Tawn, J. A. (2004) A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 66: 497-530.

Keef, C., Svensson, C., and Tawn, J. (2009a) Spatial dependence in extreme river flows and precipitation for Great Britain. *Journal of Hydrology*, 378: 240-252.

Keef, C., Tawn, J. and Svensson, C. (2009b) Spatial risk assessment for extreme river flows. *Journal of the Royal Statistical Society Series C- Applied Statistics*, 58(5): 601-618.

Keef, C., Tawn, J., and Lamb, R. (2013) Estimating the probability of widespread flood events, *Environmetrics*, 24 (1): 13–21

Keef, C., Lamb, R., Tawn, J., Dunning, P., Batstone, C. and Lawless, M. (2011). Spatial coherence of flood risk – technical methodology report, Environment Agency. Science Report – SC060088/R1.

Lamb, R., Crossley M. and Waller, S. (2009). A fast two-dimensional floodplain inundation model. *Proceedings of the Institution of Civil Engineers: Water Management*, 162(6): 363-370

Lamb, R., Keef, C., Tawn, T., Laeger, S., Meadowcroft, I., Surrendran, S., Dunning, P. and Batstone, C. (2010) A new method to assess the risk of local and widespread flooding on rivers and coasts. *Journal of Flood Risk Management*, 3(4): 323-336.

Liang, Q. H. (2010) Flood Simulation Using a Well-Balanced Shallow Flow Model. *Journal of Hydraulic Engineering-ASCE*, 136(9): 669-675.

Lloyd's. (2006) Cat models cannot replace common sense.   Retrieved 17th June 2011, from http://www.lloyds.com/News-and-Insight/News-and-Features/Archive/2006/05/Cat_models_cannot_replace_common_sense.

Merz, B., Kreibich, H., Thieken, and Schmidtke, R. (2004) Estimation uncertainty of direct monetary flood damage to buildings. *Natural Hazards and Earth System Sciences*, 4(1): 153-163.

Merz, B. and Thieken, A. H. (2009) Flood risk curves and uncertainty bounds. *Natural Hazards*, 51(3): 437-458.

Merz, B., Kreibich, H., Schwarze, R. and Thieken, A. (2010) Review article 'Assessment of economic flood damage'. *Natural Hazards and Earth System Sciences*, 10(8): 1697-1724.

Morris, M., Hassan, M., Kortenhaus, A. and Visser, P. (2009a). Breaching processes: a state of the art review, FLOODsite. Report Number: T06-06-03.

Morris, M., Kortenhaus, A. and Visser, P. (2009b). Modelling breach initiation and growth, FLOODsite. Report Number: TO6-08-02.

Muir-Wood, R. and W. Bateman (2005). Uncertainties and constraints on breaching and their implications for flood loss estimation. *Philosophical Transactions of the Royal Society A - Mathematical Physical and Engineering Sciences*, 363: 1423-1430.

Neal, J., Fewtrell, T., Bates, P. and Wright. N. (2010). A comparison of three parallelisation methods for 2D flood inundation models. *Environ. Modelling and Software*. 25(4): 398-411.

Netherlands Ministry of Public Works, IHE-Delft, et al. (2010). CRESS - Coastal and River Engineering Support System.

Oasis (2014) Loss Modelling Framework [online], https://connect.innovateuk.org/web/oasis-open-access-catastrophe-model/overview [accessed 1 November 2014]

Penning-Rowsell, E., Johnson, C., Tunstall, S., Tapsell, S., Morris, J. Chatterton, J. and Green, C. (2005). The benefits of flood and coastal risk management: a handbook of assessment techniques, Middlesex University Press.

Penning-Rowsell, E., Priest, S., Parker, D., Morris, J., Turnstall, S., Viavattene, C., Chatterton, J., Owen, D. (2013) Flood and coastal erosion management: A manual for economic appraisal, Routledge

Pitt, M. (2008) The Pitt Review - Learning Lessons from the 2007 floods. London, Cabinet Office

Sangal, B. P. (1983) Practical methods of estimating peak flow. *Journal of Hydraulic Engineering-Asce*, 109(4): 549-563.

Speight, L (2013) A multi-site methodology for understanding dependencies in flood risk exposure in the UK, PhD Thesis, Newcastle University

Svensson, C. and Jones, D. A. (2002) Dependence between extreme sea surge, river flow and precipitation in eastern Britain. *International Journal of Climatology*, 22(10): 1149-1168.

Svensson, C. and Jones, D. A. (2004) Dependence between sea surge, river flow and precipitation in south and west Britain. *Hydrology and Earth System Sciences*, 8(5): 973-992

Vorogushyn, S., Merz, B. and Apel, H. (2009) Development of dike fragility curves for piping and micro-instability breach mechanisms. *Natural Hazards and Earth System Sciences*, 9(4): 1383-1401.

Vorogushyn, S., Merz, B., Lindenschmidt, K-E .and Apel, H. (2010) A new methodology for flood hazard assessment considering dike breaches. *Water Resources Research*, 46.

Wyncoll, D. and Gouldby, B. (2013) Integrating a multivariate extreme value method within a system flood risk analysis model. *Journal of Flood Risk Management*

Table 3 Qualitative assessment of relative contribution to uncertainty from individual system components

Component	Explicitly considered by the methodology?	Potentially large impact on		Overall	Recommendation
		component uncertainty	system uncertainty		
Marginal model for extremes at each gauge	Yes	Yes	Yes	Acceptable - Limited data is acknowledged as a problem in all statistical modelling of extremes and the methods used have been shown to be suitable for this application.	Update model as data record increases. Priority: Low
Conditional dependence model of extreme spatial and temporal dependency	Yes	Yes	Yes		
Conversion of daily mean flow to peak flow and hydrographs	No	Yes	No	Acceptable – Although the peak flow and hydrograph generation methodology is simplified it does not affect the large scale multi-site dependence structure.	Further work is required on the ungauged site transfer method at the local scale between gauged locations and site of interests. Priority: Low unless very detailed local analysis is required
Conversion of still water level and waves to total onshore water level	No	Yes	No	Acceptable - Given the requirements for a multi-site model which can be applied using readily available data and larger uncertainties from simulation of extreme events. Well established methods are have been used.	Could be improved using mode detailed onshore wave modelling for coastal sites of interest if required. Priority: Low unless very detailed local analysis is required
River routing with hydraulic model	No	No	No	Acceptable - Given the requirements for a multi-site model which can be applied using readily available data and larger uncertainties from simulation of extreme events.	Could improve using more detailed hydraulic modelling for fluvial sites of interest if required. Priority: Low unless very detailed local analysis is required

Component	Explicitly considered by the methodology?	Potentially large impact on		Overall	Recommendation
		component uncertainty	system uncertainty		
Floodplain inundation modelling	Yes	No	No	Acceptable: For large events that cause high damages minor differences in floodplain inundation are unlikely to significantly affect the results.	Suitable inundation models should be chosen based on the level of detail required at specific site locations and the computational time available. Priority: Medium
Representation of flood defence crest heights	Yes	Yes	Yes	Significant - The correct representation of crest height is very important as in most cases the damage will be zero if the defence does not breach so it is important to correctly incorporate low (potential failure) points.	Use of the best available data is recommended. When detailed survey information is not available the spatially varying crest height simulation should be used to incorporate potential low points. Priority: High
Representation of flood defence fragility	No	Yes	Yes	Significant – As above the correct representation of failure probability is very significant so it is important to correctly incorporate weak (potential failure) points. This is particularly important for coastal sites.	There is a need to incorporate the available research into defence failure into system models. The spatially varying crest height methodology could be adapted for spatially varying defence reliability. The analysis should be extended to other failure modes where these are considered important to the sites of interest. Priority: High
Calculation of damage	No	Yes	Yes	Significant -The lack of data on caravan vulnerability is a significant uncertainty in the system risk model. Due to the high vulnerability of caravans to relatively low water depths their location on the site is important and could result in the different between zero damage and total write off.	More data is required on flood receptors. Following an event data should be collected on flood damages to caravans and depth-damage curves re-assessed. Priority: High