A multi-site approach to risk assessment for the insurance industry

L. SPEIGHT[1], J. HALL[1], C. KILSBY[1] & P. KERSHAW[2]

1 *School of Civil Engineering and Geosciences, Cassie Building, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK*
linda.speight@ncl.ac.uk

2 *Catlin Insurance, 3 Minster Court, Mincing Lane, London, EC3R 7DD, UK*

Abstract Using a case study of one insurance company's exposure from static caravans we propose a methodology for flood risk assessment at multiple sites nested within a national framework. Following a source-pathway-receptor approach, a mixture of statistical and physically based methods is used in a systems based model which incorporates the most important random processes associated with flood damage. Within the system meteorological inputs are modelled statistically using a conditional dependence model; the water level, floodplain inundation and damage calculations are deterministic; and the impact of flood defence failure is considered probabilistically. The methodology explicitly couples spatial dependencies between variables affecting flood risk at national and local scales. The output is greater understanding of risk, and the associated uncertainties, which can be used to inform decision making.

Keywords risk assessment; multi-site; national framework; fluvial flood; storm surge; flood defences; uncertainty; insurance

INTRODUCTION

The focus of this paper is the development of a multi-site based approach to flood risk assessment which considers all important random factors influencing flood risk. Over the past decade the importance of taking a risk based approach to flood management has been realised and new methodologies have been developed covering different national and international contexts (Hall *et al*. 2003; Apel et *al*. 2006). However as yet there is no integrated risk model that incorporates all contributing factors and the spatial and temporal dependencies between them.

The methodology development is based on a case study of insurance exposure to static caravan sites in the UK. The UK is unique in that flood insurance is widely available to property owners. In order to continue to provide insurance, companies needs to be able to accurately model and understand risk. The standard means of assessing risk is through Catastrophe (Cat) models. These are complex process based models which use a mixture of numerical and statistical methods making it difficult for the end user to fully understand the underlying processes. Therefore it is useful for an insurance company to increase its understanding of risk outside of the Cat modelling framework, for example to identify which aspects of the process contribute the most uncertainty and therefore need detailed consideration in the pricing process, or if there is potential to incorporate mitigation measures for certain risks.

Although designed for an insurance application, the strength of this methodology is that it can be used for robust flood risk assessment for any site specific locations nested within a larger framework. Three significant issues have been identified as particularly relevant to this case study and to the wider development of flood risk assessment in the UK;

1. The probability of large events affecting multiple geographical areas of the UK.
2. The sensitivity of risk estimates to assumptions made about flood defences.

3. The interactions between fluvial and coastal process.

The methodology proposed in the remainder of this paper seeks to address these issues.

NESTED MULTI-SITE MODEL

The model structure is based on a multi-site method. This allows more detail to be included at sites of interest while maintaining the large scale structure and allowing consideration of large scale dependencies between input events. If the sites of interest are selected carefully, it is assumed that understanding gained from the multi-site method will be applicable across the wider network.

The interest in this project is in risk to an insurance portfolio covering approximately 80% of static caravans in the UK worth around £4,600 million. Five risk clusters have been selected which typify particular issues within the portfolio and provide broad spatial coverage across the UK, these are shown in Figure 1.

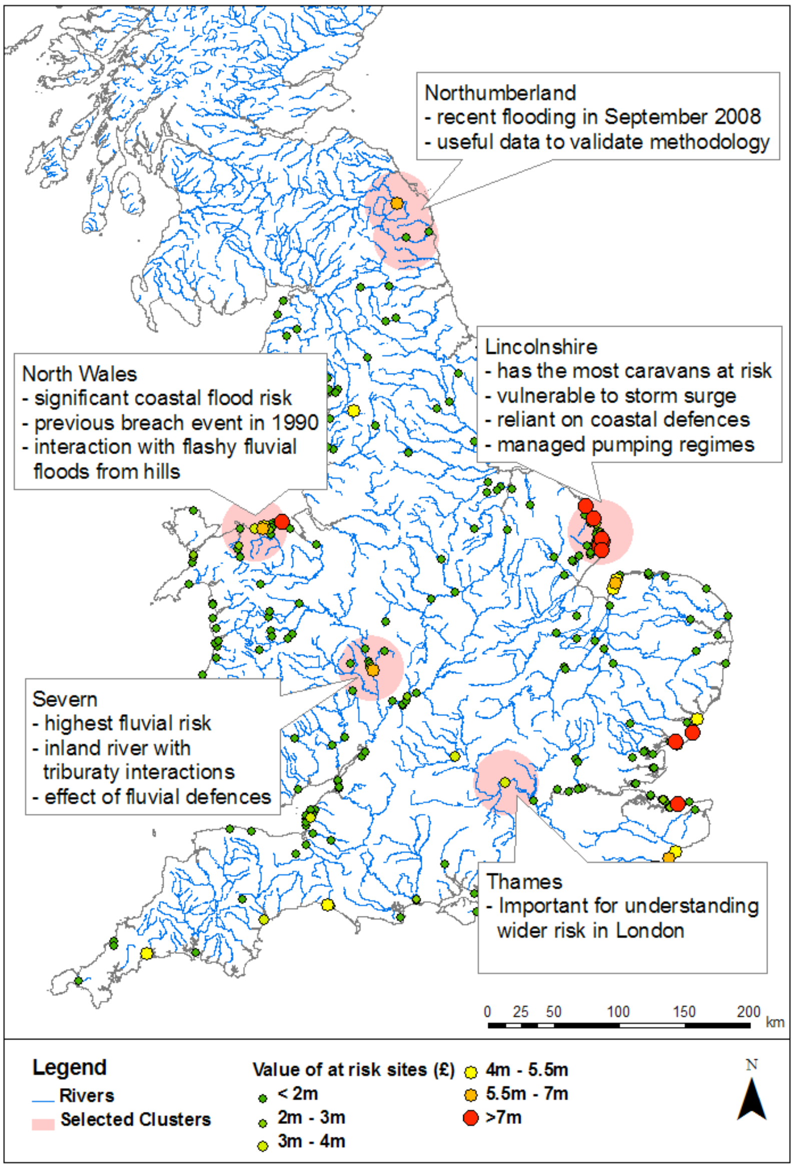


Fig. 1 Selected risk clusters

RISK BASED METHODOLOGY

The risk model takes the form of a source-pathway-receptor model. This allows explicit consideration of all random factors which affect the risk process. The sources are meteorological inputs, initially river flows and sea levels, other meteorological sources can be included at a later date. The pathways are the catchments, rivers and defence systems the flood water passes through and the receptors are static caravans distributed across the UK. The various stages of the process are shown in Figure 3.

Sources

For each event, i, daily mean flows are simulated at all the gauges in the network using a conditional dependence model. The vector of flows across the network, X_i, represents a large scale spatial event with probability, $P(X_\mathrm{i})$.

The local dependencies are then addressed at each site or risk cluster, s. The daily mean flow at the nearest gauge to the site is labelled x_i. The procedures for transferral to ungauged sites are discussed later. The daily mean flow is then converted into a peak flow and hydrograph, Q_i, such that;

$$Q_i = f(x_i, k_s) \tag{1}$$

where k_s is the ratio of daily mean flow to flood peak specified for each site.

Pathways

Figure 2 shows a flood defence system with n defence section, d_1 to d_n, characterised by their construction type and standard of protection (SOP). Any one of the defences can fail in one or more locations resulting in inundation of the floodplain. Within each defence section it is assumed that the crest height and defence reliability will gradually vary along the length of the defence. The degree of variation depends on the defence type and condition such that;

$$c_{j,i} = f\ (defence\ type, defence\ condition)$$

(2)

$$r_{j,i} = f\ (defence\ type, defence\ condition)$$

(3)

Where c and r are the simulated crest height and resistance to load of any given defence, j. The vectors of crest height and reliability state for the whole system are referred to as C_i and R_i accordingly.

The water level throughout the modelled reach, L_i, or at a particular defence section, $l_{j,i}$, is simulated using the hydraulic model for the specified flows and defence crest heights;

$$L_i = f(C_i, Q_i)$$

(4)

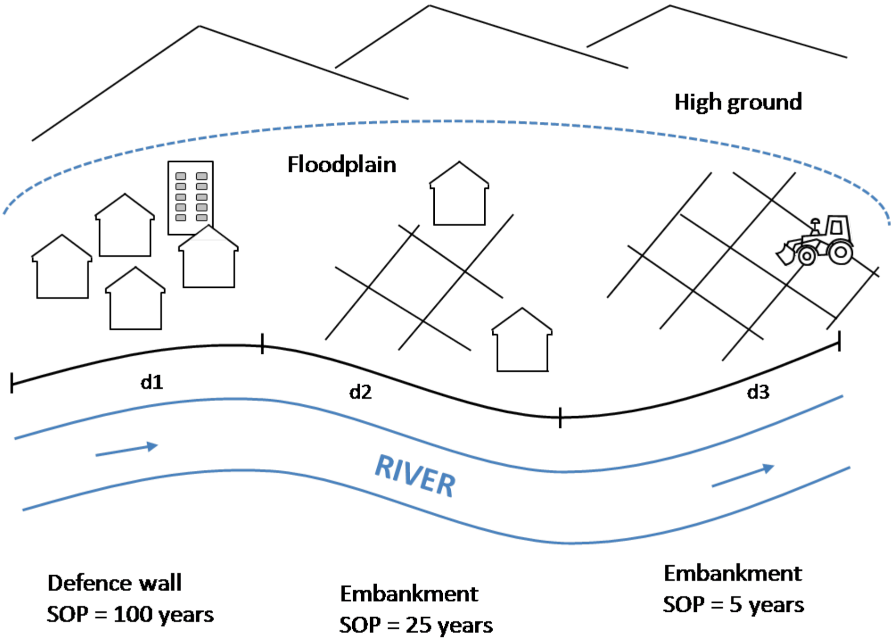


Fig. 2 Schematisation of flood defence system

Overtopping of defences is considered deterministically through the use of the hydraulic model conditional on the simulation of the defence crest height. The probability of overtopping, $P(OT_i)$ depends on the probability of the extreme event, however due to the variation in crest height $P(OT_i|L_i = SOP) \neq 1$. 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.

$$OT_i = f(C_i, L_i) \tag{5}$$

$$P(OT_i) = P(X_i) \tag{6}$$

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

The probability of failure for any given event, $P(D_{j,i})$ is conditional on the loading variable, amount of overtopping, and defence reliability such that for any given defence;

$$P(D_{j,i}) = f(l_{j,i}, ot_{j,i}, r_{j,i}) \tag{7}$$

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

$$P(D_{j,y,i}) = f(l_{j,y,i}, ot_{j,y,i}, r_{j,y,i}) \tag{8}$$

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

$$P(D_{j,1,i}) \times P(D_{j,2,i}) \times P(\bar{D}_{j,3,\dots,t,i}) \tag{9}$$

The probability of defence failure for any given load is defined using a fragility curve, giving the conditional failure probability $P(D_j|l_j)$, in this case the load is defined as the water level at defence j. The unconditional failure probability, $P(D_j)$ is given by;

$$P(D_j) = \int_0^{\infty} p(l_j) P(D_j\,|l_j)\,dl \tag{10}$$

Where $p(l_j)$ is the probability density function of the water level, l, at defence j.

The water level at any given defence section is dependent on previous upstream breaches. To account for this an iterative sampling procedure is adopted. Firstly the hydraulic model is run assuming no failures to give $L_{i,\overline{D_1,\dots,D_n}}$. Successive sequences of one or more defence breaches, D_{ss}, are sampled based on the conditional failure probabilities at each point in the defence $P(D_{j,y}|L_{i,\overline{D_1,\dots,D_n}})$ until $P(D_{ss}|L_{i,\overline{D_1,\dots,D_n}}) \rightarrow 0$. 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 probability of each failure sequence $P(D_{ss})$,is the product of the failure probability at each breach point $P(D_{j,y,i})$, 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 $p(D_{ss})$ for each event. Each breach in $D_{ss,i}$ is assigned a maximum width such that;

$$W_{ss,i} = f(\,L_i, defence\ type, floodplain\ size\ and\ shape, timing\ of\ breach) \tag{11}$$

Receptors

For each event, X_i, for the specified defence state variables, C_i, R_i, $D_{ss,i}$, and $W_{ss,i}$, the 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 caravan specific depth damage curves these floodplain depths are converted to loss estimates at each site and summed together to give the loss across the portfolio conditional on the event.

The methodology explicitly couples large scale spatial dependencies in extreme events with local scale dependencies in defence crest height, defence reliability and defence loading. The novel aspects of the methodology are in the simulation of extreme events and consideration of the defence system. These aspects will be discussed further in the sections below.

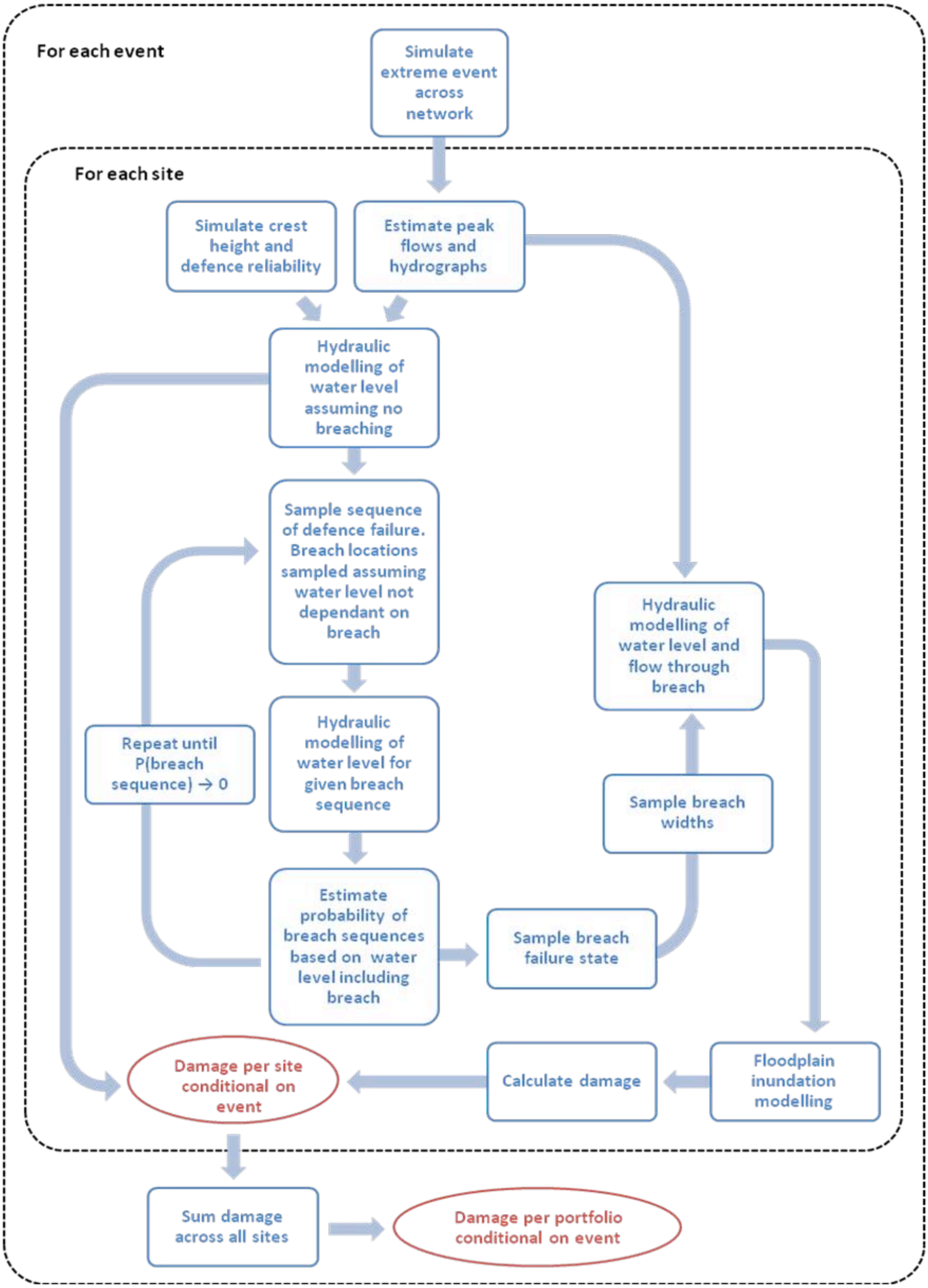


Fig. 3 Overview of risk methodology

SIMULATION OF EXTREME EVENTS

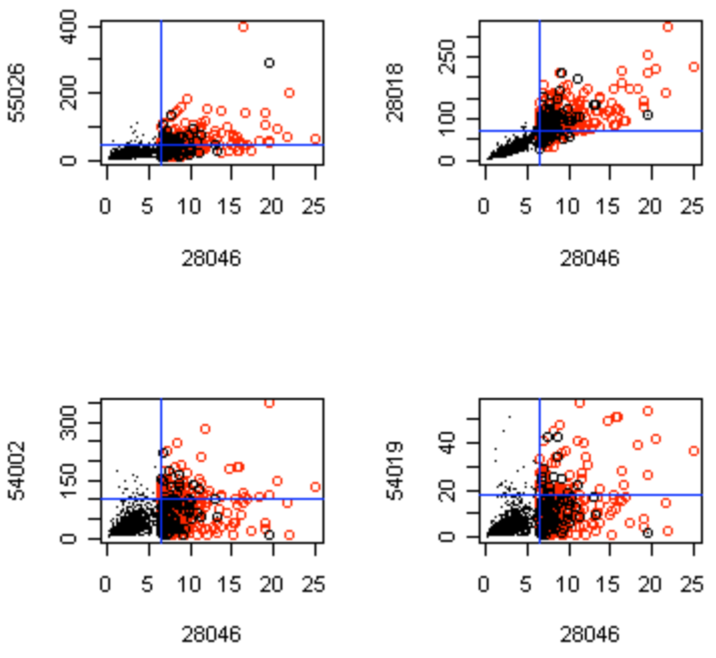
Within the modelling process, meteorological inputs are modelled statistically using the conditional dependence model proposed by Heffernan and Tawn (2004) fitted to daily mean flow data as demonstrated by Keef *et al* (2009a; 2009b; 2009c). The method is a flexible model that is able to account for both asymptotically dependent and independent data and can model the changes in dependences as events get more extreme. The model describes the distribution of **H** | G, given that G is large for independent, identically distributed data without missing values (Keef *et al*. 2009a). This methodology allows simulation of extreme events which maintain the correct spatial correlation of events across the network and can also be used to simulate events more extreme than those observed in the data. The model is fitted to each conditional gauge G in turn for each dependent gauge H, therefore G is a single variable and **H** is a set of variables.

The method separates the marginal and dependence characteristics and models them separately. The marginal characteristics are modelled through a standard general parento distribution. The dependence model takes the form of a multivariate semi parametric regression model;

$$\boldsymbol{H} = \boldsymbol{a}(g) + \boldsymbol{b}(g)\boldsymbol{Z} \tag{12}$$

Where ***a*** and ***b*** are normalising functions such that *a* describes the overall strength of dependence between the two variables, and *b* describes how the dependence changes with the threshold. **Z** is a vector of residuals between each modelled pair of gauges. Further details of how to fit the model are provided by Keef *et al* (2009a; 2009b; 2009c). An example of flow data simulated from the model is shown in Figure 4.

For multivariate and time lagged data the dependencies are modelled non-parametrically through the residuals. By using a conditional dependence model of this type it is therefore possible to simulate multi variable extreme events without the need for large scale weather simulation.



Legend

Blue lines are gauge thresholds
. are data under the threshold
0 are observed data
0 are simulated data

Gauge references are SRC numbers

The data is simulated conditional on gauge 28046

Fig. 3 Example of flow data simulated from conditional dependence model

To date the Heffernan and Tawn methodology can only simulate flows at gauged sites. For site based assessment it is necessary to use the model to produce estimates of extreme events at ungauged sites. Keef *et al* (2009a) use a routine which interpolates flow estimates between gauges and to ungauged catchments. This methodology serves as a useful demonstration of what is possible at a national scale however it suffers from several assumptions especially in the volume and timing of flow from tributaries. The advantage a nested multi site approach rather than a national model is that more detail can be used at the sites of interest. Therefore it is possible to use the recommended methodologies for ungauged sites in the Flood Estimation Handbook (FEH, Robson and Reed 1999), modified for use with daily mean flow data.

Fitting the model to daily mean flow data is problematic as this is of limited use for flood risk assessment however due to the complexity of the conditional dependence model it is not feasible to use continuous or peaks over threshold data on a national scale. The daily mean flows are therefore converted to flood peaks after the dependence model has been fitted using a deterministic relationship based on catchment characteristics.

Once a peak flow has been established a hydrograph is needed for input to the hydraulic model. There are various methods outlined in the FEH for borrowing hydrograph shapes (Robson and Reed 1999). For this initial development of the methodology these methods are appropriate. However to take account of all random processes affecting flood risk, the shape of the hydrograph needs to be considered. For example flashy events with large flood peaks might cause overtopping of a flood defence whereas longer duration events are more likely to cause defence failure due to piping. Vorogushyn (2009) and Apel *et al* (2006) presents a promising methodology to this ends which uses historic flood hydrographs to generate a set of representative hydrograph shapes to sample in the risk model.

CONSIDERATION OF DEFENCE FAILURE

Flood defence system reliability

Unlike the previous defence system discretisation methods of RASP (Risk Assessment of flood and coastal defences for Strategic Planning, (Hall *et al.* 2003; Gouldby *et al.* 2008)) and Apel *et al* (2006; 2009), there is no limit on the length of each defence section. These previous methods were based on the assumption that for any given defence section the whole section would respond to loading in the same way. This assumption was only valid for defence sections of less than 500m, above this defence resistance parameters are independent (Centre for Civil Engineering Research and Codes. and Technical Advisory Committee on Flood Defence. 1990). In the proposed methodology the assumption of fully dependent resistance is lifted therefore there is no requirement to restrict defence section length. This methodology also allows for incorporation of random sources of defence failure for example following the 1953 flooding in the UK, one cause of failure along the Ouse was reported to be a fallen bush exposing roots and allowing a preferential route for erosion to initiate (ICE 1953)

A full risk based approach should consider all possible sources of failure which includes the probability that a defence will fail below its design standard. Reports by FLOODsite (2009) list numerous defence failure methods, the most common of which is overtopping, followed by breaching caused by overtopping. A defence may also breach without overtopping due to internal erosion caused by pore water, most commonly due to piping. Correctly modelling breaching is acknowledged to contribute the most critical sources of uncertainty in flood risk quantification (Muir-Wood and Bateman 2005).

Fragility curves are use to characterise the reliability of defences. The RASP based methods used 1D fragility curves where the probability of failure is dependent on the water level. Apel *et al* (2006, 2009) have expanded on this approach and used 2D fragility curves including the duration of high water levels, this is particularly important for failure due to internal pore water.

Overtopping

The degree of overtopping at any given location depends on the crest height which may deviate from the design crest height along the length of the defence. Realisations of the potentially varying crest height can be simulated along the defence or between known surveyed points using an order autoregressive process. This allows for a gradually varying crest level where by the elevation z_y depends on the previous value using a first order autoregressive process;

$$z_y = \varphi_1 z_{y-1} + \alpha_y + \beta_y \tag{13}$$

Where φ_1 is the autoregressive parameter, α_y is a random number generated from a normal distribution with mean zero and a specified standard deviation for each given defence, and β_y is the design or surveyed crest height. Figure 5 shows example simulations using different values for the standard deviation of the error and the autoregressive coefficient. Defence A is an old, poorly maintained earth embankment showing large changes in crest height. Defence B is a well maintained or harder defence, such as a sea wall, with very little variation from the design crest level. The process can also be modified to take account of varying design crest levels.

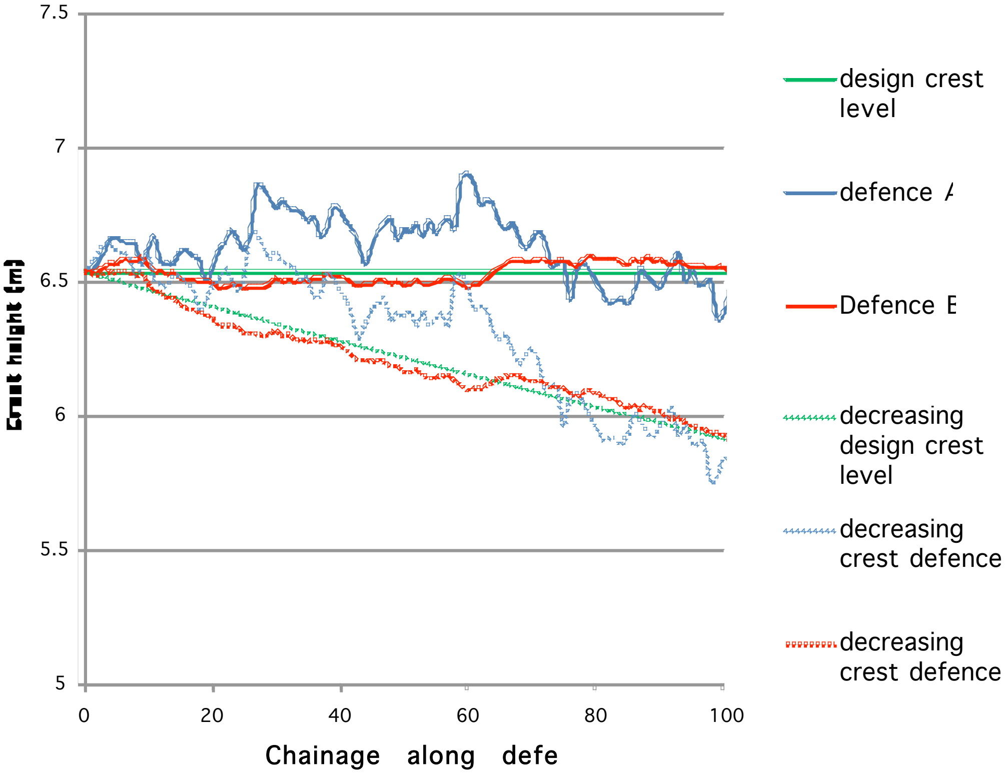


Fig. 4 Example simulations of gradually varying crest height

Breaching

There are two key factors to understanding breaching, firstly the structural and loading conditions required to initiate a breach and secondly how the breach develops over time.

The erodability of a flood defence is dependent on the construction type, materials used and surface protection in the form of vegetation or harder materials such as rock armour (FLOODsite 2009). The load the defence is subject to is influenced by potential upstream breaching as the reduction in flood peak caused by floodplain storage behind upstream breached defences can result in fewer breaches downstream (Apel *et al* . 2009).

In a similar way to the crest height, the strength of the defence is not uniform along its entire length, for example some areas may be subjected to animal burrowing or poor drainage. An autoregressive process can also be used to account for these semi-random factors. A standard fragility curve is established for each defence section. The autoregressive process is then generated to modify the fragility curve along the length of the defence. This process can be carried out as many times as necessary based on the number of strength factors to account for. For any given water level the probability of failure varies based on the crest height as shown by the P(failure) line in Figure 6. The fragility curve is then modified by the autoregressive defence strength function. For example although the defence crest is relatively low between 60 and 100m, the defence strength in this section is stronger than average and therefore the varying probability of failure is lower. The final stage is to sample breach points based on the varying failure probability, these are shown as red triangles in Figure 6.

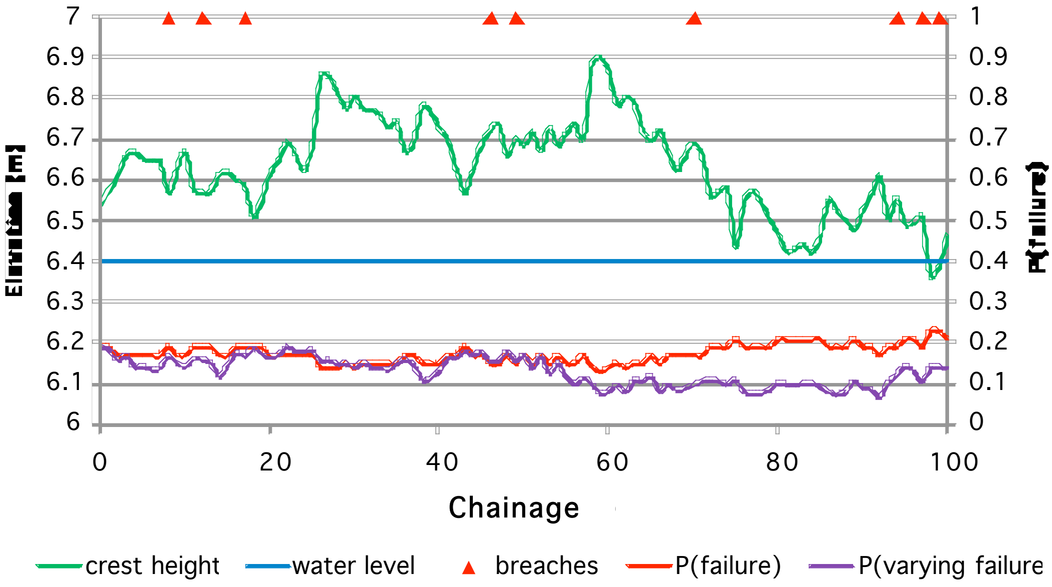


Fig. 5 Example simulations of spatially varying defence strength

Once a breach has initiated it grows laterally and vertically at a rate dependant on the material properties and pore water, with complex interactions occurring between these different factors (FLOODsite 2009).

As well as the defence type and construction, the hydraulic gradient across the breach is also critical in determining the breach size and depth. Muir-Wood and Bateman (2005) found that the largest breaches formed in wide flat floodplains where flood water quickly travels away from the breach location. In the same way neighbouring breaches are in competition with each other, since water from the initial breach reduces the hydraulic gradient across the defence erosion rates for subsequent breaches will be lower introducing further temporal dependencies in the breaching mechanisms.

One of the major limitations of the RASP methodology is that is assumes that all defences are subjected to the same loading at the same time and will respond to loading in the same way. The iterative sampling procedure for breach initiation and widths allows incorporation of the temporal and spatial dependencies between flood breaches. Breach width is sampled based on the timing of breach, floodplain size and shape, defence construction, and load (Equation 11).

The rate of growth is less critical than the breach width (Safe Coast 2008). RASP assumes that all breaches instantly erode to their full width and Apel *et al* (2006, 2009) assume the breach is complete in one hour. The sensitivity to breach growth rate will therefore depend on the time step of the hydraulic model. Commercial hydraulic modelling packages now include the function to specify a breach growth rate so a suitable rate will be defined.

INFORMING DECISION MAKING

The risk method described above is a flexible component based model. This means that the inputs into any component can be changed to investigate uncertainty and/or to allow exploration of management options. For the insurance industry this is important as it enables a company to price its premiums competitively without exposing itself to undue risk.

The insurance premium for any given portfolio is calculated as (Grossi and Kunreuther 2005);

$$Insurance\ premium = AAL + Risk\ Load + Expense\ Load \tag{14}$$

The risk load is the surplus required by the insurance company before they are happy to cover the loss, and *AAL* is the Average Annual Loss which is the expected loss each year assuming the company insures the portfolio with no changes over an infinite number of years. Therefore to correctly price the policy, an insurance company needs to be able to accurately predict the *AAL* and understand what factors contribute to the risk load.

As identified above there is limited data available on flood defence condition and performance. The flexible model structure allows for uncertainties of this type to be explicitly included within the risk assessment and for the significance of these uncertainties to be identified. For example if analysis reveals that correctly modelling the defence crest height is the most significant factor in determining flood risk the insurance company would either need to put more resources into modelling this component or increase their risk load accordingly. Alternatively premiums could be reduced by reducing the *AAL* through mitigation measure, for caravans this might include the installation of floatation devices or raising the floor level of the caravans. By allowing sampling of multiple damage curves to account for these factors the impact of reducing the vulnerability to flood damage can be assessed.

Outside of the insurance industry this methodology provides a robust framework for decision making by explicitly modelling all important random factors associated with flood risk and allowing these factors to vary within realistic bounds.

CONCLUSIONS

This paper has outlined a new risk based methodology that explicitly considers all important random processes associated with flood risk. In particular the proposed flood defence failure methodology provides significant improvements over existing risk based failure methods by removing the artificial restrictions on defence section length and the number of breaches per defence section. Another improvement is in the incorporation of spatial and temporal dependencies at multiple scales. The benefits of the nested approach combined with the spatial dependence modelling of extreme events means that an insurance company is able to simultaneously review their exposure to national risk as well as look in more detail at particular locations. This also means that particular high risk locations can easily be identified.

The development of a fully risk based model of this type is of wider benefit outside of the insurance industry particularly for applications involving spatial dependencies at multiple scales for example identifying the risk across a network of electricity sub stations and identifying which sites within the network are the most vulnerable. The flexible nature of the methodology means it can be adapted to look in more detail at particular aspects of interest and can be used to gain an understanding of the importance of different factors in contribution to the uncertainty involved in flood risk estimation.

The model structure is flexible and through the conditional dependence model can be linked to other extreme weather events. It is hoped to also be able to link storm surge and wind storm risk into the model in the future.

Acknowledgements Linda Speight's PhD project is funded by EPSRC and Catlin Insurance.

REFERENCES

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

Apel, H., Thieken, A. H., *et al*. 2006. A probabilistic modelling system for assessing flood risks. *Nat. Hazards,* **38**, 79-100.

Centre for Civil Engineering Research and Codes and Technical Advisory Committee on Flood Defence. 1990. Probabilistic Design of Flood Defences. Gouda. CUR.

FLOODsite. 2009. Integrated flood risk analysis and management methodologies. Retrieved 1st April 2010, from www.floodsite.net.

Gouldby, B., Sayers, P., *et al*. 2008. A methodology for regional-scale flood risk assessment. *Proc. Inst. Civ. Eng. Water Management,* **161**(3), 169-182.

Grossi, P. and Kunreuther, H. 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.

Hall, J. W., Dawson, R., *et al*. 2003. A methodology for national-scale flood risk assessment. *Proc. Inst. Civ. Eng. Water and Maritime Engineering,* **156**, 235-247.

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.

Institution of Civil Engineers (1953). *Conference on the North Sea Floods of 31 January / 1 February 1953: A collection of papers presented at the institution,* London, The Institution of Civil Engineers.

Keef, C., Lamb, R., *et al*. 2009a. Spatial coherence of flood risk – Methodology report. *Science Report – SC060088/SR.,* Environment Agency

Keef, C., Svensson, C., *et al*. 2009b. Spatial dependence in extreme river flows and precipitation for Great Britain. *J. Hydrology,* **378**, 240-252.

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

Muir-Wood, R. 1999. Employing catastrophe loss modelling to price and manage European flood risk. *Global Change and Catastrophe Risk Management: Flood Risks in Europe* Retrieved 17th April, 2009, from http://www.iiasa.ac.at/Research/RMS/june99/papers/muirwood.pdf.

Muir-Wood, R. and Bateman, W., 2005. Uncertainties and constraints on breaching and their implications for flood loss estimation. *Phil. Trans. R. Soc. A. Mathematical Physical and Engineering Sciences*, **363**, 1423-1430.

Robson, A. and Reed. D. , 1999. *Flood Estimation Handbook, Vol 3: Statistical procedures for flood frequency estimation*. Institute of Hydrology.

Safe Coast. 2008. Comparison between different flood risk methodologies: action 3B report. Retrieved 17/3/2010, from http://www.safecoast.org/editor/databank/File/a3breport_finalversion%5B1%5D.pdf.

Vorogushyn, V. 2009. Analysis of flood hazards under consideration of dike breaches. *PhD*. University of Potsdam