Rapid evidence assessment of the use of emulators in flood risk analysis

Date: June 2025

Version: FRS19218a/TR

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

We help people and wildlife adapt to climate change and reduce its impacts, including flooding, drought, sea level rise and coastal erosion.

We improve the quality of our water, land and air by tackling pollution. We work with businesses to help them comply with environmental regulations. A healthy and diverse environment enhances people's lives and contributes to economic growth.

We can't do this alone. We work as part of the Defra group (Department for Environment, Food & Rural Affairs), with the rest of government, local councils, businesses, civil society groups and local communities to create a better place for people and wildlife.

Published by:

Environment Agency
Horizon House, Deanery Road,
Bristol BS1 5AH

www.gov.uk/environment-agency

Further copies of this report are available from our publications catalogue: www.gov.uk/government/publications or our National Customer Contact Centre: 03708 506 506

Email: enquiries@environment-agency.gov.uk

Authors: HR Wallingford

Keywords: modelling, emulators, coast, surface water, river, flood risk analysis

Research contractor: HR Wallingford. Howbery Park, Wallingford, Oxfordshire, OX10 8BA

Environment Agency's Project Sponsor: Ben Lukey, Project Executives: Jonathan Boyes, Shirley Greenwood, Project Managers: Sue Manson, Ian Perkins, Jenny Kirwan-Evans, Sally Brown

Research 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. Our vision is that the nation is recognised as a world leader in researching and managing flooding and coastal change.

The Joint Programme is overseen by Defra, the Environment Agency, Natural Resources Wales and the Welsh Government on behalf of all risk management authorities in England and Wales.

You can find out more about our current science programmes at [Research at the Environment Agency](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 flood and coastal erosion risk management work, please contact fcerm.evidence@environment-agency.gov.uk.

Dr Robert Bradburne	Julie Foley
Chief Scientist	**Director of Flood Strategy and Adaptation**

Contents

Executive summary

Emulators are increasingly being used in flood risk management, particularly for extending the observation record. In the UK, we have an extensive record of observations, but the record of high and extreme flows is more limited, both inland and at the coast. Emulators could be an option to address this data sparsity.

To do this, we need to better understand emulators and their use. They can be viewed as a "black box" and need more scrutiny and testing before they can be widely adopted within the Environment Agency and other Risk Management Authorities.

This evidence review looked to better understand how emulators can be used in flood risk activities to provide an evidence base on the available approaches. To this aim, we synthesised current knowledge on different flood modelling emulators using a systematic approach. We followed the rapid evidence assessment (REA) process, which has been developed by the Department of Environment, Food and Rural Affairs (Defra) to provide a rigorous, transparent, and exhaustive synthesis of evidence from scientific literature.

An REA follows a similar process to a systematic review but introduces some restrictions so that it does not take as long or cost as much. The process involves definition of research questions, development of a protocol, a systematic search for evidence, screening of the evidence, extraction of evidence into a systematic map, critical appraisal of the evidence, synthesis and, finally, drawing of conclusions.

The primary question that we addressed in this project is: what is the evidence for the successful application of emulators in the context of analysis of present and future flood risk?

The evidence shows emulators are used extensively within flood risk modelling and analysis. The main findings that have been identified are that:

- there are different ways of developing emulators dependent on the required predictive limits and data available
- emulators are currently used more widely within coastal modelling, than hydrological or hydraulic modelling
- there has been limited demand for emulators in the "consequence" components of the source-pathway-receptor-consequence model; models are already computationally efficient and therefore emulators are not required
- it is important to consider the accuracy and uncertainty of any emulator in context of the overall analysis
- there is significant evidence of emulators being combined with evolutionary optimisation algorithms (artificial intelligence) in flood risk analysis - examples include:
    - flood defence failure during Hurricane Katrina (Kingston and others, 2011)
    - calibrating a MIKE 11 rainfall or run-off model (Khu and others, 2004)

    - selecting an optimum flood mitigation strategy (Woodward and others, 2013)
- emulators have been used to assess climate change impacts, both in terms of the whole source-pathway-receptor model as well as just the “source” component

During the review, it was apparent that there was no readily available industry guidance document that covers the application of emulator techniques to flood risk modelling and analysis. Given the widespread application of emulators, the review recommended the Environment Agency, along with other Risk Management Authorities, consider developing appropriate guidance. This could cover:

- the range of available techniques, with a discussion of the pros and cons of each
- a standard template for developing an emulator including input and output parameters, design points, validation, and error metrics
- basis for error acceptance, considering the overall application and including model chain uncertainties
- a description of modelling problems that are or are not suitable for the application of emulators
- limits of emulators application
- an overview of software libraries, capabilities, and verification

1 Introduction

1.1 Motivation

Emulators are increasingly being used in flood risk management, particularly for extending the observation record. In the UK, we have an extensive record of observations, but the record of high and extreme flows is more limited, both inland and at the coast. Emulators could be an option to address this data sparsity.

To do this, we need to better understand emulators and their use. They can be viewed as a “black box” and need more scrutiny and testing before they can be widely adopted within the Environment Agency and other Risk Management Authorities.

This evidence review looked to better understand how emulators can be used in flood risk activities to provide an evidence base on the available approaches. To this aim, we synthesised current knowledge on different flood modelling emulators using a systematic approach. We followed the rapid evidence assessment (REA) process, which has been developed by the Department of Environment, Food and Rural Affairs (Defra) to provide a rigorous, transparent, and exhaustive synthesis of evidence from scientific literature.

1.2 Rapid Evidence Assessment

Defra has worked with partners to develop methods for conducting evidence reviews that are designed to make the most of existing research investment. One of these methods is a REA, which follows a systematic review approach but is less resource intensive, while maintaining rigour and transparency. Detailed REA guidance is provided in Collins and others (2015).

Typically, REAs consist of a series of steps common to the systematic review process, but the aims and objectives of the study are defined so that it can be completed on a relatively short timescale. While an REA should be as rigorous and exhaustive as possible, restrictions can be applied to reduce the time and expense of delivery. This flexibility means that although the conclusions can be translated into practice in a reasonable timeframe, they are not as robust as results of a systematic review.

For this study, we have broadly followed the methodology of Collins and others (2015), which describes in clear terms the necessary steps of a REA, along with the roles and responsibilities of all parties involved. The main parties are the review team, who undertake the review, and the steering group, a group of technical experts who guide and assist the review team where necessary to ensure the outputs of the REA meet the needs of end users.

Following this methodology, the main tasks of the REA review team are to:

- define the primary and secondary questions to be addressed in the study

- establish the protocol outlining their approach to the study and agree it with the steering group
- complete a search for relevant evidence
- screen the evidence, retaining only evidence relevant to the research questions
- systematically extract evidence relating to the research questions into a systematic map
- critically appraise the evidence, evaluating it in terms of relevance to the research questions and robustness of the methodology applied
- synthesise the evidence to produce summary information describing the volume and characteristics of the evidence base
- draw conclusions from the results of the evidence review
- communicate the evidence review findings

This report describes the findings of the project.

1.3 Background to emulators

Collins and others' approach (2015) identifies the development of a conceptual model as an important aspect in aiding the evidence review and providing an overview of interactions between different components.

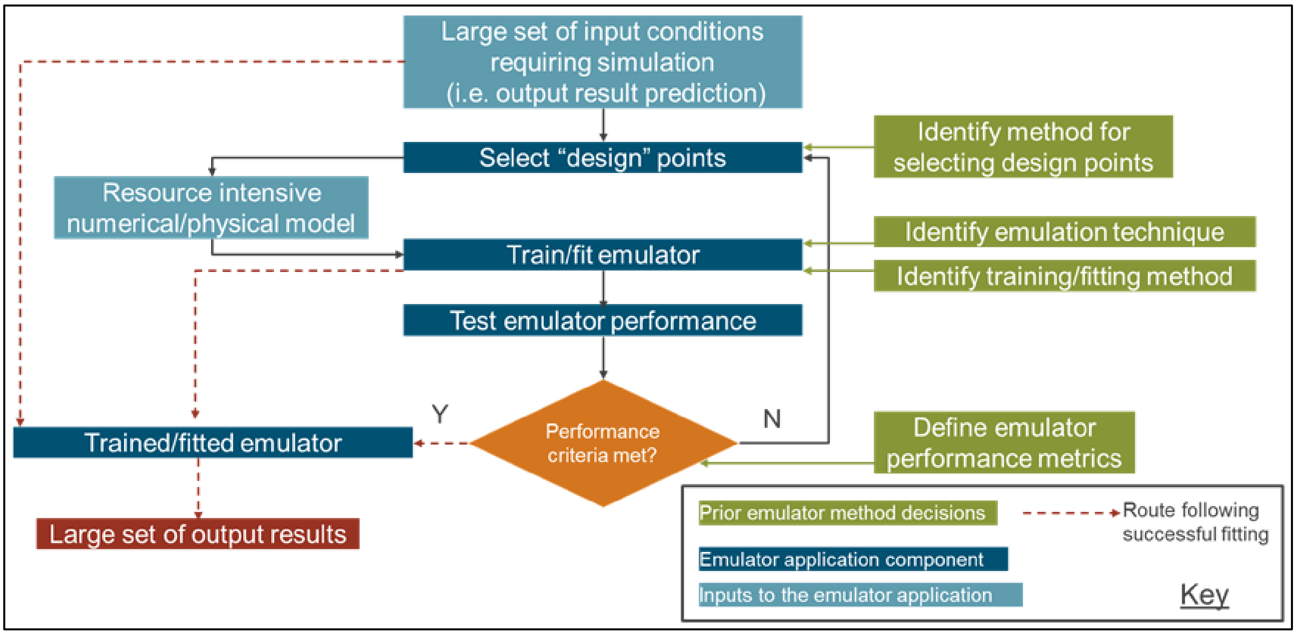


Figure 1: Conceptual diagram of components of an emulator (key: green boxes - prior emulator method decisions; royal blue boxes - emulator application component; pale blue boxes - inputs to the emulator application; red dashed line – route following successful fitting)

Figure 1 shows the main elements within the framework of emulators in a flow chart. This indicates the prior emulator method decisions, the application content and inputs to the emulator application. These are brought together to consider whether the criteria are met and a set of possible outputs. The basic problem is the need to run many (physical process) resource intensive (either computational or human) model simulations, or even a single model simulation that requires extensive computational resource (for example, 3D numerical models of wave overtopping that solve the full Navier-Stokes equations). In practice, it is not possible to run the model for all the simulations needed. To overcome this, the model is only run for a selected set of input values. These can include:

- model boundary conditions relating to flood sources (for example, rainfall, river flow, sea level or wave conditions)
- model parameters (for example, roughness, soil moisture, bed friction)

Emulators can be, and have been, applied with the following types of data:

- stochastically generated Monte Carlo event sets
- time series data (real-time and non-real time)
- single event or boundary condition input prediction

These input values, and the corresponding output results, are referred to as the ‘design points’ (in references to designing the emulator). They are illustrated in Figure 2.

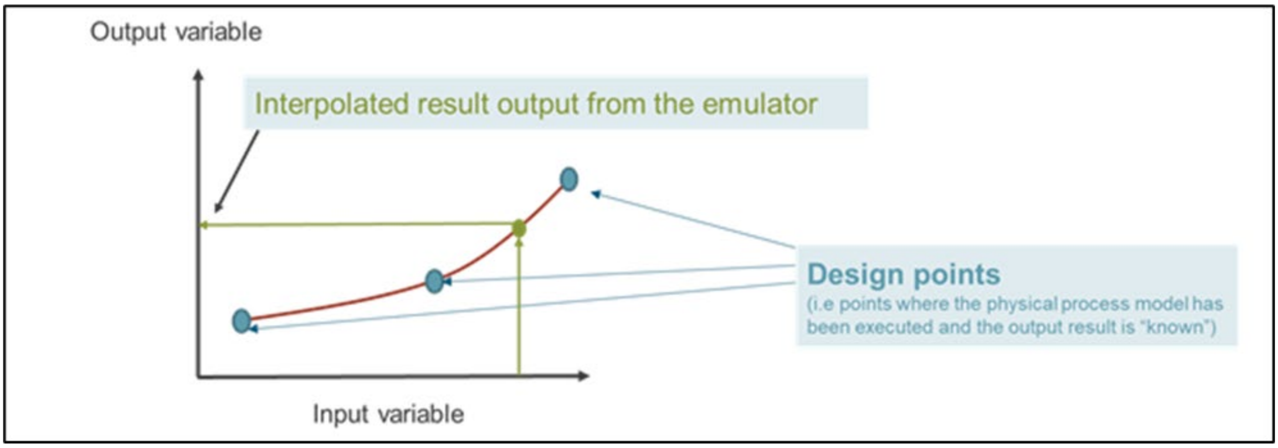


Figure 2: Conceptual illustration of design points

Figure 2 plots a graph of the output variables (y-axis) against the input variables (x-axis). It illustrates how an emulator uses the link between design points to interpolate from the input variable. The design points are points where the physical process model has been executed and the output result is ‘known’. It notes the design points where the physical processes model has been executed. ‘Design point’ in this context is unrelated to the expression used in reliability analysis (for example, Melchers and others, 1999). Moreover, it is unrelated to any aspect of the design of any structures.

Emulators, which require a fraction of the resource compared with equivalent (physical process) models, interpolate between the design points. They therefore offer efficiency savings. As the emulators are interpolating, they introduce an error. The error can be managed through appropriate strategies for selecting the design points (that is, ensuring enough are selected and that they are selected appropriately). These error management aspects are all components of the emulator application and these are shown within the ‘emulator application components’ (navy blue) boxes in Figure 1.

As emulators only interpolate, it is essential for the physical process model to remain constant for the design point simulations. Changes to the underlying physical representation of a model is not appropriate when applying emulators. In the context of probabilistic flood risk analysis there is a well-known problem relating to simulating flood inundation from breaches or failures in separate flood defence segments.

For probabilistic analysis, this means having to simulate a significant number of flood inundation simulations to represent the different breach scenarios. Because each inundation simulation requires a different physical system (that is, the flood defence system is physically different for each simulation), emulators do not offer a credible approach to solving this problem. To overcome this computational burden, flood risk assessments have a simplified physical process representation of flooding (sometimes referred to as reduced complexity, or reduced physics).

The Foresight approach for calculating climate change risk, adopted and evolved by Sayers and others (2017) is illustrated in Figure 3.

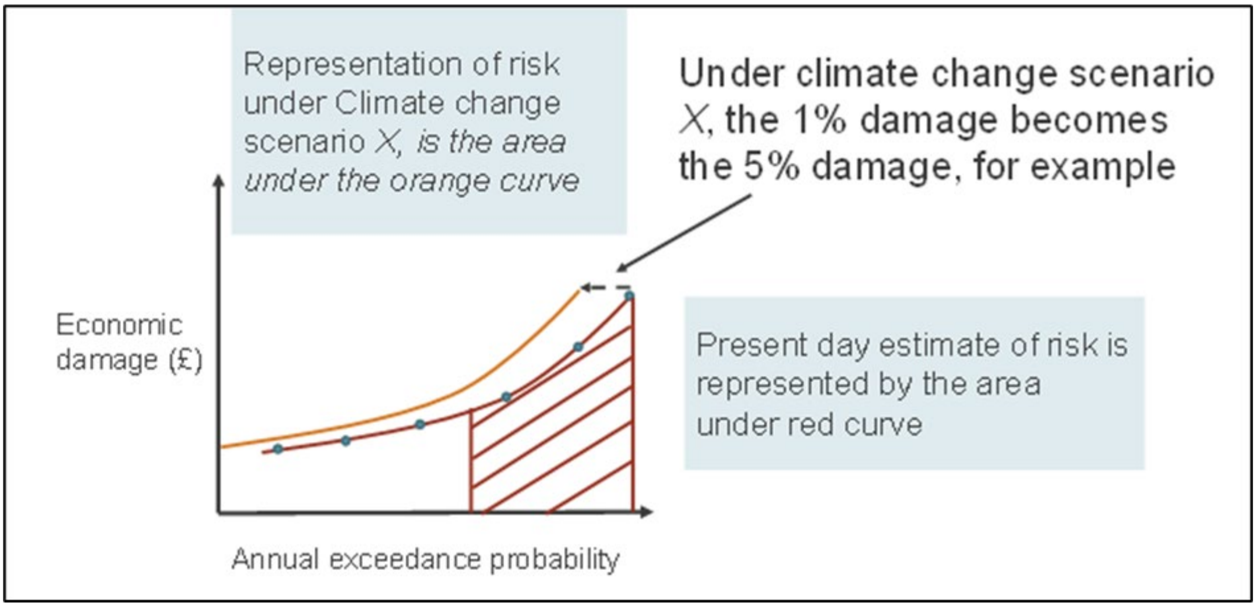


Figure 3: The Foresight Climate Change approach, evolved for use by Sayers and others (2017)

The graph in Figure 3 shows the link between annual exceedance probability (x-axis) and economic damage (in pounds on the y-axis). Representation of risk under climate change scenario ‘X’ is the area under the orange curve. Present day estimate of risk is represented by the area under the lower red curve. Under the climate change scenario, the damage increases. It shows, for example, a change in the standard of protection of a defence from 100 years (1% AEP) currently to 20 years (5% AEP) in the future. This change is used to imply an equivalent change in economic damage and a revised loss distribution (Figure 3). The revised loss distribution can be integrated to evaluate the future climate change risk. This method does not conform to the framework for emulators described in this report because:

- it does not interpolate between known design points and therefore does not directly replicate or emulate a specific model
- there is no process of training or model fitting of a model used for interpolation

For these reasons, it is not considered further in this study.

1.4 Practical application and limits of emulators

Three scenarios were identified in relation to using emulators in practice. The scenarios include developing an emulator:

- for a new application where the required predictive limits are known in advance of the design process
- for a new application where the required predictive limits are not known in advance of the design process
- based on existing data or model outputs

In the first situation, it is standard practice for the design points of the emulator to be selected to minimise the extent of any extrapolation. This is because the emulator is a statistical model with no capability to model the underlying physical processes and is specifically intended for estimation. Small deviations in terms of extrapolation may be justified in some circumstances but care is required.

This paper describes the application of a parameter space-filling algorithm, the maximum dissimilarity algorithm (MDA) (see Camus and others, 2011 for a discussion in relation to waves), that selects design points that capture the limits of the input parameter data set. This is illustrated in Figure 4 with the MDA and weighted maximum dissimilarity algorithm (WMDA).

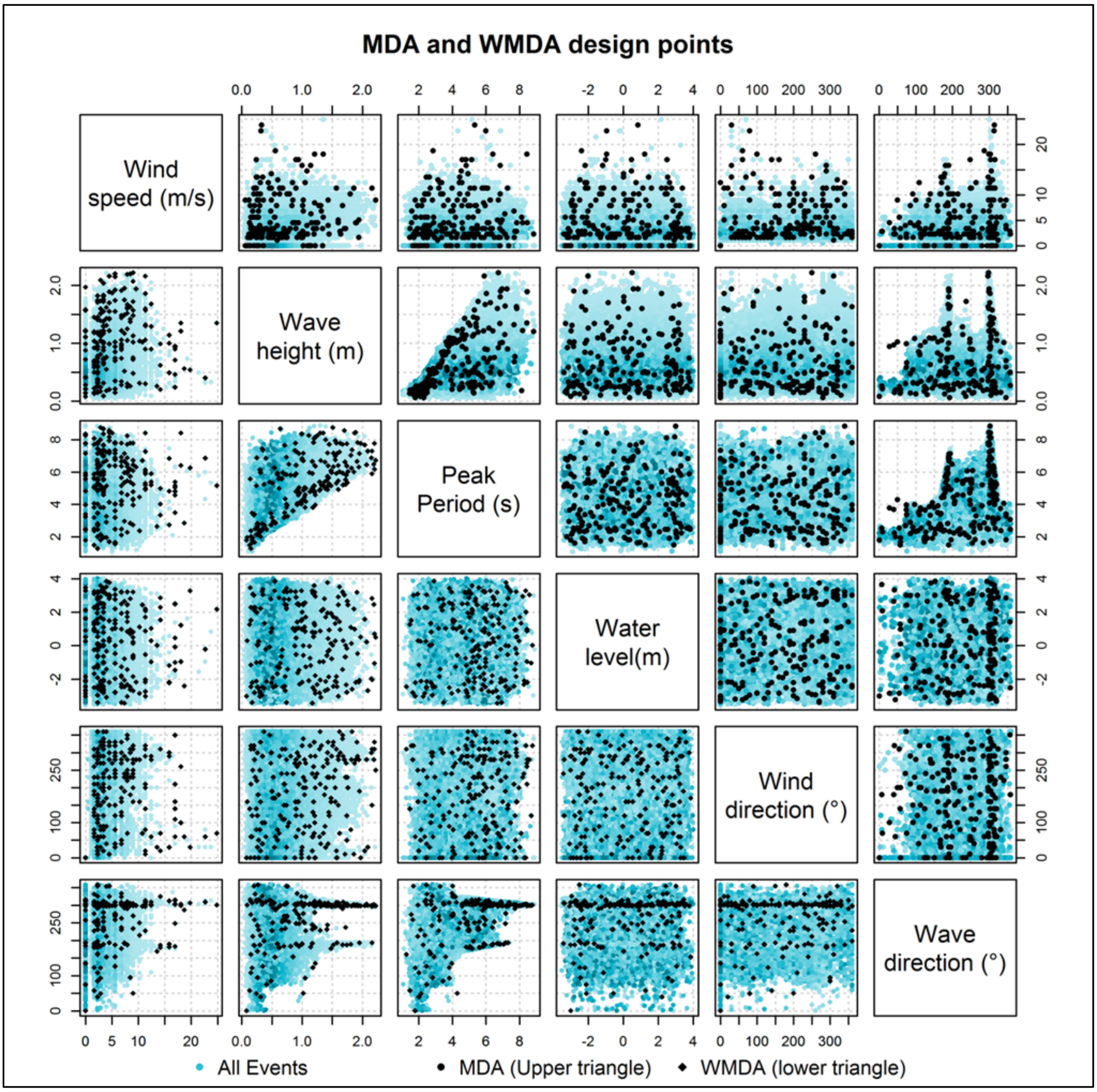


Figure 4: Illustration of a space filling algorithm to select design points (black points). Upper right triangle unweighted MDA, lower left triangle weighted MDA (WMDA) by direction. Source: Malde and others, 2016

In Figure 4 there are 30 postage stamp graphs showing a space filling algorithm being used to select design points from a wind and wave dataset. Both an MDA and WMDA algorithm are compared, with the WMDA choosing more points along dominant wave directions. It shows the underlying data (generated by a Monte Carlo simulation procedure) requiring prediction (blue) and design points selected by the algorithm (black). In some cases, it may be possible to identify areas of the input parameter space that are more important than others. In this case, a weighting can be applied to the algorithm to prioritise certain areas of the input parameter space. The figure shows a comparison between a standard MDA application (upper right triangle) and a weighted (WMDA) application (lower left triangle). The weighting has been applied to prioritise design points associated with specific wind directions (lowest row in the figure).

In other situations, however, the important region of the input parameter space may not be known at the outset when designing the emulator. There is evidence of an example of this situation, in the context of flood risk analysis, described by Kingston and others (2011). This paper describes the application of a neural network to emulate a geotechnical model that was used to explore the reliability of a New Orleans flood defence embankment that catastrophically failed during Hurricane Katrina.

In reliability analysis, the so-called ‘limit state’ defines the region that separates failure from non-failure (for example, Melchers, 1999). In the New Orleans case study, a computationally expensive geotechnical model was constructed and an emulator of this model was required. In reliability analysis, it is important for the emulator to perform well along the limit state, as the failure probability is most sensitive to errors in this region. As the location of the limit state was not known at the outset, an iterative process was implemented to select the design points. This required the application of an optimisation algorithm (genetic algorithm) that was used to search the input parameter space to identify the limit state (region between safe and failed structure). This is discussed further in section 6.1.

In the final case, where existing data are used to form the basis of the emulator, there is perhaps a greater motivation or requirement to extrapolate. There is evidence that emulators have been inappropriately applied in practise, beyond their recommended limits. This relates to the development and widespread use of a wave overtopping neural network (Van Gent, 2007). There has been a specific effort to provide guidance in relation to the limits of applying this type of model (Pullen and others, 2018). These aspects are discussed in more detail in section 6.2.

2 REA objectives

The Environment Agency has identified the use of emulators in flood risk analysis as a topic of interest. The project objective was:

"to synthesise current knowledge on emulation and emulators and their use in flood (fluvial and coastal) and rainfall extremes in the UK using a systematic approach. The project should follow the guidance presented in Collins and others (2015), which outlines the governance required, how to define key questions and how to carry out, synthesise and communicate the evidence review findings."

The term 'emulator' is often applied to a specific class of statistical models or mathematical techniques. These models are all focused on estimating a result between known values.

Other terms used to describe the same class of model include:

- metamodels
- surrogate models
- lookup tables
- response surface
- simulation library

Specific mathematical techniques are associated with this class of model. These mathematical techniques include:

- linear interpolation (often used in lookup tables or simulation libraries)
- (piecewise) polynomial interpolation
- artificial neural networks (ANNs)
- Gaussian process emulators (GPEs) (also known as Kriging)
- radial basis functions (RBF)

These techniques are widely applied to a range of different environmental models, including those relating to flood risk analysis. From now on in this document, the term 'emulator' is used to refer to this class of models.

'Emulator' has also recently been applied within the context of climate change risk analysis (Sayers and others, 2017). Here, it is distinct from the range of mathematical techniques used above and the framework shown in Figure 1. In this specific climate change context, it is used to describe a type of approach first applied on the Foresight Future Flooding project (OST, 2004). This Foresight method uses estimates of changes of the standard of protection (SOP) of flood defences to estimate changes in flood risk.

Review questions

The Environment Agency identified an initial primary question and related secondary questions to form the basis of the evidence review. It was however, noted that these were suggestions that needed considering further. These questions were reviewed at a project workshop meeting and revised questions were agreed. The initial and revised questions are provided below. The method also requires the elements of population, intervention, control and outcome (PICO) to be defined. These are also detailed in this section.

2.1 Initial questions

Primary question:

What is the evidence of the effectiveness of emulators in extending data records outside our current observation range for use in flood risk understanding?

Secondary questions:

- where do we find the golden middle between reduced simulation time and error associated with approximation?
- how do we define these ratios or proportions? - is this the same for the fluvial (and rainfall) modelling as it is in coastal modelling?
- how big does our training data set need to be? - what difference does it make to uncertainty and accuracy in results?
- is there an appetite (how much) to accept simplification errors?

2.2 Revised questions

Primary question:

What is the evidence for the successful application of emulators in the context of analysis of present and future flood risk?

Secondary questions:

- what evidence is there of successfully finding an appropriate balance between reduced simulation time and increased error or uncertainty associated with approximation and what criteria have been used to determine this balance?
- what is the evidence for successfully using emulators to predict flood sources (for example, waves, river flows), pathways (wave overtopping, breaching, flood depth, flood velocity), consequences (economic damage, loss of life) and overall flood risk (based on whole system models)?
- what is the evidence for using emulators for static versus dynamic (time-stepping) models in flood risk analysis and forecasting?

2.3 Population, intervention, control and outcome (PICO)

The PICO components used in developing the search terms for the REA are defined as:

- population – flood risk or forecasting numerical models

- intervention – application of emulator techniques
- control – application without emulator techniques
- outcome – effective application of emulator techniques

3 Evidence collation process

To undertake the REA, a literature study was carried out. All retrieved studies were assessed for relevance using the following inclusion or exclusion criteria:

- relevant subjects: flood risk management, flood risk
- geographical reference: fluvial, pluvial, coastal, regional, national and international
- climatic conditions: present day and climate change
- language: primarily English
- date: 1990s onwards

The information sources or databases used as part of the study were:

- Google Scholar
- Scopus
- Web of Science
- grey literature
- papers provided by the project steering group

3.1 Search keywords and strings

The main search terms were proposed and discussed in the initial workshop. A list of keywords was devised based on the outcome of the workshop. These have been combined into search strings and used in the literature search.

The 2 elements of the search terms are the statistical models of interest (emulators and alike), and the application field of the statistical model (for example, flood risk management). The search terms are summarised below. The search strings were constructed by joining the statistical models and the applications using an AND operator. Where equivalent terms from the same statistical model were identified, these were joined using an OR operator. This logic was also applied to the set of terms listed under each application, for example, (metamodel OR meta-model OR 'meta model') AND ('flood risk' OR 'flood hazard'). The search terms comprised:

statistical models or techniques:

- emulator: also known or referred to as statistical emulator
- metamodel: also known or referred to as meta model, meta-model
- surrogate model
- response surface methodology
- design selection or choice: also known or referred to as design point or training data selection

applications:

- general: flood risk, flood hazard

- present and future flood risk
- coastal: coastal flood, sea level, surge, wave height, wave period, wave direction, sea level, tide or tidal, surge
- fluvial: fluvial flood, river flow, discharge, streamflow, run-off, catchment
- pluvial: pluvial flood, rain or rainfall, snow or snowfall, surface water, precipitation
- flood inundation: flood inundation, flood depth, flood velocity
- defence: defence breach, defence failure, overtopping

Appendix A contains the search strings that were applied.

3.1.1 Search engines

The search strings were used for searches in Google Scholar, Web of Science and Scopus. According to a recent study carried out by the London School of Economics and Political Science (Martín-Martín, Alberto and others, 2018), the 3 sources combined provide comprehensive coverage. In particular, the authors recommended including Google Scholar for these reasons:

“The selective approach of Web of Science and Scopus produces a curated collection of documents, but is sensitive to biases in the selection criteria. Indeed, evidence has shown that these databases have limited coverage in the areas of social sciences and humanities, literature written in languages other than English, and scholarly documents other than journal articles. For its part, Google Scholar’s inclusive and unsupervised approach maximises coverage, giving each article “the chance to rise on its own merit”. – Martín-Martín, Alberto and others, (2018)

The search process followed 4 steps.

1. An initial search was carried out to identify the volume of ‘hits’ for each search string.
2. The search queries were submitted to the search engine or platform and the returned results were downloaded into a CSV file using a suitable web text scraper.
3. The CSV file was then automatically screened to remove results that were not relevant to the REA.
4. The remaining queries in the CSV were then subjected to a manual screening to provide the evidence base.

These steps are described briefly below.

3.1.2 Initial searches

Each search string was passed through Google Scholar, Web of Science and Scopus. These searches were carried out to ensure that the search strings provided a suitable body of evidence on which further screening could be based. The results of the searches are shown in Table 1. This details string ID, search term and the number of hits in Google Scholar, Web of Science and Scopus.

Table 1: Search hits for each search string[1]

String ID	Search term[2]	Google Scholar	Web of Science	Scopus
1	(‘flood risk’ OR ‘flood hazard’)	1,550	27	40
2	‘present and future flood’	8	0	0
3	(‘coastal flood’ OR ‘wave height’ OR surge OR tide OR tidal OR ‘sea level’)	68,100	200	278
4	(‘pluvial flood’ OR rainfall OR precipitation)	28,700	759	804
5	(‘fluvial flood’ OR ‘river flow’)	1,660	9	14
6	(‘inundation’ OR ‘flood depth’ OR ‘flood velocity’)	3,520	23	25
7	(‘defence breach’ OR ‘defence failure’)	24	0	0
8	‘hydrolog*’	333	4	9

3.1.3 Compiling the evidence base

Following the initial searches, the search strings were then run through a web text scraper and the outputs extracted to CSV files. This allowed a literature database to be compiled from which primary and secondary screening could be performed. When performing the searches, only the first (up to) 1,000 returned results were incorporated into the literature database. To prevent the database from being dominated by a small set of statistical models or applications, the searches were done iteratively through each (statistical model AND application) combination. The output literature database was a combination of all returned results.

[1] The syntax for the search strings needs to be verified before using

[2] Each search also includes ((emulator OR emulation) OR (‘meta-model’ OR ‘meta model’ OR ‘metamodel’) OR ‘surrogate model’ OR ‘response surface methodology’ OR ‘design selection’)

3.1.4 Primary screening and extraction

The primary screening was an automatic process that used the procedure outlined below.

1. Papers from journals that were deemed inappropriate were excluded; these journals were filtered based on their titles and included words such as ‘molecular’, ‘psychology’ and ‘food’.
2. Each paper was then evaluated according to the total number of search string keywords that appeared in the title and abstract and the number of citations per year. For each set of keywords, a paper would score either a 1 if any of the keywords appeared or a 0 if none of the keywords appeared. A sum of the keywords was then calculated. The number of citations per year was calculated by dividing the number of citations for each paper by the number of years since the published date.
3. The results were then sorted based on the sum of keywords in the title, citations per year and then the sum of keywords in the abstract and the top 100 were selected.
4. The contributions from the steering group were added to the screened list of papers.
5. Papers were given an ID based on the origin of the search (GS for Google Scholar, WoS for Web of Science, SG for steering group contributions) and the search string they are relevant to in the following format: <origin of search>_<search ID>_<paper ID>.

3.1.5 Secondary screening and extraction

Entries in the literature database then went through a secondary manual screening process led by the project scientists where each paper was assessed according to its relevance to the primary and secondary questions.

The majority of the process involved removing obviously irrelevant references. There were numerous examples of papers that had many keywords that were relevant, but the subject of the paper was not relevant to the review.

Eliminating these types of papers formed the majority of the secondary screening process. The remaining process involved reading the abstract and assessing, using expert judgement, the paper’s direct relevance to answer the evidence review questions, that comprised:

- the discussion regarding the balance between reduced simulation time and increased error or uncertainty associated with approximation
- the application of emulators (including the design selection) in the context of present and future flood risk, or the modelling of the flood model components, including sources, pathways and consequences
- the application of emulators for static versus dynamic (time-stepping) models in flood risk analysis and forecasting

The final literature database is included in Appendix B

4 Evidence summary (primary question)

From the screened evidence list, the papers were then assessed against the primary and secondary questions.

What is the evidence for the successful application of emulators in the context of analysis of present and future flood risk?

The review of the evidence shows extensive use of emulators within flood risk modelling and analysis. The evidence shows that these techniques have been applied to the different sources of flooding. This is show in Figure 5.

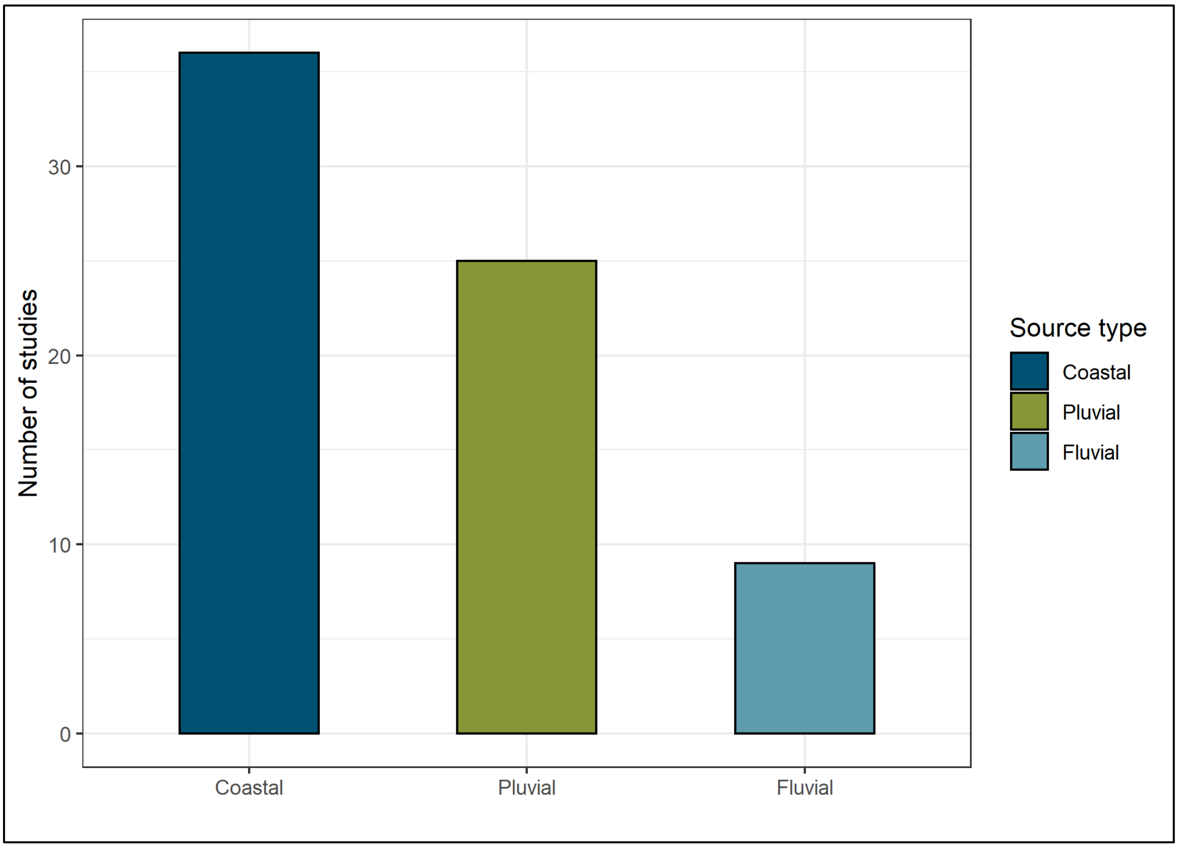


Figure 5: Evidence breakdown by flood source

Figure 5 shows the breakdown of evidence studies by coastal, pluvial and fluvial flooding sources. There are over 30 coastal studies, 25 pluvial studies and fewer than 10 fluvial studies.

The range of different components within the source-pathway-receptor conceptual model are show in Figure 6.

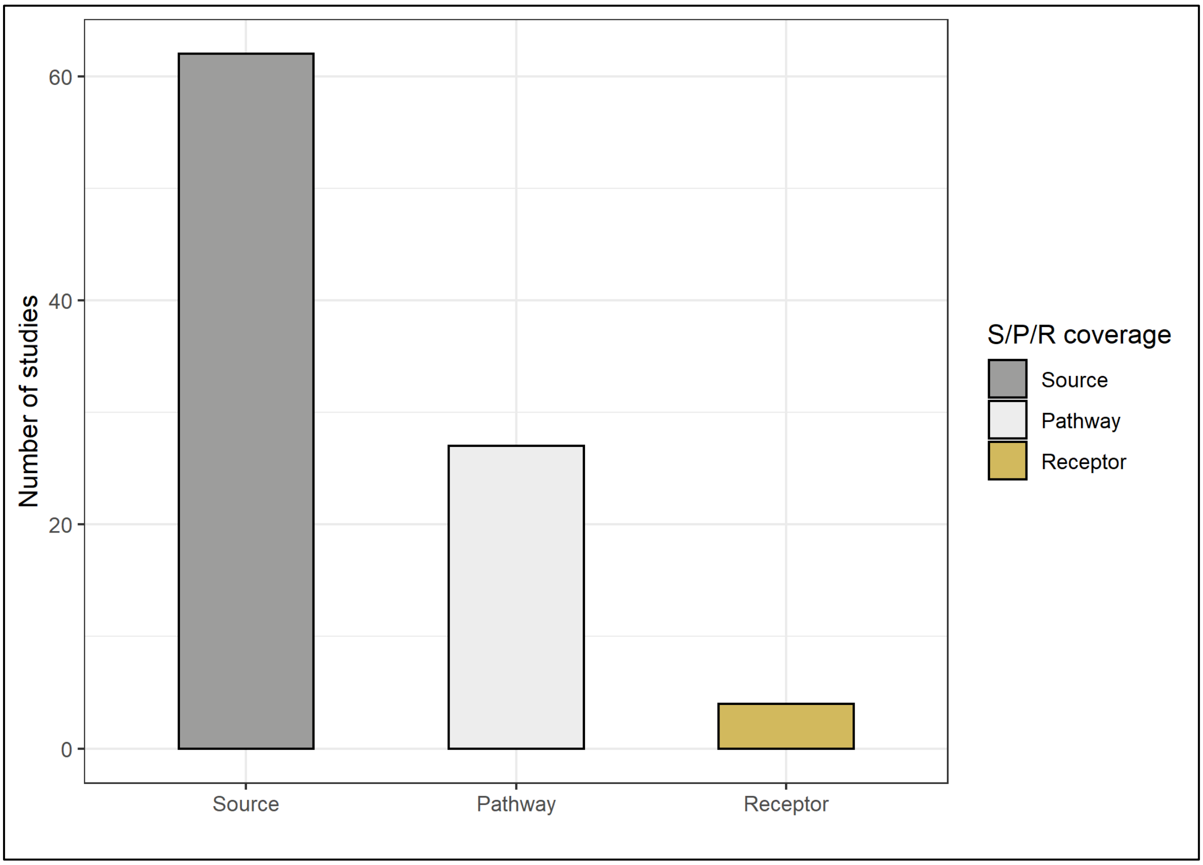


Figure 6: Evidence breakdown by source, pathway and receptor

Figure 6 shows that the majority (80%) of evidence relates to the 'source' component, and of these, approximately 50% are coastal.

Figure 7 and Figure 8 show the range of emulator techniques that have been applied across the different 'sources' of flooding and different components within the source-pathway-receptor (SPR) conceptual model.

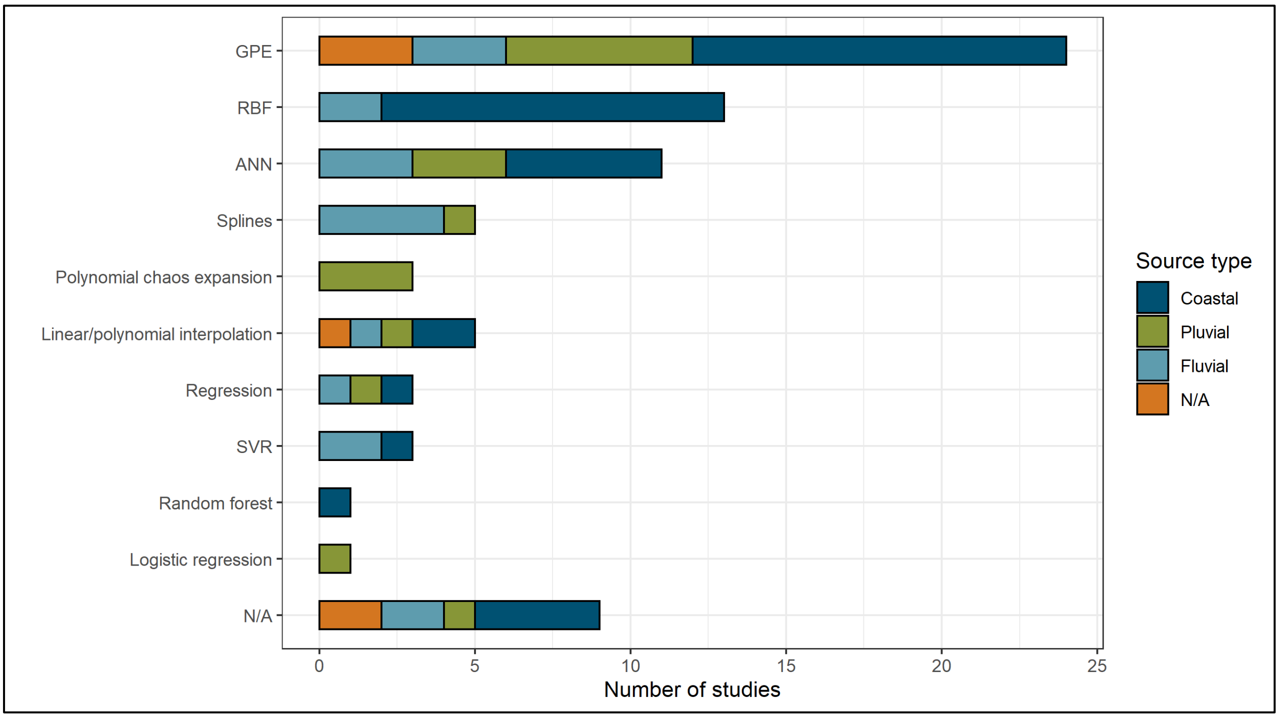


Figure 7: Breakdown of emulator techniques and flood sources

Figure 7 shows that the GPE emulator technique is by far the most used, particularly in coastal studies. There are very few examples of the use of SVR, random forest and logistic regression emulators.

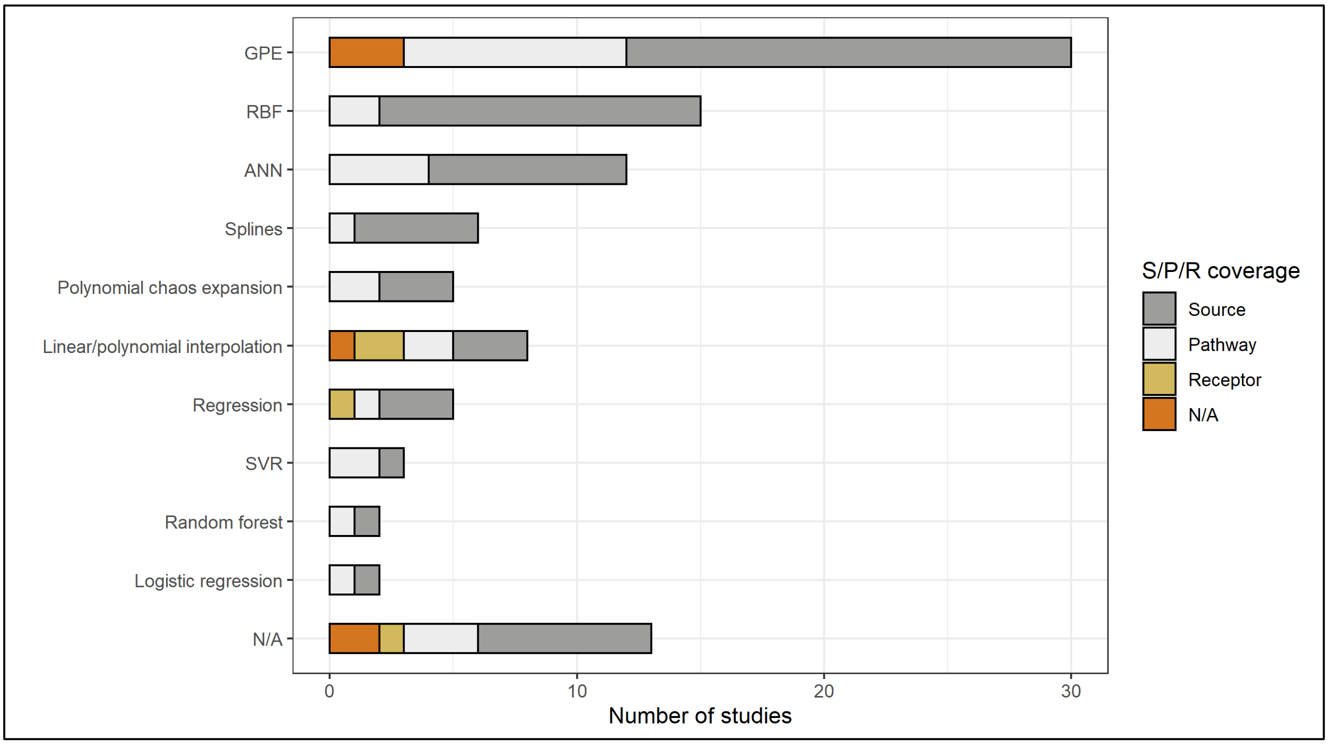


Figure 8: Breakdown of emulator types and source, pathway and receptor components

Figure 8 shows that the majority of the studies using the GPE emulator are for the source component, as opposed to pathway or receptor. Only linear interpolation and regression have been used for receptor studies.

Figure 9 and Figure 10 show a breakdown by climate change.

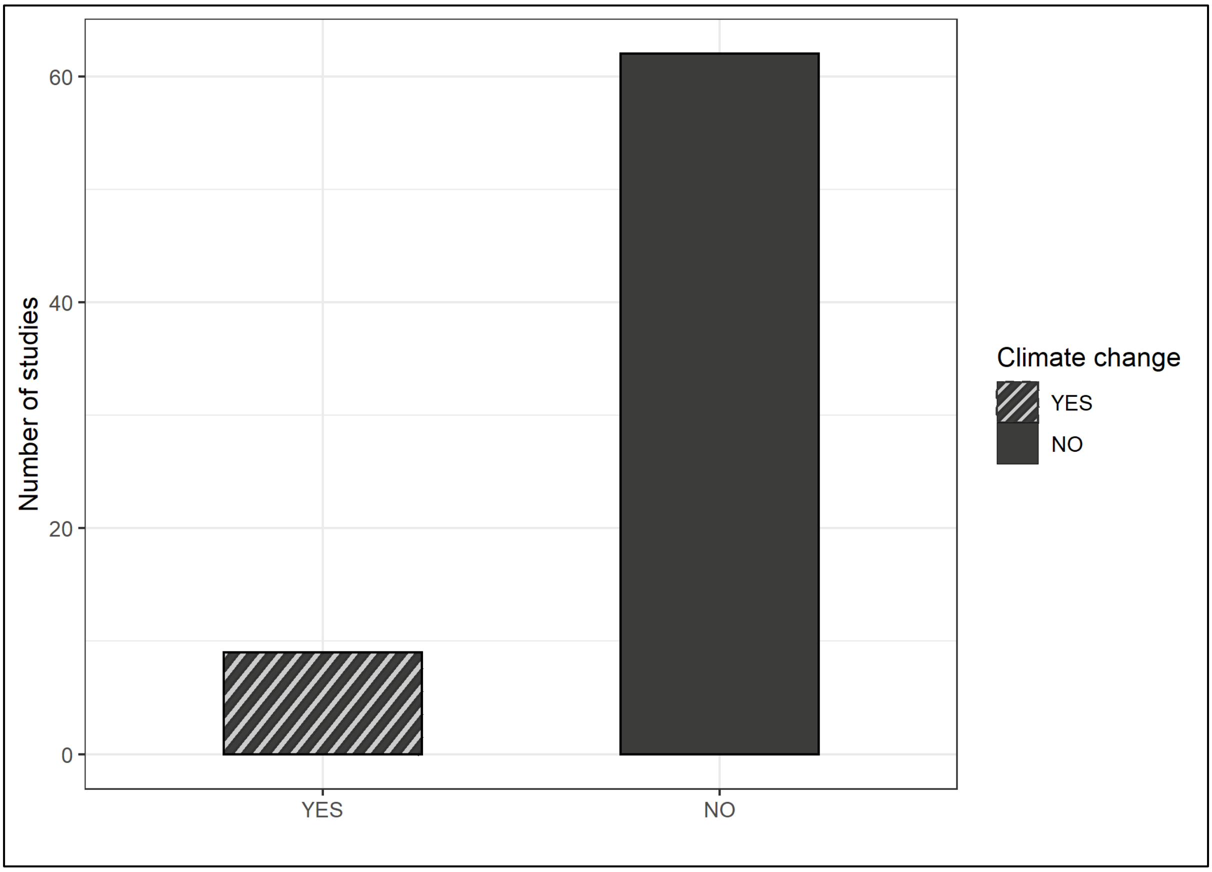


Figure 9: Evidence breakdown by climate change

Figure 9 illustrates that very few studies (about one sixth) included climate change in the use of emulators.

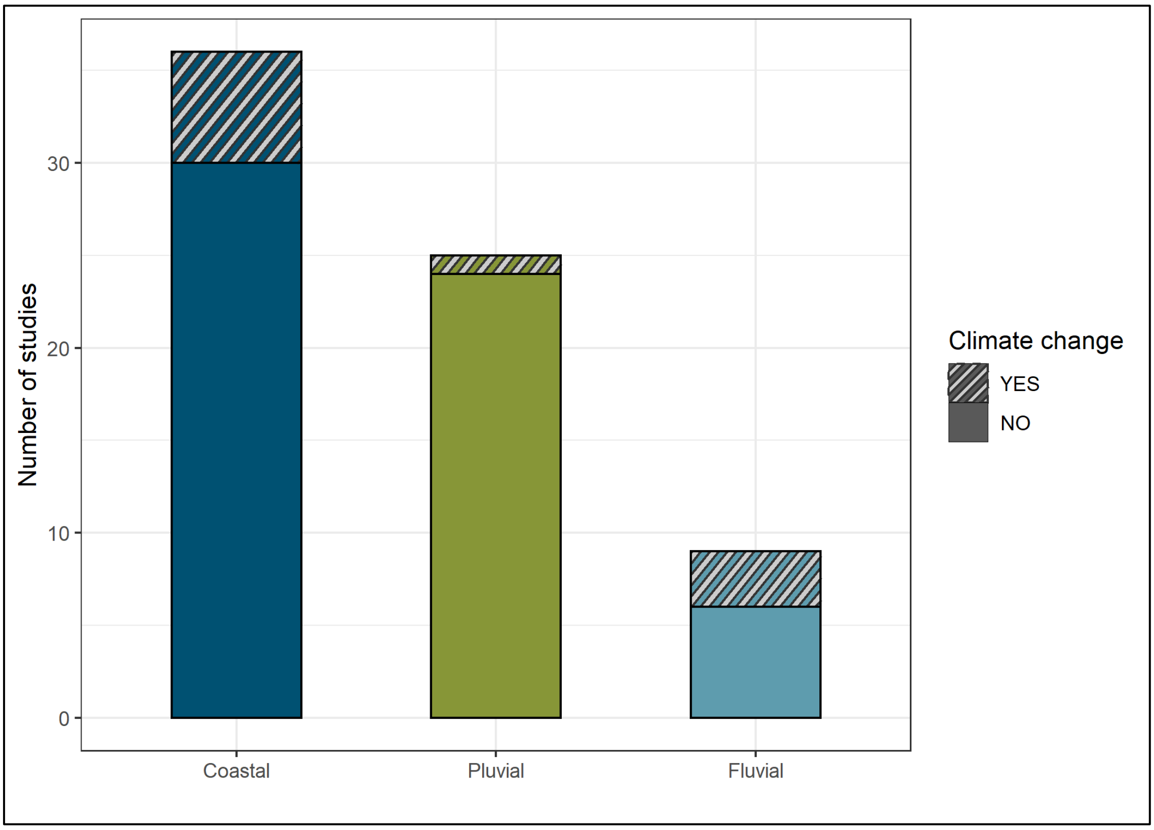


Figure 10: Evidence breakdown by flood source and climate change

Figure 10 indicates that where climate change was considered, this was mostly from coastal studies, followed by fluvial then pluvial. An interpretation and detailed discussion of these results is provided in response to the secondary questions in section 6.

5 Evidence summary (secondary questions)

5.1 Emulator uncertainty

What evidence is there of successfully finding an appropriate balance between reduced simulation time and increased error or uncertainty associated with approximation, and what criteria have been used to determine this balance?

There is some evidence (approximately 20% of papers) to show the application of methods that consider the trade-off between the number of design points and the accuracy of the emulator. These are shown in terms of a breakdown by flood 'source' in Figure 11.

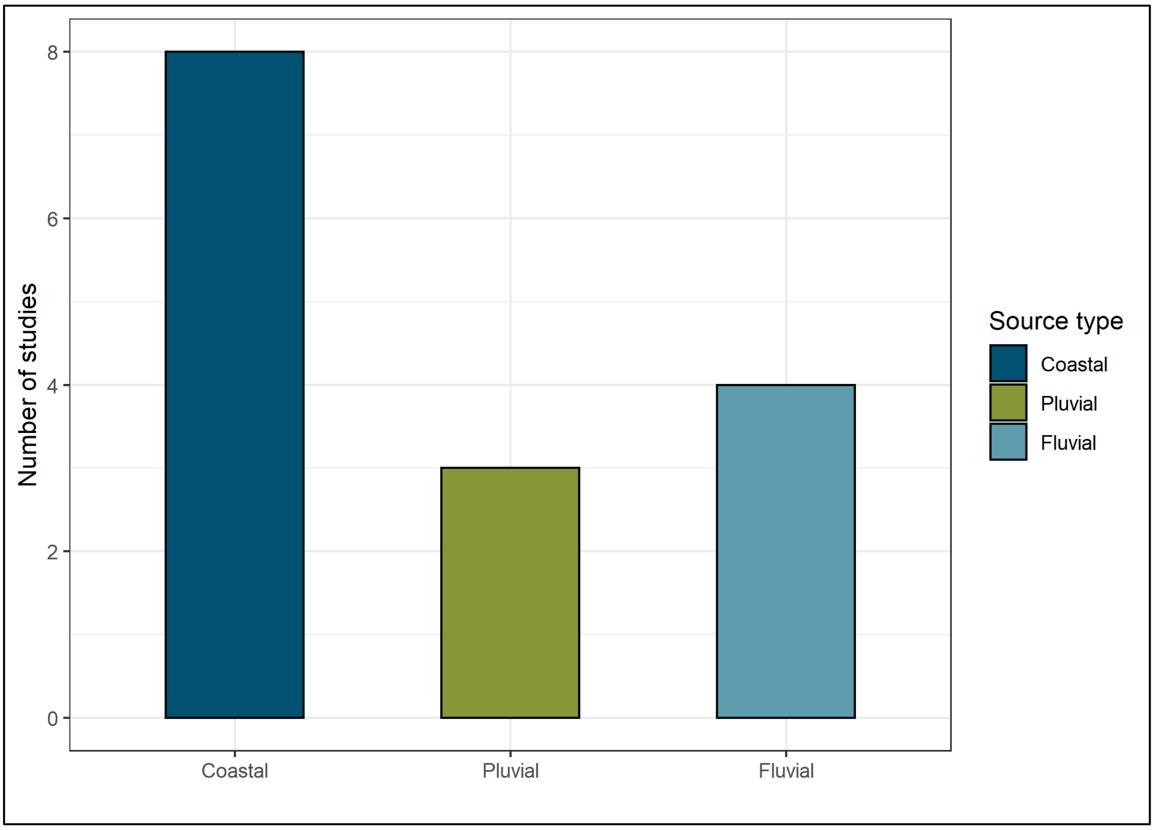


Figure 11: Breakdown by flood "source" of papers that seek to balance runtime and accuracy

Figure 11 shows that when focussing on studies balancing runtime and accuracy by flood source, most studies consider coastal, followed by fluvial and then pluvial flood sources.

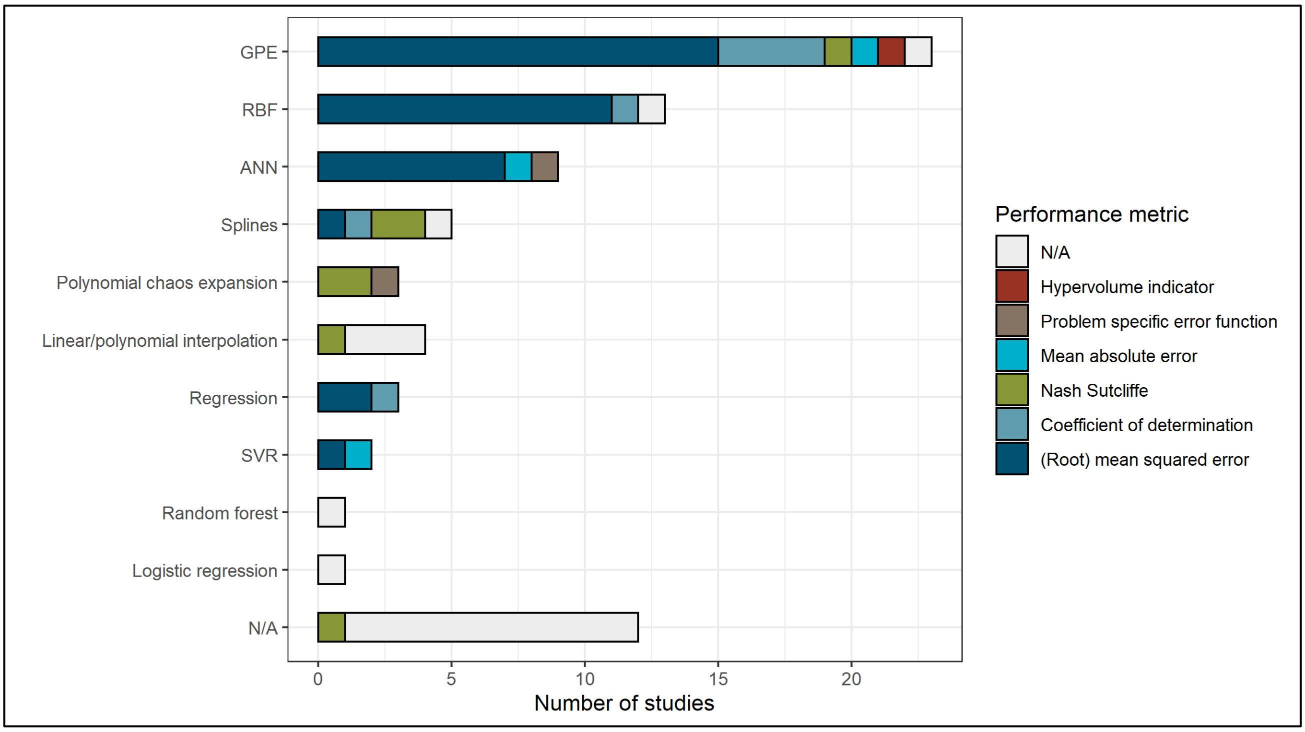


Figure 12: Breakdown of emulator techniques and performance metrics used in their evaluation

Figure 12 shows that a range of performance metrics have been used to assess the uncertainty associated with emulators, with root mean square error (RMSE) being the most popular, followed by mean absolute error, coefficient of determination and Nash Sutcliffe. Further discussion on methods for determining error and placing these errors in the context of the overall analysis is provided below.

5.1.1 Validation of emulators

Cross validation, such as leave-one-out or K-fold cross validation, is a technique that can be effectively applied to evaluate the predictive error associated with emulators. In the context of emulators, the technique involves a series of steps.

1. Run the computational model for a set of design points.
2. Remove a random subset from the original set of design points.
3. Fit the emulator to the remaining design points.
4. Use the newly fitted emulator to predict the value at the design points that were removed in step 1.
5. Compare the predicted value at the removed design points with the actual value to calculate the prediction error.
6. Repeat steps 1 to 3, by selecting different subsets of design points to remove.
7. Evaluate the overall predictive error for all removed design points.

With the leave-one-out cross validation, only one design point is removed each time, whereas 1/K of the design points are removed with the K-fold cross validation. The advantages and disadvantages of these approaches are discussed in Hastie and others (2013), together with recommendation on the choice of K. A specific example related to the methodology applied on the Environment Agency's national flood risk assessment is shown in Figure 13 (Malde and others, 2016).

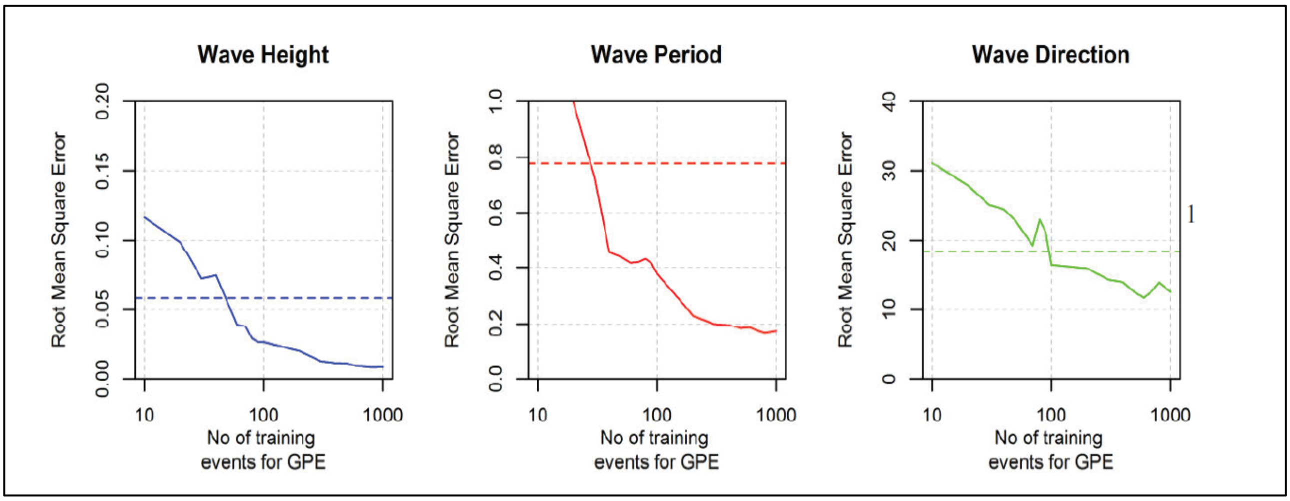


Figure 13: Reduction in emulator prediction error in relation to number of model simulations (Malde and others, 2016)

Figure 13 shows 3 graphs plotting root mean square error against the number of training events for GPE. In this example, the cross validation technique was applied to sets of design points that have been increased. The figure shows the reduction of the RMSE in emulator predictions of wave height, period and direction, with an increase in the number of design points. Because a space filling algorithm has been used, the RMSE calculated can be considered a 'worst case'. This is because the design point that is left out, by definition of the algorithm, is far away from the other design points and therefore likely to have most error associated with it.

This traditional approach involves using a regular grid (not space-filling algorithm) and linear interpolation (as opposed to a GPE). The dashed lines were generated using 1,200 design points. Therefore, what the graphs show is that the MDA or GPE approach reduces the required number of design points (computational model simulations) to achieve the same RMSE. In terms of wave height, for example, this reduction is by more than an order of magnitude (approximately 70 versus 1,200), based on where the dashed and solid lines cross each other).

While Figure 13 provides an indication of the error of the emulator for this particular component, it does not say what is an acceptable error in terms of a practical application. Answering this type of question can be complex. In specific relation to coastal modelling, the Environment Agency has published standards (Environment Agency, 2016). These standards specify an acceptable error for a specific model component in relation to a specific activity. For example, for strategic analysis, 0.5m RMSE is prescribed for nearshore wave conditions. This error is based on a judgement about the influence of an error in a model component when related to the uncertainty in an overarching calculation.

As emulators are computationally efficient, they facilitate uncertainty and sensitivity analysis. In a recent study for the Environment Agency, the wave overtopping model, BAYONET GPE (Pullen and others, 2018) has been implemented within a chain of coastal models to undertake uncertainty and sensitivity analysis (Environment Agency, 2020). The chain also includes an emulator of the SWAN wave transformation model. The results show that some model components have little influence on the output result, in terms of their contribution to uncertainty, when compared to other components. An example of this is shown in Figure 14.

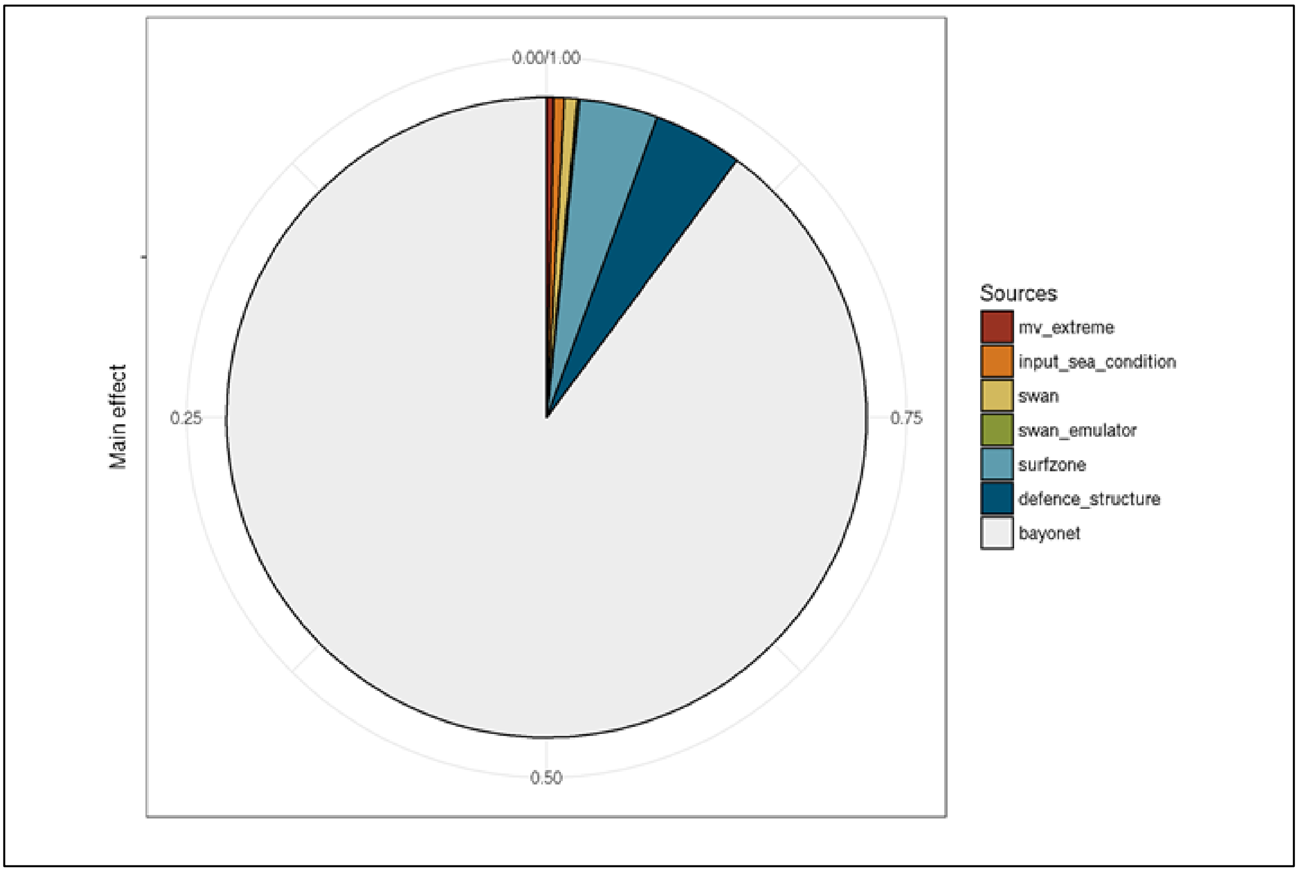


Figure 14: Example of emulators being applied in the context of uncertainty and sensitivity analysis. In this example, BAYONET GPE contributes most to the overall uncertainty. Environment Agency, 2020

The pie chart in Figure 14 represents the overall uncertainty. Each section of the chart shows the contribution different sources of uncertainty make to the overall uncertainty. The figure shows the SWAN model and related SWAN emulator contribute a small fraction to the overall uncertainty in wave overtopping rate. This is indicated by the small ‘pie slices’ of green and yellow colours, respectively. This highlights the importance of considering model errors in relation to the context of the application and uncertainties that are present elsewhere within modelling chains. In other words, errors associated with emulators should be placed into context with other uncertainties.

In this particular study (Figure 14), the uncertainty associated with the predictions of the wave overtopping model, BAYONET GPE, contributed most to the overall uncertainty. This is primarily due to the limited range of physical model tests on which the emulator is based. To reduce the uncertainty, a range of further model experiments that are focused on the areas where data is currently lacking is required.

5.1.2 Combining emulators with evolutionary optimisation algorithms

To increase computational efficiency, without loss of accuracy, there is evidence of emulators being combined with evolutionary optimisation algorithms (artificial intelligence), in the context of flood risk analysis. Two examples of this are described in relation to:

- reliability analysis of a New Orleans levee ('17th Street Canal') that failed during Hurricane Katrina (Kingston and others, 2011)
- calibration of a MIKE 11 rainfall run-off model (Khu and others, 2004)
- identification of an optimised flood risk mitigation strategy, taking account of climate change uncertainties (Woodward and others, 2013)

The development of an emulator, of the geotechnical model, used to assess the reliability of the New Orleans levee provides an example of a specific type of problem that can arise. In this example, an emulator of a geotechnical model was required. To solve the specific problem, however, it was not necessary for the emulator to perform equally well across the full range of input parameter space. This is because probabilistic reliability analysis requires a binary 'fail or no fail' output of a limit state equation (LSE) (for example, see Melchers and others, 1999). The probability of failure output from the reliability analysis is particularly sensitive to evaluations in the region of the limit state (that is, the surface that distinguishes failure from non-failure), and training of the emulator is particularly important in this region. Therefore, to define the emulator, it was important to have a high density of design points along and around the limit state.

Unfortunately, because of the complexity of the problem and the number of input parameters to the geotechnical model, the location of the limit state was not known at the outset. To solve this problem, an evolutionary optimisation algorithm, in this case a genetic algorithm (GA), was applied to help select the design points and ensure that these were focused along the limit state. There are 6 steps in the process.

1. Specifying an objective function of the optimisation algorithm in terms of the LSE output (Z). More specifically, Z=0 defines the limit state and values of Z close to 0 are close to the limit state.
2. Randomly selecting a limited set of design points to cover the input parameter space.
3. Running the geotechnical model to evaluate the LSE for the initial population.
4. Evaluating the objective function and applying the GA to identify a new set of design points that are likely to be closer to the limit state.
5. Running the geotechnical model for the new set of design points.
6. Repeating steps 4 and 5.

Figure 15 shows the steps in the process.

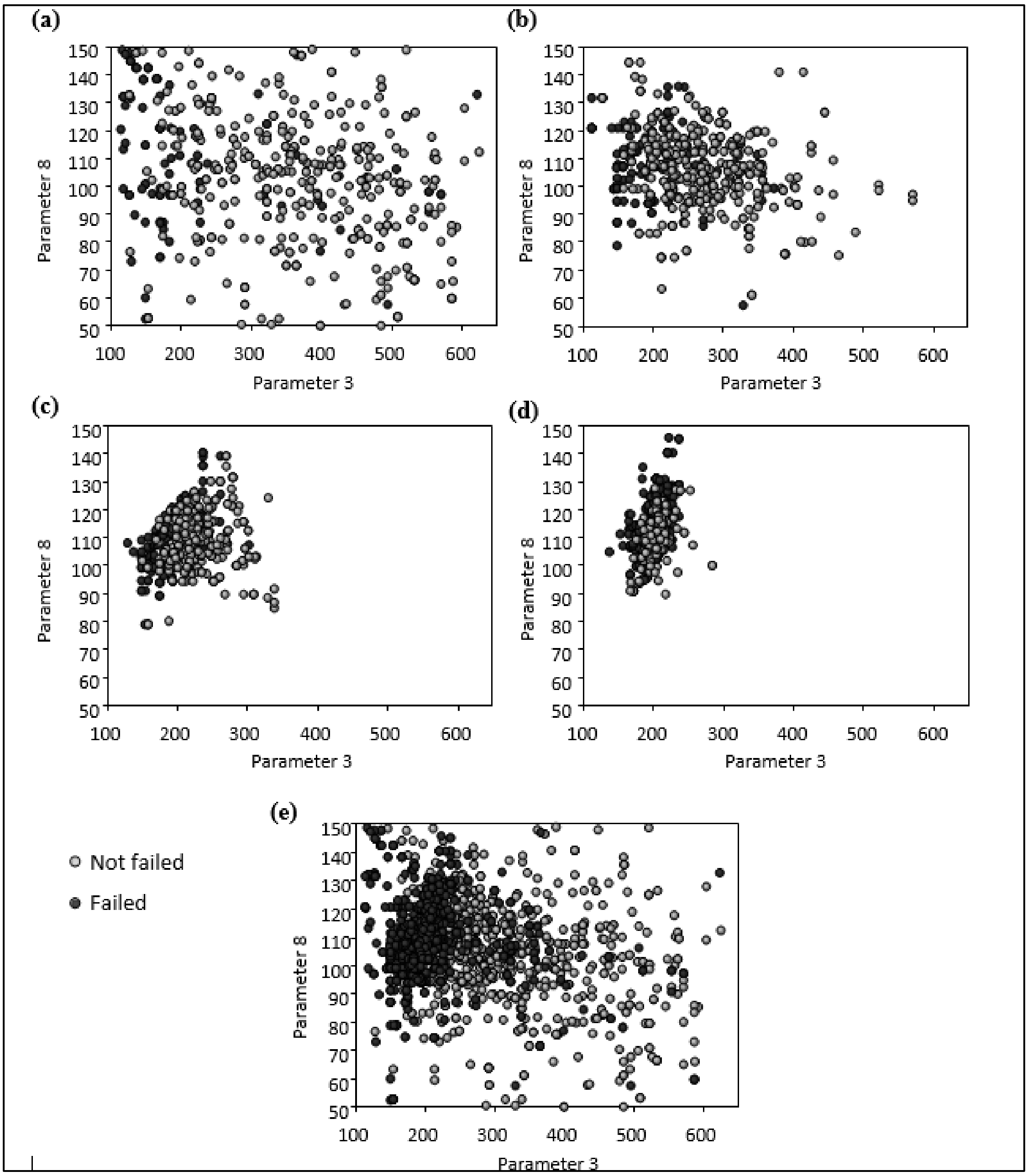


Figure 15: Iterative selection of design points for an emulator of a geotechnical flood defence failure model, New Orleans (source: Kingston and others, 2011)

The 5 graphs in Figure 15 show improvement of an emulator used in a flood defence failure study in New Orleans. Panel A (top left) shows the initial random selection of design points (illustrated using 2 of the more dominant input parameters) covering the input parameter space. Panels B (top right), C (middle left), D (middle right) and E (bottom middle) show successive iterations of design points identified by the GA. The tight clustering in panel D illustrates how the GA has successfully identified the limit state. All points illustrated in panel D give Z values that are close to 0. Panel E combines all other panels and shows the overall set of design points that were used to train the emulator. The characteristics of this set comprise a high density of design points along the limit state, where greatest accuracy is required.

Another example of the coupling of an emulator with an optimisation algorithm is described by Khu and others (2004). This example relates to the calibration of a rainfall run-off model (MIKE 11) using an ANN and a GA. The purpose of the study was to achieve an optimum, and computationally efficient, calibration using the emulator in place of the rainfall run-off model. Several years of validation data was available and there were a number of different parameters that could be adjusted to obtain a good fit to the validation data.

The steps in the process are similar in many regards to the New Orleans levee example. In this case, however, the objective function of the optimisation algorithm was defined based on the RMSE associated with the MIKE 11 model when compared to the validation data. The initial population of design points was selected at random and used to train the emulator. The GA was then used to identify model parameter sets that were likely to give a lower RMSE. The emulator was used instead of the MIKE model to evaluate the new set of design points, thereby making substantial computational savings. This process was repeated to define an optimum calibration parameter set. This set was then evaluated using the MIKE model to ensure the emulator was performing as required.

5.2 Emulators and the source-pathway-receptor model

What is the evidence for successfully using emulators to predict flood sources (for example, waves, river flows), pathways (wave overtopping, breaching, flood depth, flood velocity), consequences (economic damage, loss of life) and overall flood risk (based on whole system models)?

Emulators were applied to all components of the source-pathway-receptor conceptual model. The breakdown is shown in Figure 16.

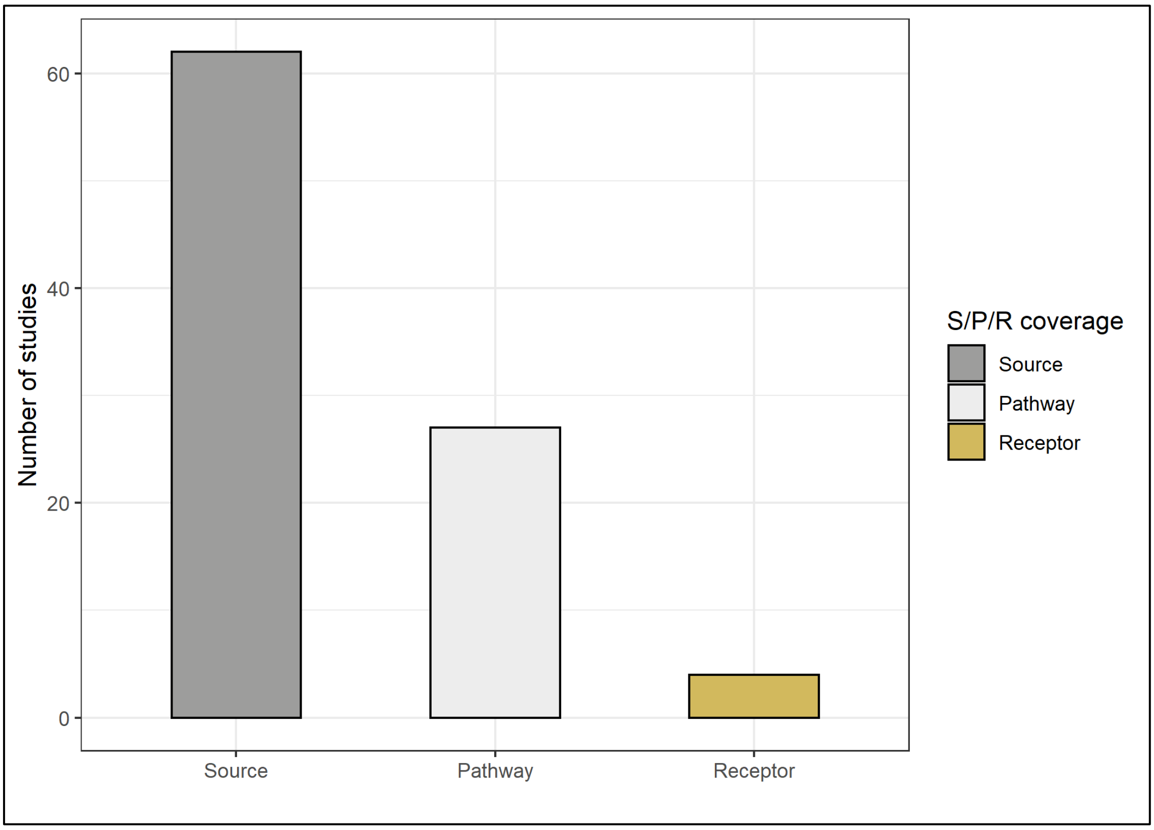


Figure 16: Evidence breakdown by source, pathway, receptor

In Figure 16, most of the papers (approximately 75%) relate to the ‘source’ component, with the remainder primarily focused on the ‘pathway’ component.

A breakdown of the different pathway components is shown in Figure 17.

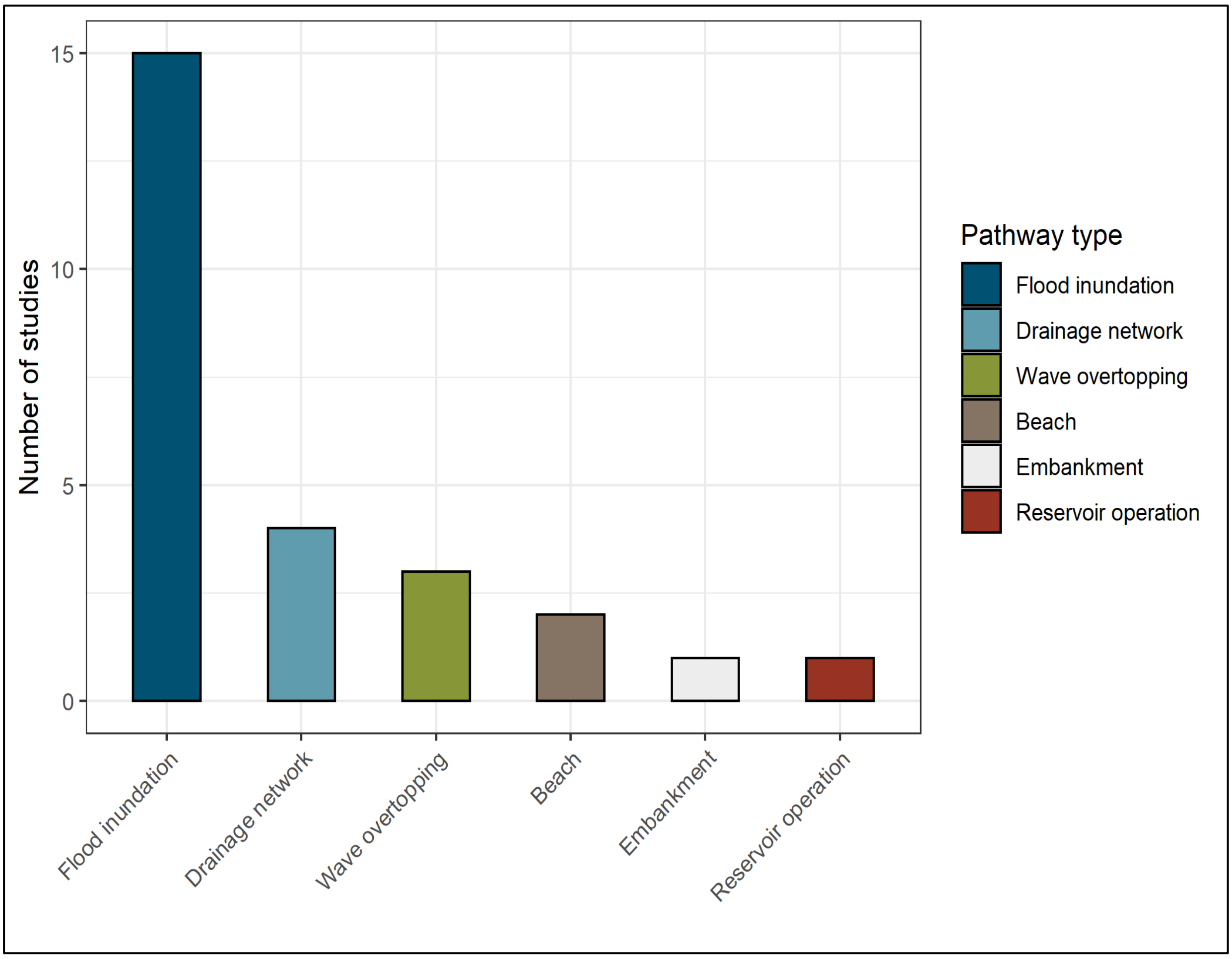


Figure 17: Emulator evidence defined by pathway model type

Figure 17 indicates that the greatest number of emulator evidence came from flood inundation, followed by drainage networks, wave overtopping, beaches, embankments and finally reservoir operations.

The distribution of the application of emulators over the components of the SPRC conceptual model is likely to arise due to a number of reasons. In particular, these can relate to the:

- complexity of the models that are applied
- maturity of the application of emulators versus reduced complexity models in the field
- maturity of drivers requiring the use of emulators (for example, probabilistic and uncertainty analysis)

Consequence models (loss of life or economic damage) tend to be fairly simplistic in practice. As a result, there is often no requirement or motivation to use emulators. For example, the widely-applied Middlesex multicoloured damages (Penning-Rowsell and others, 2013) are simple functions that give economic damage in relation to flood depth. The calculations can be carried out in one or two lines of computer code and are exceptionally fast to calculate. In general terms, even for probabilistic analysis, it is not necessary to further simplify this model to reduce computational runtime and therefore there is no motivation to use emulators for this situation.

Similarly, loss of life models can be simplistic in nature (for example, Defra and Environment Agency, 2005) and there is little in the way of computational demand. For this reason, there has not been extensive development of emulators in the context of receptor model components.

With regard to hydrological modelling, there is a long tradition of using reduced complexity, 'lumped models' that average spatial characteristics. As these are relatively simplistic to apply, it could be that this has led, to a certain extent, to less focus in terms of the development of emulators in this field. More recently, however, gridded models are becoming more widely applied in practice, together with a recognition of uncertainties and probabilistic analysis, and this has led to increased interest in emulators.

In contrast, even the earliest wave transformation models capture the 2D spatial characteristics of the bathymetry and the processes of refraction and shoaling, that are well understood from linear wave theory. As the inputs to even the earliest models comprised multiple variables (wave height, period, direction and wind speed) and there was often a requirement to consider time series data, emulators (in the form of simple lookup tables comprised input parameters spaced on a regular grid) have been widely applied since the late 1980s. There is evidence of this in the grey literature (for example, historical HR Wallingford consultancy reports), but not in the widely available literature.

Moreover, techniques for the joint probability analysis of waves and sea levels have employed Monte Carlo approaches since the mid-1990s (HR Wallingford, Lancaster University, 1998, and Hawkes and others, 2002). These statistical (Monte Carlo) simulation approaches require many different input sea states to be evaluated. This demand has also stimulated the field to explore emulation approaches to limit the associated computational burden.

It is possible that this early adoption of emulator type approaches gives rise to the increase in evidence obtained from coastal models from this review.

The evidence shows that a wide range of emulator techniques have been applied, the most popular being GPE, which has been applied in approximately 30% of the papers (see Figure 7 and Figure 8). There is no clear evidence to suggest the choice of emulator technique is dependent on the flood sources or the model components, that is, the proportion of studies on different flood sources or components is similar across all emulator techniques.

The choice of a particular emulation method used is complex and depends on:

- the time the paper was developed
- the background field of the researchers
- the availability of software

Pre-widespread use of the internet in the late 1980s and mid-1990s, the emulation method chosen depended largely on the background of the researchers. In contrast to today, there were not freely available toolboxes in high level programming languages that facilitated the efficient adoption and comparison of a wide range of different methods. For example, the response surface method and related use of piecewise polynomials was highly prominent in the engineering community relating to the influential paper of Myers (1971) and the subsequent highly cited paper of Faravelli (1989). This approach formed the basis of an early application of the response surface method type approach in relation to flood risk analysis.

HR Wallingford (2002) describes the application of piecewise polynomials to a fluvial or tidal joint probability problem on the tide estuary. Joint extremes of fluvial flows and sea levels were evaluated and then a Monte Carlo simulation of extreme (joint) events was undertaken using an established approach (HR Wallingford, Lancaster University, 1998).

In that study, an existing 1D hydraulic model of the estuary was used to calculate extreme water levels at different cross sections in the river. Given the computational speeds available at the time, it was computationally intensive to run the hydraulic model for each Monte Carlo event. To overcome this problem, piecewise polynomials were fitted to a small sample of design points created using the 1D hydraulic model. This fitted surface was then used in the place of the hydraulic model to evaluate the remaining events at each cross section. The method was implemented in Fortran software specifically developed at the time for the project.

In time, as methods have evolved, there is now the opportunity to undertake a more objective analysis in terms of identifying the most applicable method. The methods are themselves more readily accessible in terms of open source software libraries. Examples of this include:

- Manage Uncertainty in Complex Models (MUCM)
- DiceKriging (Roustant and others, 2012)
- BACCO (Hankin, 2005)

An example of this evolution and decision-making process is provided in the next section through a case study using wave overtopping models.

5.2.1 Wave overtopping case study

Emulators, predicting the process of wave overtopping of coastal structures, have been widely applied in practice for well over a decade. This section provides an historical insight into:

- motivation for using emulators
- motivation of the choice of emulator
- practical applicability and limitation issues that arose
- how these application issues have been addressed

Wave overtopping of coastal structures is a physically complex process. The predictions involve the highly dynamic process of extreme waves impacting and interacting with coastal structures. The input parameters to the calculations include sea conditions (height, period, direction, sea level) and structure geometry (for example, crest level, slope, freeboard, berm width).

The area has been extensively studied using small-scale physical models and computationally intensive numerical models, for example, by Dodd (1998), Hubbard and Dodd (2002), Hu and others (2000), Liu and others (2000), Garcia and others (2004), Mattis and others (2018), Cozzuto and others (2019) and Dimakopoulos and others (2019). Both of these models are resource intensive. As there is an extensive requirement for wave overtopping predictions for designing coastal structures and flood risk analysis, empirical models (EurOtop, 2018) and emulators have been a natural choice for providing these predictions.

Van-Gent and others (2007) describe the first development of a widely used neural network wave overtopping model. An executable software version of the model was freely available for download and this supported its widespread use in practice. The model is based on a database of physical model experiment results that had been collated from laboratories across Europe. The experimental results were an artefact of the various research and consultancy studies that had been carried out at each laboratory. Therefore, the underlying data (design points of the emulator) were not chosen in advance, rather they were just the data available. This subsequently led to issues relating to limits of applicability, and this is discussed further below.

It is apparent, like many similar studies at the time, that the choice of the neural network technique was not based on a careful consideration of the merits of different types of techniques. In this example, the choice of a neural network is attributed to the successful application of this technique in the field. More specifically, it is attributed to the design of rock armour structures (Mase and others, 1995). The discussion within this paper does not mention that other alternative techniques were considered.

Following its introduction, Kingston and others (2008) continued to use a neural network but extended the approach of Van-Gent and others (2007) to capture uncertainties associated with the fitting process. These are important in terms of the overall representation of uncertainty within the model.

Experience on the Environment Agency’s State of the Nation, national flood risk project highlighted significant issues relating to the application of these emulator models. Consultants, following the guidance provided by Van-Gent and others (2007), derived coastal structure geometry data for the majority of coastal flood defences in the country. It became apparent that over 70% of the data provided to the study was substantially outside the range of physical model experimental data that was available within the database and on which the emulator was based. This situation prompted further development.

Pullen and others (2018), using the same base data set, opted to replace the neural network with a GPE. The reasons for this related to an existing framework for capturing uncertainties associated with distance from design points and also uncertainty within the design points themselves. The latter arises as the underlying database is derived from small-scale physical model tests carried out in different laboratories. Where the same tests were undertaken in different laboratories, this did not necessarily lead to the same results, and therefore there was uncertainty in the design points. This feature was captured through the use of the so-called 'nugget' (Andrianakis and Challenor, 2012).

Pullen and others (2018) also sought to address the issue relating to the limits of application of the emulator. The original guidance accompanying the model of Van-Gent and others (2007) specified the upper and lower limit of each input parameter as being within the predictive capability of the model. As there are so many input parameters, and correlation within the underlying data, this gave rise to a situation where the model was being routinely applied in practice, with users apparently unaware that there were no underlying data to support the model outputs. In general terms, this is inappropriate for emulators that are developed to estimate values between known data points. This situation is conceptually illustrated in Figure 18.

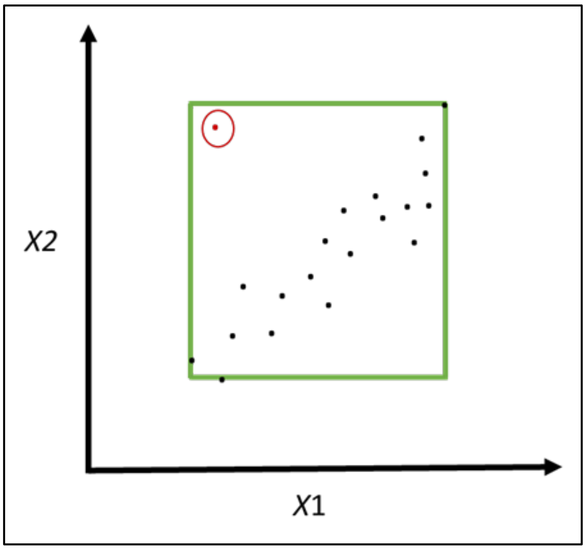


Figure 18: Conceptual illustration of the inappropriate use of the limits of input parameters being used to define the limits of applicability of an emulator (source: Pullen and others, 2018)

In Figure 18, a green square contains black dots indicating a positive correlation between the x-axis (parameter X1) and y-axis (parameter X2). Towards the top left corner of the green square there is a red dot. The red point is within the defined limits of applicability, based on the guidance of Van-Gent and others (2007), but there are no underlying data to support the predictions.

To overcome this situation, Pullen and others (2018) used a statistical measure, Mahalonobis distance, to give model users clear guidance where they were requesting predictions that were too far from the underlying data. This is illustrated in Figure 19.

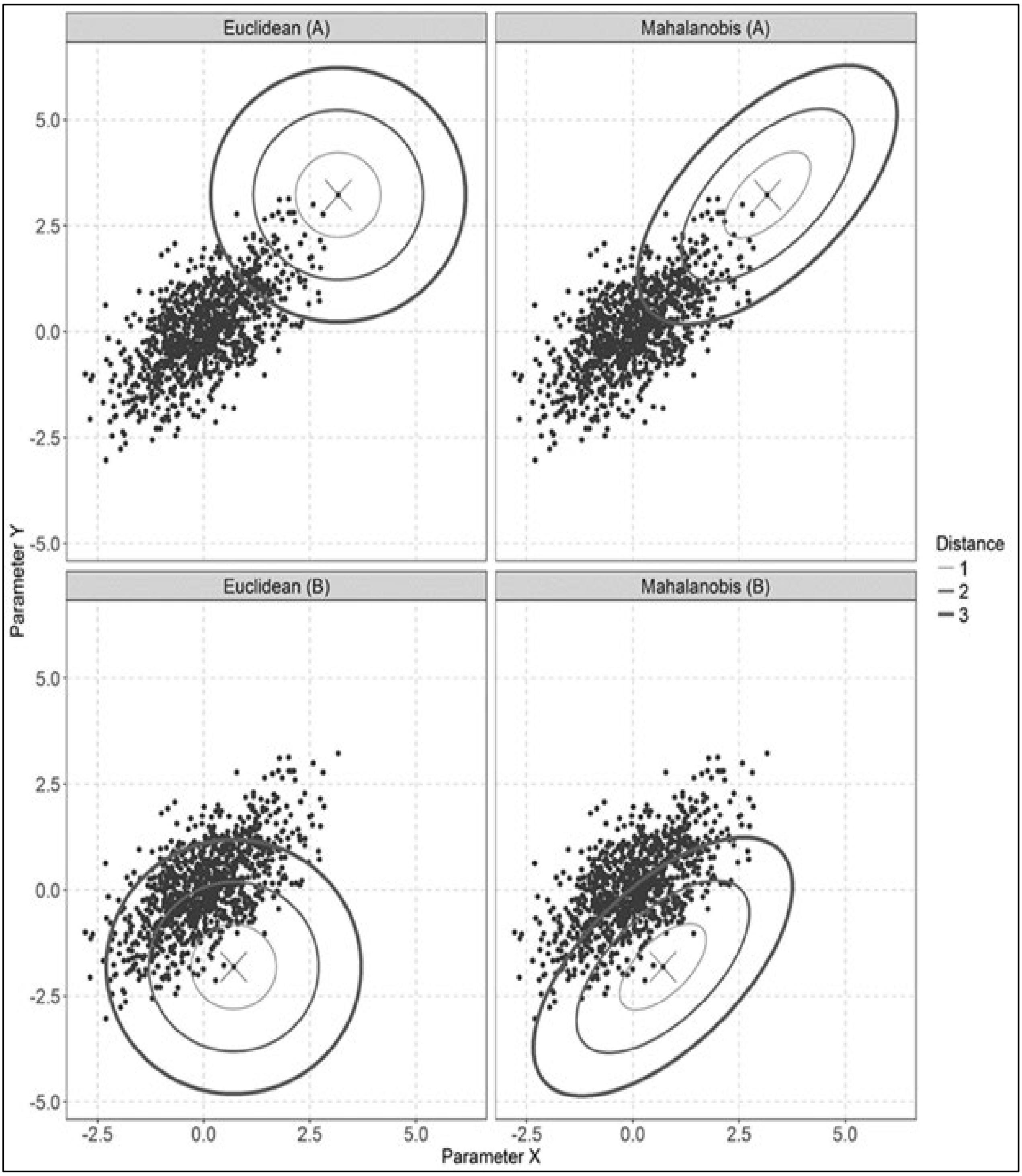


Figure 19: Graphical illustration of the Mahalonobis distance measure, used to assess the limits of applicability of a wave overtopping emulator (source: Pullen and others, 2018)

Figure 19 shows 4 panels. The 2 panels on the left show Euclidian distance from a parameter set of datapoints - these are 3 equal concentric circles around a central point. The other 2 panels on the right show the Mahalonobis distance measure around a point, which is more oval and elongated.

The use of this measure is captured within an online version of the BAYONET GPE model. This is illustrated in Figure 20.

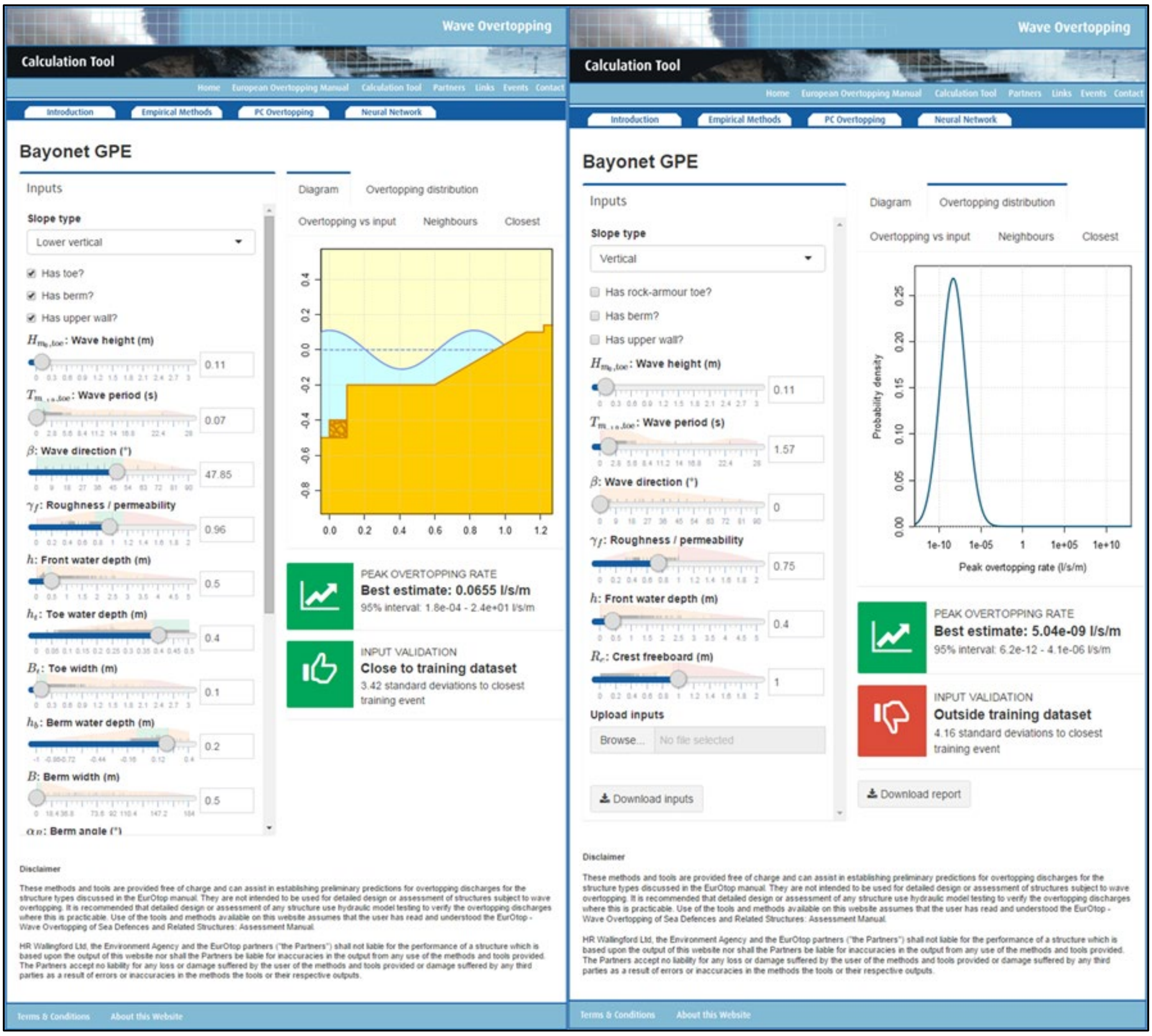


Figure 20: Screenshot of a translation of the Mahalanobis distance measure user guidance within an online wave overtopping model (source: https://www.overtopping.co.uk)

Figure 20 shows the model set-up as seen on the overtopping webpage. This provides evidence to highlight the importance of clear guidance on the application limits of emulators in practice.

5.3 Dynamic and static emulators

What is the evidence for using emulators for static versus dynamic (time-stepping) models in flood risk analysis and forecasting?

Dynamic emulators comprise a time dimension. These can be considered in 2 categories.

1. Emulators whose input parameters include variables specified at multiple preceding time steps and, potentially, multiple points in space. The emulators are then used to predict model outputs at future points in time (and potentially space). These

emulators can seek to replicate grid hydrological or hydraulic models that solve partial differential equations (for example, dynamic flood inundation models).

2. Static emulators that generate time series outputs but whose input parameter set is based on a single point in time.

The benefit of the former is that the ‘memory’ of the system that is being modelled is captured within the emulator. This is an important requirement for many hydrological, hydraulic and coastal models that are seeking to predict the evolution of a system in time. The challenge is that the models are often run on a grid and for long periods of time (days, weeks or even years), which means the number of input parameters (one for each times-step) becomes large. Training emulators with a large number of parameters can become inefficient. Situations can arise where more model simulations are required to train the emulator than to solve the problem using a more conventional method. For these reasons, these types of dynamic emulator models are not widely applied in practice unlike the static emulators.

There is evidence of both of these types of emulators that produce time series (dynamic) outputs, in particular in relation to coastal and pluvial sources. This is illustrated in Figure 21 and examples and further discussion is provided below.

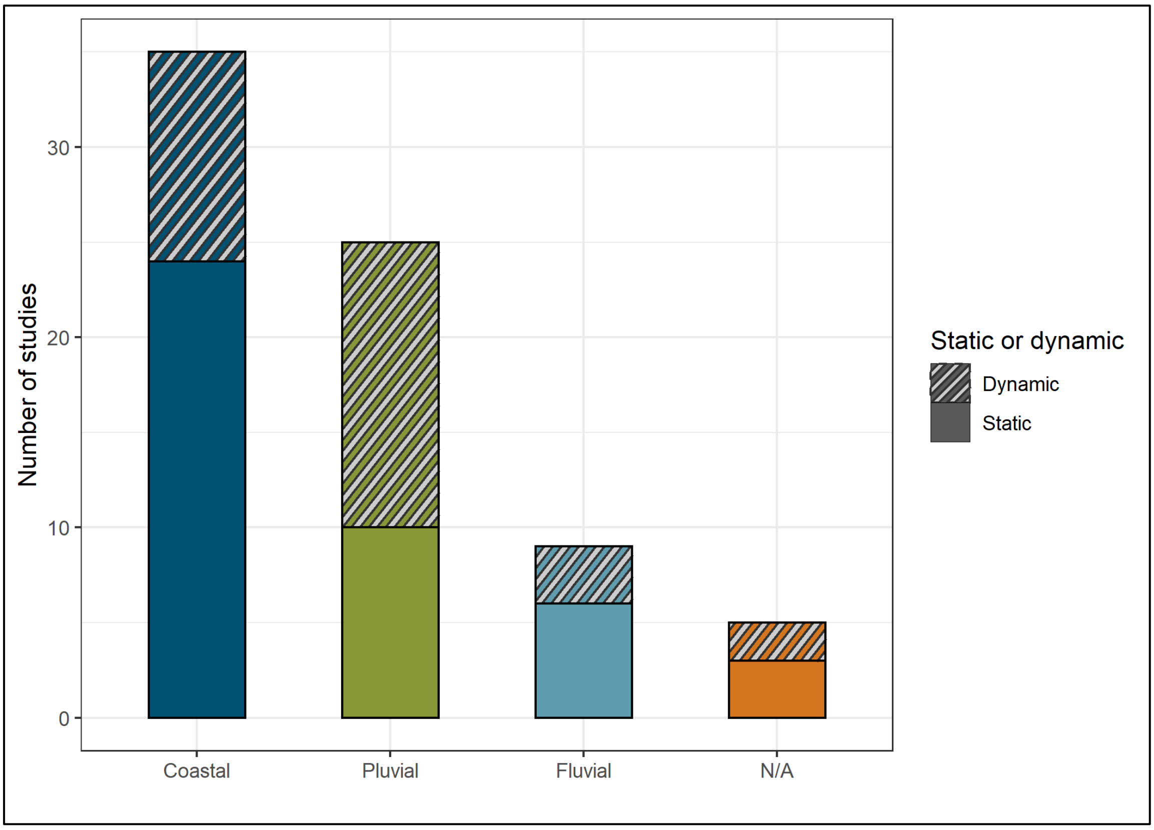


Figure 21: Evidence of dynamic emulator application by flood source

Figure 21 illustrates that the greatest number of studies where there is evidence of dynamic emulator application is in coastal, followed by pluvial and fluvial flooding. Coastal and fluvial flood sources have a greater proportion of studies related to static applications, whereas for pluvial it is dynamic.

Castelletti and others (2012) describe the development of a framework for the implementation of dynamic emulators in environmental modelling. In their paper, they consider 2 categories of emulators:

- structure-based
- data-based

Their description of a structure-based emulator is one that is often referred to as reduced physics, simplified model or reduced complexity model. This type of approach does not conform to the conceptual framework describing emulators defined for this project (section 3). This highlights the importance of having clear terminology when discussing emulators.

Castelletti and others' data-based description of dynamic emulators does conform to the conceptual framework used here, and this paper sets out a framework (DEMo) for developing emulators whose input parameters are specified at different points in time.

Liu and Pender (2015) describe the application of a support vector regression (SVR) emulator to a flood inundation model. The results of the study are presented as a time series of depths and velocities of flooding. It is important to recognise that the underlying emulator is, however, static. It takes as input a time series of flood depths and velocities, and estimates, using an emulator, at each time step. The point here is that it would not be appropriate to use a different shape of input hydrograph, and the same emulator, to predict the flood extent of the revised hydrograph.

Other examples of static emulators being used to generate time series output are described in a series of papers (Rueda and others (2017), Antolinez and others (2016), Perez and others, (2015). In this approach, statistical models that incorporate a time component (autoregressive-moving-average (ARMA) type methods) are used to generate a time series of climatological variables that are then translated to daily and hourly offshore sea conditions; emulators are then applied to transform the time series data into a time series of nearshore sea conditions and coastal flooding and coastal process variables. This type of time series data can be applied in the context of continuous simulation risk analysis methods, which have been the subject of significant research in hydrology (for example, McMillan and Brasington, 2008).

6 Climate change

6.1 Introduction

The project steering group highlighted a specific interest in climate change and emulators. It is understood that this was motivated by considering applying emulators relating to climate change, for extrapolation purposes, on an existing ongoing project.

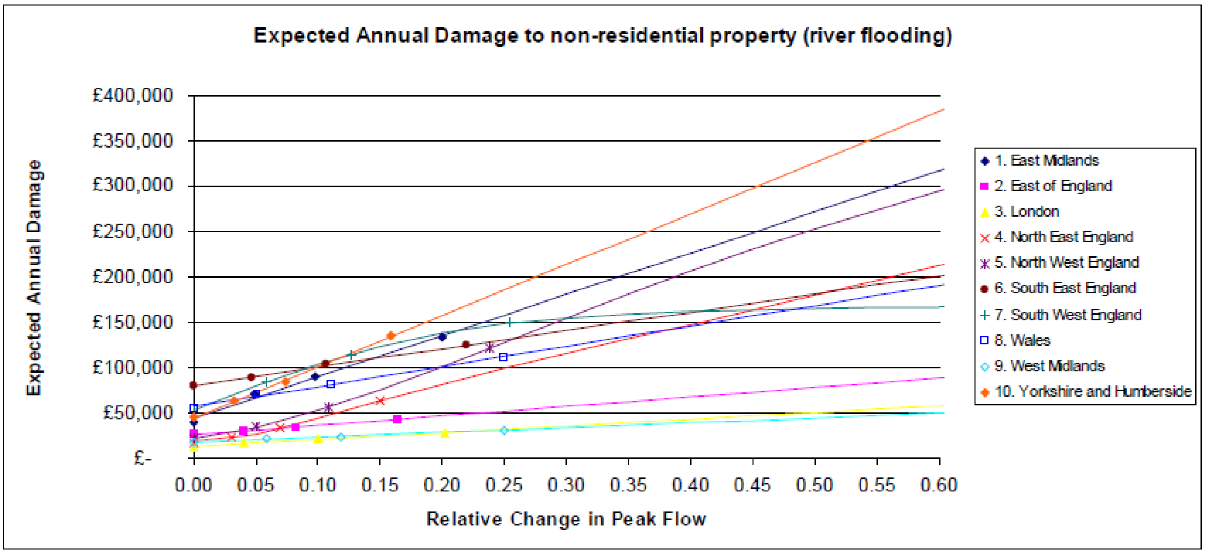
In general, emulators are not applied for the purposes of extrapolation. Extreme value statistics not emulators are used for estimating beyond existing data. Emulators are methods for interpolation, not extrapolation. Different situations can, however, arise in practice and these were described in section 3.3. In particular, there are occasions where emulators have been trained based on existing data, rather than design points being carefully selected to ensure performance of the emulator over the required range of prediction limits. Section 6.2 describes the case history, and provides evidence, of wave overtopping where emulators were being applied in practice, well outside the region where data were supporting the predictions. Measures that were then introduced to ensure predictions are now within the range of the underlying data were also described in the wave overtopping case study example.

In general terms, small deviations outside the strict range of design points may be justifiable and acceptable in practice in some circumstances. However, care should be taken here. Expecting emulators to generate reliable results well beyond the limits of the design points is inadvisable.

Despite this discussion, this review found evidence of emulators being applied in the context of climate change, and this is described in more detail below.

6.2 Source-pathway-receptor-consequence emulators

The government’s first Climate Change Risk Assessment (CCRA) made wide use of response functions to assess climate change impacts for each sector (Defra, 2012). In flooding, these were typically one dimensional, relating a flood driver to flood consequence (for example, expected annual damage (EAD) and number of properties flooded), as shown in Figure 22.



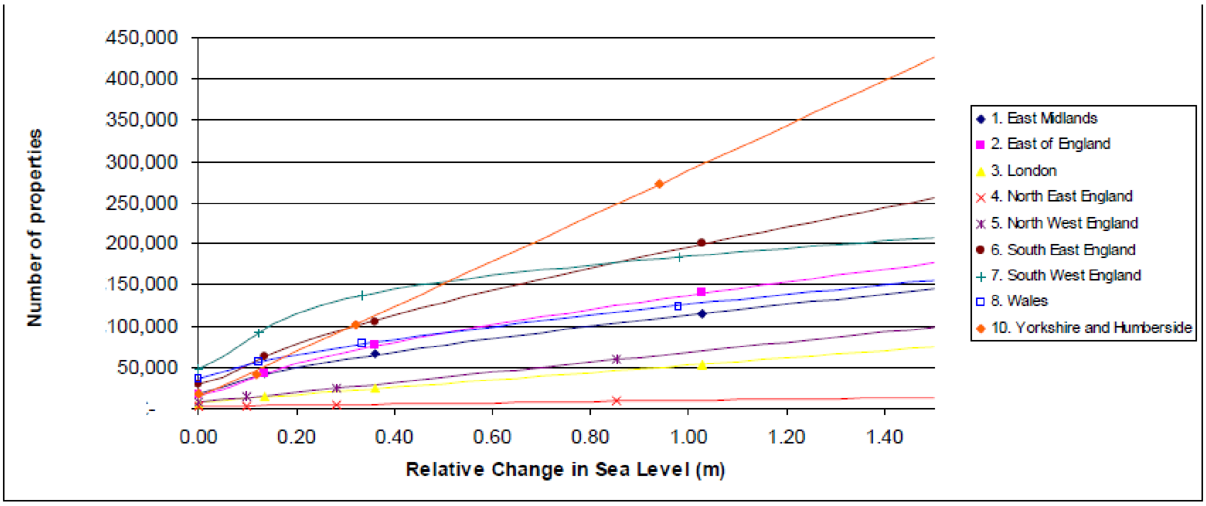


Figure 22: Example response function (emulator) derived in the first CCRA (Defra, 2012) for 10 regions across England and Wales

Figure 22 shows 2 graphs. The top graph illustrates the relative change in peak flow (on the x-axis) against expected annual damages (in GBP on the y-axis). The data shows an increase in damages as peak flow increases across all regions of England and Wales. The second graph shows a relative change in sea level (in meters on the x-axis) against number of properties flooded (on the y-axis). The data shows an increase in properties flooded as sea level increases across all regions of England and Wales.

This type of emulator is seeking to replicate the whole of the source-pathway-receptor-consequence modelling system. The input to the emulator in this example is a single 'source' variable (for example, sea level or peak flow) and the output is the flood risk. The underlying data points were derived using the Environment Agency's long-term investment scenarios (LTIS) model simulations that involved applying the NaFRA model. As the emulators are constructed independently of any climate change scenario, they allow a fast evaluation of different climate change scenarios that are within the limits of the LTIS model runs.

One of the main drawbacks of this approach is that because the whole model system has been emulated, it is not always straightforward to explore mitigation options. For example, using only the relationship depicted in Figure 22, it would not be possible to explore mitigation options that involved raising flood defences. A new set of emulator design points that incorporated different levels of raised defences would be required.

Examples of the application of combined evolutionary optimisation algorithms and emulators were described in the context of emulator uncertainty (section 6.1). This combination has also been applied in the context of options appraisal, considering climate change.

Woodward and others (2013) describe the development of an emulator of a risk-based probabilistic model, based on the Environment Agency's MDSF2 model, for analysis of different flood risk mitigation options for an area on the Thames Estuary. The input to the emulator was sea level rise and the output was flood risk defined in terms of expected annual damage. This is shown in Figure 23.

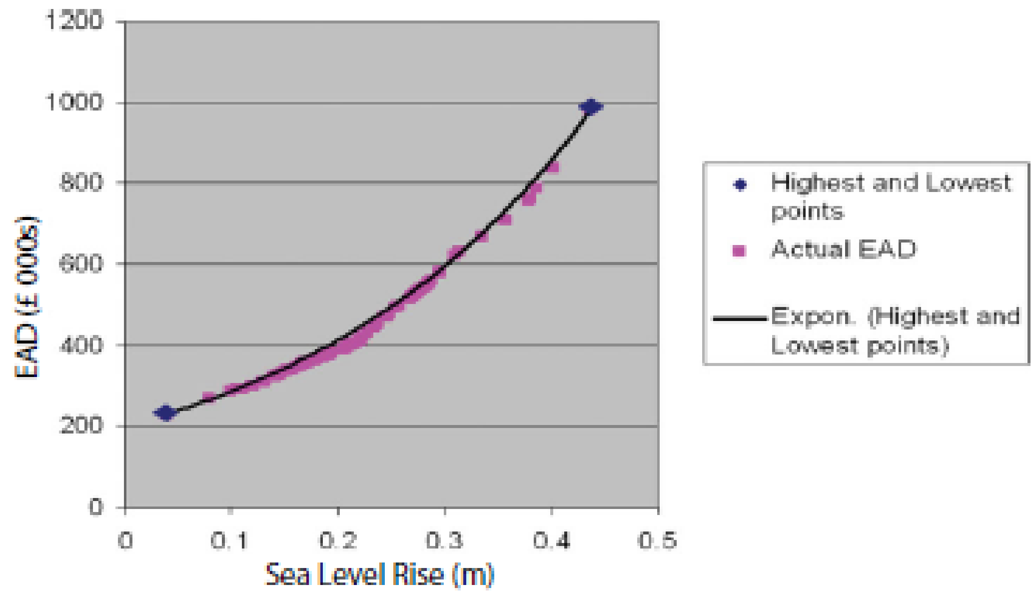


Figure 23: An emulator of a probabilistic risk analysis model to support climate change risk mitigation optimisation

The graph in Figure 23 plots sea level rise (in meters) is plotted on the x-axis against EAD (in GBP) on the y-axis. It shows results of Woodware and others (2013) with good agreement between emulator output and actual annual damages.

Identifying an optimum solution, in terms of economic benefits and costs, is challenging because there are many different mitigation options that can be implemented at many different points in time, and multiple climate change scenarios. In principle, this requires the risk analysis model to be run for many 100s or 1,000s of scenarios (combinations of different mitigation options, intervention time points and climate change scenarios). The combination of the evolutionary optimisation algorithm and the emulator dramatically reduced computational time in this example.

The optimisation algorithm makes the search more efficient, and the emulator allows different potential solutions to be rapidly evaluated. In a similar way to the Khu and others (2004) calibration described in section 6.2, this also involved a final iteration stage that replaced the optimum solution chosen by the emulator with the risk analysis model.

6.3 Flood “source” emulators

Climate change emulators that focus on the flood ‘source’ component have also been developed. These offer more flexibility in terms of exploring mitigation measures than those that are constructed on the whole system. An example of an emulator that focuses on the pluvial or fluvial flood ‘source’ described by Prudhomme and others (2013) is shown in Figure 24.

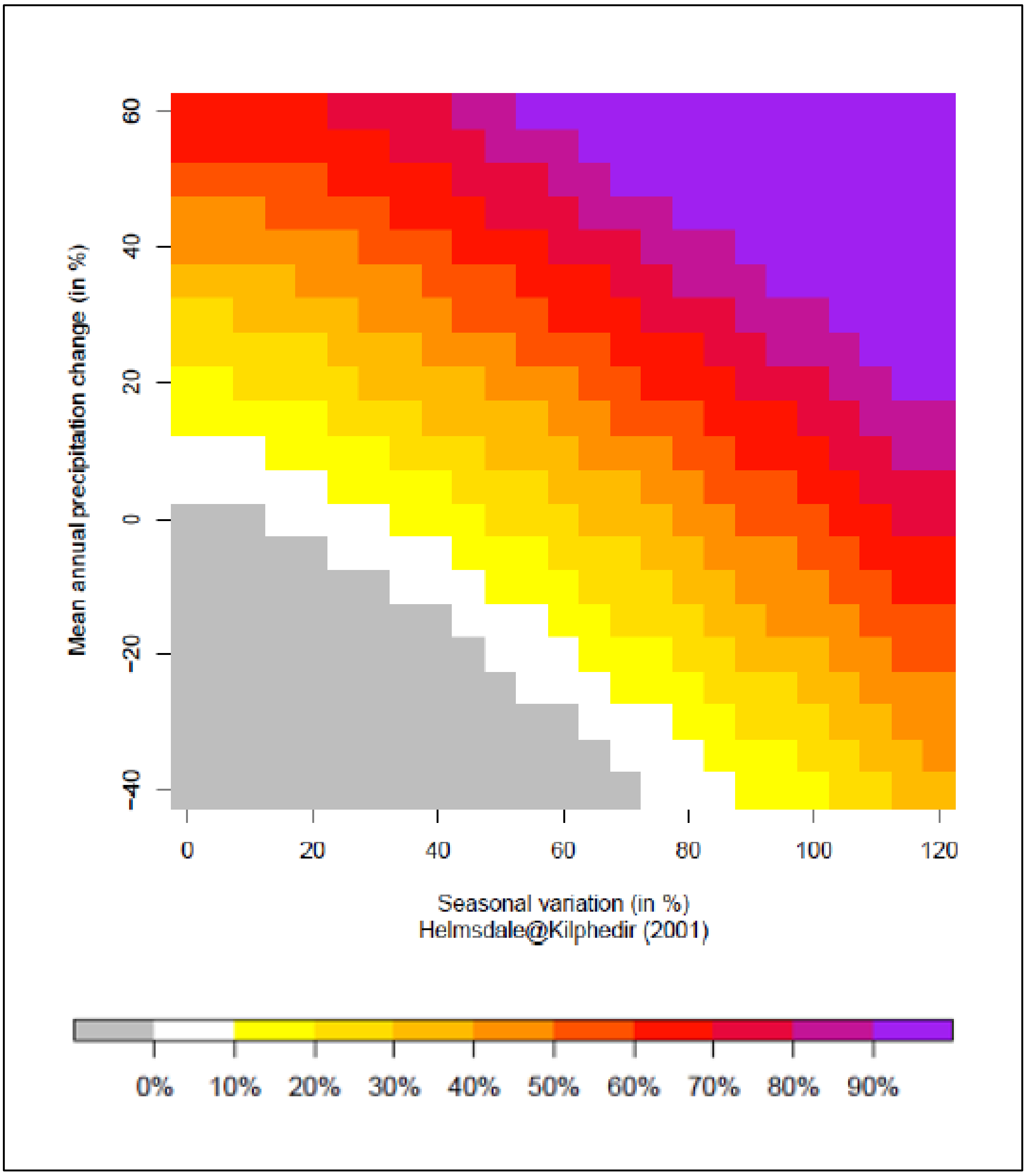


Figure 24: Example peak fluvial flow response function (Prudhomme and others, 2013)

Figure 24 shows a graphical illustration (response surface) for different catchment types of an emulator of mean annual precipitation change, plotted against seasonable variation in precipitation. It shows the percentage change in the 20-year return period peak fluvial flow in relation to seasonal variation of precipitation (on the x-axis) and mean annual precipitation (on the y-axis). Where there is a large (for example 70%) seasonal variation in precipitation, and this relates to a moderate (for example 40%) mean annual change in precipitation, then the predicted change in the 20-year return period peak flow rate is a 75% increase. The figure also shows an example response surface for a specific catchment type. It plots mean annual precipitation change (%) against seasonal variation (%). Nine different catchment types were identified in this study. These relationships were developed to be independent of a climate change scenario. Therefore, they can be readily applied to obtain rapid results for different climate change scenarios.

The subsequent integration of the response surfaces with climate change projections described by Kay and others (2014) are shown in Figure 25. It overlays the output from UKCP09 probabilistic projections onto the response surface.

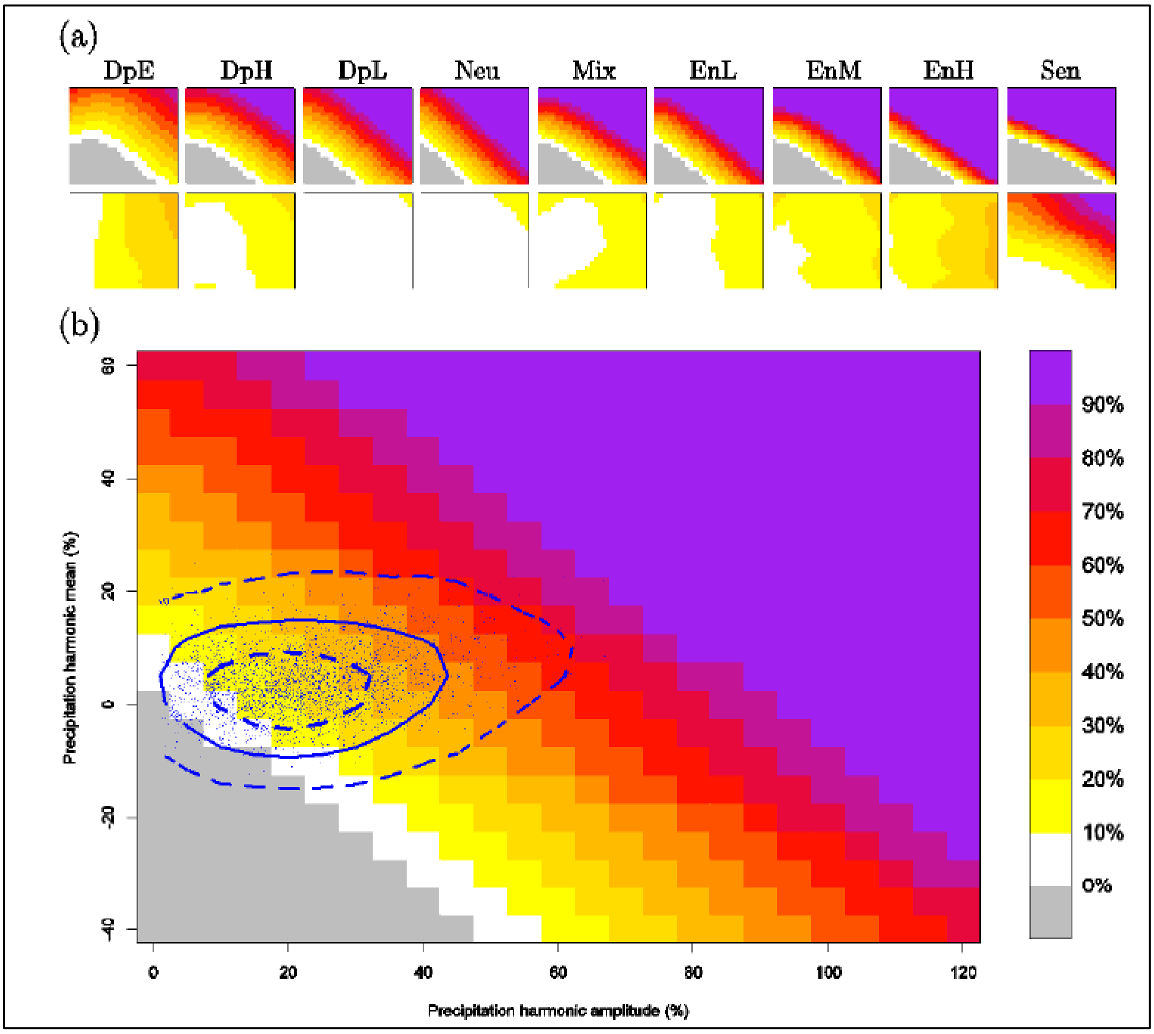


Figure 25: Integration of climate change projections with peak 20-year flow response function (Kay and others, 2014)

In Figure 25 percentage precipitation harmonic mean (y-axis) is plotted against percentage precipitation harmonic amplitude (x-axis). The contours are probability density, and each individual point represents a single climate change projection. The integration of the probability density with the response surface gives a probability density function of the percentage change in the peak 20-year flow rate.

In this example, the emulator replaces the need to run a gridded hydrological flow model for every single UKCP09 climate change projection, thereby significantly reducing the computational burden.

More recently, in relation to coastal climate change and sea level rise, response functions for wave overtopping rates as a result of sea level rise have been generated (Hames and others, 2021). This has been done using the coastal modelling system that was developed and applied in the Environment Agency (Gouldby and others, 2017). The modelling system, for reasons of computational efficiency and dimension reduction, comprises 2 separate stages of integration (Figure 26). The first integration stage comprises 2 separate emulators of 2 separate physical processes; wave transformation and wave overtopping.

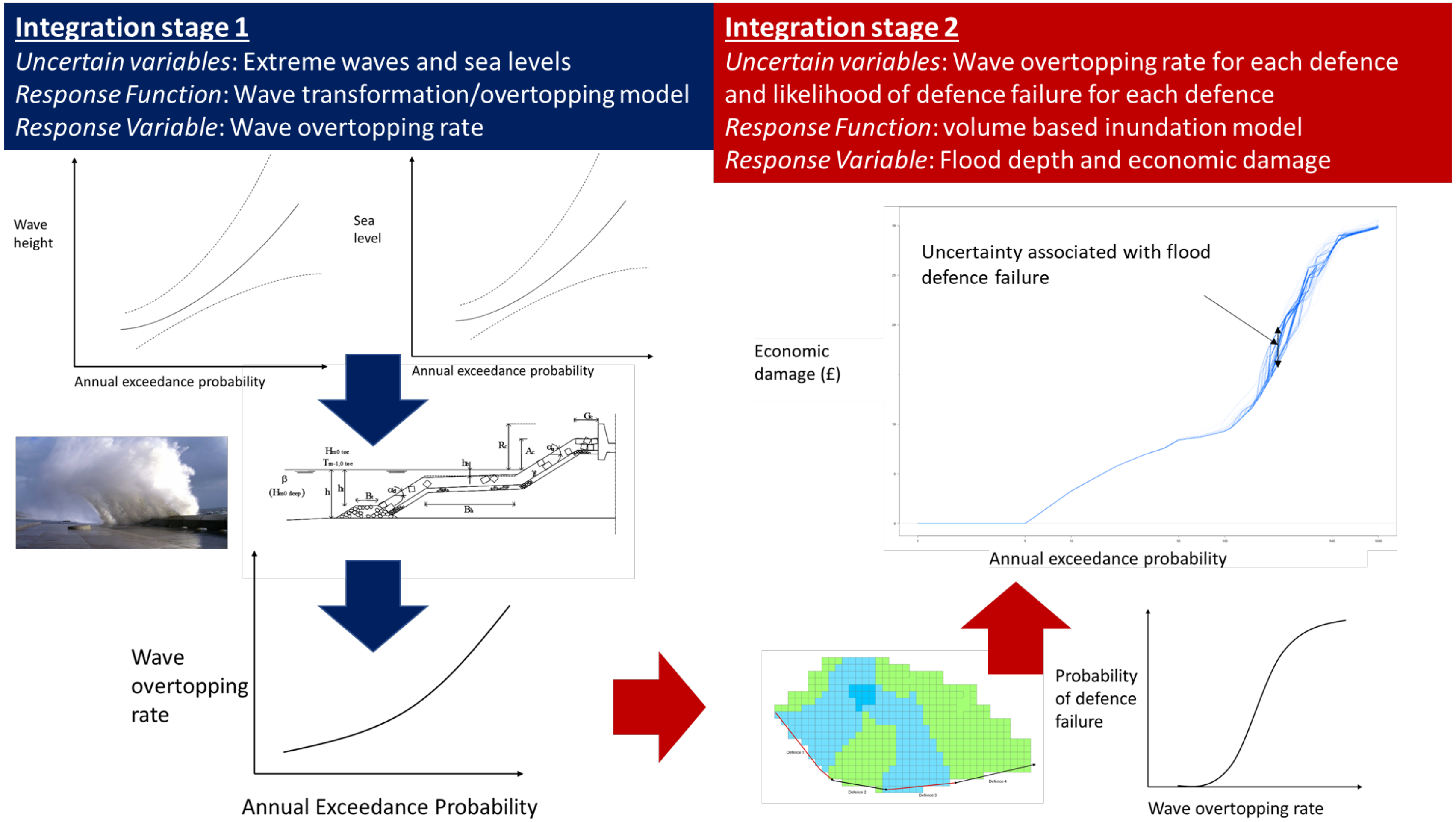


Figure 26: Illustration of the stages involved in the Environment Agency's NaFRA coastal flood risk analysis. Two emulators are applied: one for wave transformation and the other for wave overtopping

Figure 26 illustrates a conceptual flow diagram showing the use of wave transformation and wave overtopping emulators in NaFRA coastal flood risk analysis. It firstly integrates extreme waves and sea levels to produce a response function through a wave transformation model - or wave overtopping model - leading to the response variable of the overtopping rate. Secondly, it takes that overtopping rate for each defence and the likelihood of failure to produce a response function of the volume-based inundation. This further produces the variable of flood depth and economic damage.

The emulator of the SWAN wave transformation model (Booij, 1999) was developed by carefully selecting design points to cover the input parameter space (see Figure 4). This enabled computationally efficient transformation of offshore Monte Carlo event sets (1,000s of sea condition events) to the nearshore. The second emulator in the modelling chain was the BAYONET wave overtopping model (Kingston and others, 2008).

The climate change emulator that was developed (Hames and others, 2021) is therefore an example of an emulator of a modelling chain that comprises of 2 emulators (an emulator of emulators). When developing the climate change emulator, a decision was needed about when to apply estimates of sea level rise. The possible options were:

- offshore, so the wave transformation aspects included the sea level rise
- nearshore, ignoring the deep water wave transformation effects

To test this, a sensitivity study was undertaken that involved adding 1m of sea level rise offshore and then transforming the wave conditions to the nearshore. These results were then compared with the distribution of nearshore sea conditions with no sea level rise added. The outcome of the sensitivity study indicated that deep water wave transformation effects were not overly significant. In the light of additional sensitivity analysis that had already been carried out - and that showed the contribution of the uncertainty in wave transformation to be a small fraction of the contribution to the overall uncertainty in wave overtopping (Figure 14) - it was, for the purposes of the study, considered reasonable to apply the sea level rise at the nearshore.

Emulators for different structure types were developed for different regions around the country. An example of one of these is shown in Figure 27.

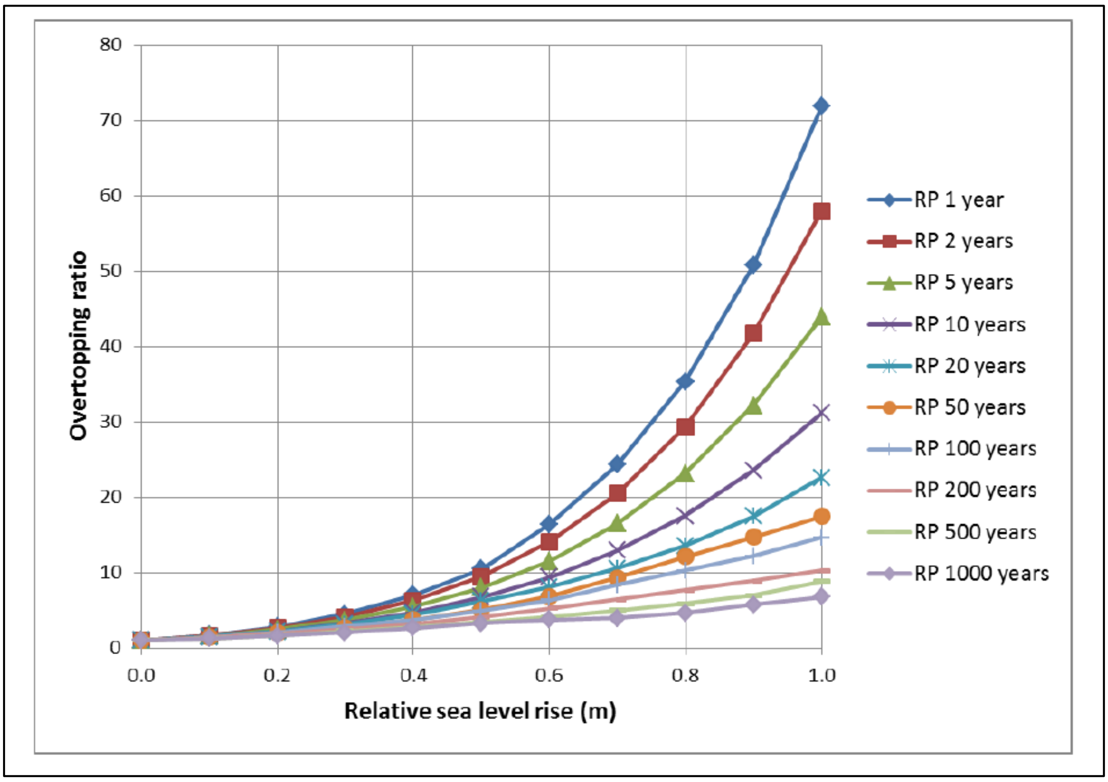


Figure 27: Coastal wave overtopping rate response functions, for different return periods in relation to sea level rise (Hames and others, 2021)

Figure 27 plots overtopping ratio (y-axis) against sea-level rise (in meters, x-axis). Ten different return periods are shown (ranging from 1 year to 1000 years), with the steepest function for return period of 1 year. Each increase in the return period results in a less steep function.

7 Conclusions

Collins and others' (2015) rapid evidence review method was applied to search for evidence of the use of emulators in flood risk analysis. The evidence shows emulators are used extensively within flood risk modelling and analysis.

Three scenarios were identified in relation to using emulators in practice. The scenarios include developing an emulator:

- for a new application where the required predictive limits are known in advance of the design process
- for a new application where the required predictive limits are not known in advance of the design process
- based on existing data or model outputs

The review highlighted that emulators are used more widely within coastal modelling than hydrological or hydraulic modelling. This could be the result of the traditional and widespread use of reduced physics or parametric hydrological models. These models are naturally more computationally efficient and therefore, traditionally emulators may not have been in high demand. In contrast, however, even early wave transformation models were grid based, covering extensive spatial areas. This, together with the early requirement for time series analysis, and Monte Carlo simulation data sets, output from joint probability studies may have significantly stimulated demand for emulators.

In terms of the source-pathway-receptor components, there has been little demand for emulators of the 'consequence' component. This is to be expected as, in general terms, these models are already computationally efficient and therefore emulators are not required.

A case study example relating to wave overtopping highlighted 4 points:

- early applications of emulators were generally not based on considering the different types of emulators, and the selected method was not the result of an informed decision; knowledge of a particular type of emulator that was applied in a certain field was a contributing factor in the choice of emulator (Van-Gent, 2007)
- limitations relating to the potential range of application in practice was not duly considered during the early stages of development
- methods to guide users to ensure emulators are applied within appropriate ranges have been developed and applied in practice, in response to point 2 (Pullen and others, 2018)
- recent developments have made the specific capability of different emulators clearer; in this example, the ability to capture uncertainty in the design points was an important consideration (Pullen and others, 2018)

Evidence of exploring the performance of emulators in terms of trading off the error introduced with the number of design point simulations has been described. An example of this, that uses a leave-one-out cross-validation approach, was described in the context of wave transformation modelling (Malde and others, 2016).

It has, however, also been identified, that it is important to consider the accuracy or uncertainty of any emulator in the context of the overall analysis. Where chains of models are created, some sources of uncertainty can dominate other sources. An example of this was demonstrated by comparing wave transformation model uncertainty with overtopping model uncertainty (Environment Agency, 2020).

Emulators that produce time series output were identified, but these tended to be static emulators applied to time series data. The emulators tended not to be formulated in terms of variables or parameters that vary in time. A framework for the latter has, however, been identified (Castelleti and others, 2015) and this describes the challenges of this type of approach.

The review found significant evidence of emulators being combined with evolutionary optimisation algorithms (artificial intelligence). This combination was used to:

- analyse the catastrophic failure of a flood defence structure during Hurricane Katrina: in this example, the optimisation algorithm was used to focus the selection of the design points for the emulator in a region of values that was of most importance (Kingston and others, 2011) - this was applied to a problem where the area of importance for the design points was not known in advance
- calibrate a MIKE 11 rainfall or run-off model, to minimise the computation time required to obtain an optimum calibration parameter set (Khu and others, 2004)
- select an optimum flood mitigation strategy by emulating a probabilistic risk analysis model (Woodward and others, 2013) - the study considered a range of different mitigation options, climate change scenarios and time horizons for option implementation

Significant evidence was found for applying emulators to assess climate change impacts. The evidence included 2 distinct categories:

1. Emulators of whole system source-pathway-receptor models.
2. Emulators of the 'source' component.

The whole system source-pathway-receptor emulators allow a range of different climate change scenarios to be rapidly considered. However, they offer limited flexibility in exploring mitigation options unlike the source component models.

8 Future considerations

This review has identified that emulators are widely applied within models of flood and coastal erosion risk analysis. It was apparent that there was no readily available industry guidance document that covered the application of emulator techniques to flood risk modelling and analysis.

For potential future use of emulators for flood modelling, a first step would be to develop appropriate guidance covering:

- the range of available techniques, with a discussion of the pros and cons of each
- a standard template for developing an emulator, which could include:
    - several input parameters
    - several output parameters
    - several design points
    - a method for selecting design points
    - a method of validating emulator
    - error metrics
- basis for error acceptance, while considering the context of overall application, including model chain uncertainties
- description modelling problems that are or are not suitable for the application of emulators
- methods to advise users when predictions are not supported by design points (limits of application)
- overview of software libraries, capabilities and verification

References

ANDRIANAKIS, I. AND CHALLENOR, P.G., 2012. The effect of the nugget on Gaussian process emulators of computer models. Computational Statistics and Data Analysis 56(12): 4215–4228, https://doi.org/10.1016/j.csda.2012.04.020.

ANTOLINEZ, J.A.A., MENDEZ, F.J., CAMUS, P., VITOUSEK, S., GONZALEZ, E.M., RUGGIERO, P., BARNARD, P., 2016. A Multi-scale Climate Emulator for Long-term Morphodynamics (MUSCLE-morpho), Journal of Geophysical Research, 10.1002/2015JC011107.

BOOIJ, N., RIS, R.C. AND HOLTHUIJSEN, L.H., 1999. A third-generation wave model for coastal regions: 1. Model description and validation. Journal of Geophysical Research: Oceans, 104(C4): 7649-7666.

CAMUS, P., MENDEZ, FERNANDO, J., MEDINA, RAUL., COFIÑO, ANTONIO S., 2011. Analysis of clustering and selection algorithms for the study of multivariate wave climate, Coastal Engineering, Volume 58, Issue 6, Pages 453-462, ISSN 0378-3839, https://doi.org/10.1016/j.coastaleng.2011.02.003.

CASTELLETTI, A., GALELLI, S., RATTO, M., SONCINI-SESSA, R., YOUNG, P.C., 2012. A general framework for Dynamic Emulation Modelling in environmental problems, Environmental Modelling & Software, Volume 34, 2012, Pages 5-18, ISSN 1364-8152, https://doi.org/10.1016/j.envsoft.2012.01.002.

COLLINS, A., MILLER, J., COUGHLIN, D., KIRK, S., 2015. The Production of Quick Scoping Reviews and Rapid Evidence Assessments: A How to Guide. The Joint Water Evidence Group.

COZZUTO, G., DIMAKOPOULOS, A., DE LATAILLADE, T., MORILLAS, P.O. AND KEES, C.E., 2019. Simulating Oscillatory and Sliding Displacements of Caisson Breakwaters Using a Coupled Approach. Journal of Waterway, Port, Coastal, and Ocean Engineering, 145(2), p.04019005.

DEFRA/ENVIRONMENT AGENCY, 2005. Flood Risks to People: The flood risks to people methodology, Phase 2 FD2321/TR1.

DEFRA, 2012. Climate Change Risk Assessment for the Floods and Coastal Erosion Sector, Defra Project Code GA0204.

DIMAKOPOULOS, A.S., DE LATAILLADE, T. AND KEES, C.E., 2019. Fast random wave generation in numerical tanks. Proceedings of the Institution of Civil Engineers-Engineering and Computational Mechanics, pp.1-11.

DODD, N., 1998. A numerical model of wave run-up, overtopping and regeneration. Proc ASCE, Jo. Waterway, Port, Coast & Ocean Eng., Vol 124, No 2, pp 73-81, ASCE, New York.

ENVIRONMENT AGENCY, 2004. National Flood Risk Assessment 2004, Supported by the RASP HLM plus: Methodology.

ENVIRONMENT AGENCY, 2020. Coastal flood risk modelling: Uncertainty and sensitivity analysis, Horizon House.

ENVIRONMENT AGENCY, 2016. Standards for modelling and forecasting flooding on open coasts, report prepared for the Environment Agency by CH2M.

EUROTOP, 2018. Manual on wave overtopping of sea defences and related structures. An overtopping manual largely based on European research, but for worldwide application. VAN DER MEER, J.W., ALLSOP, N.W.H., BRUCE, T., DE ROUCK, J., KORTENHAUS, A., PULLEN, T., SCHÜTTRUMPF, H., TROCH, P. AND ZANUTTIGH, B., www.overtopping-manual.com.

FARAVELLI, L., 1989. Response-Surface Approach for Reliability Analysis, Journal of Engineering Mechanics Vol. 115, Issue 12.

GARCIA, N., LARA, J.L., LOSADA, I.J., 2004. 2-D Numerical analysis of near-field flow at low –crested permeable breakwaters. Coastal Engineering 51, 991-1020.

GOULDBY, B., MENDEZ, F.J., GUANCHE, Y., RUEDA, A., MINGUEZ, R., 2014. A methodology for deriving nearshore sea conditions for structural design and flood risk analysis. Coastal Engineering, 88. pp. 15-26.

GOULDBY, B.P., WYNCOLL, D., PANZERI, M., FRANKLIN, M., HUNT, T., HAMES, D., TOZER, N.P., HAWKES, P., DORNBUSCH, U. AND PULLEN, T.A., 2017. Multivariate extreme value modelling of sea conditions around the coast of England. Proceedings of the Institution of Civil Engineers - Maritime Engineering, 170 (1). pp. 3-20.

HAMES, 2021. Impacts of Sea Level Rise on wave overtopping rates around the coast of England, ICE Mar. Eng., in press.

HANKIN, R.K.S., 2005. Introducing BACCO, an R bundle for Bayesian analysis of computer code output. Journal of Statistical Software, 14(16).

HASTIE, T., TIBSHIRANI, R. AND FRIEDMAN, J., 2013. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer New York, ISBN 9780387216065.

HAWKES, P.J., GOULDBY, B.P., TAWN, J.A. AND OWEN, M., 2002. The joint probability of waves and water levels in coastal defence design. J. of Hyd. Res. Vol.40, Issue 3.

HR WALLINGFORD/LANCASTER UNIVERSITY, 1998. The joint probability of waves and sea levels: JOIN-SEA, A rigorous but practical new approach, HR Wallingford Report SR537.

HR WALLINGFORD, 2002. Joint Probability of Sea Levels and Fluvial Flows in the Clyde Estuary, Report EX 4564.

HU, KEMING, MINGHAM, CLIVE G. AND CAUSON, DEREK M., 2000. Numerical simulation of wave overtopping of coastal structures using the non-linear shallow water equations. Coastal engineering, 41 (4). pp. 433-465. ISSN 0378-3839.

HUBBARD, M., DODD, N., 2002. A 2D numerical model of wave run-up and overtopping. Coastal Engineering Volume 47, Issue 1, November 2002, Pages 1-26.

KAY, A.L., CROOKS, S.M., DAVIES, H.N., PRUDHOMME, C., REYNARD, N.S., 2014. Probabilistic impacts of climate change on flood frequency using response surfaces I: England and Wales. Regional Environmental Change, 14 (3). 1215-1227. 10.1007/s10113-013-0563-y.

KHUU S.T., SAVIC, D., LIU, Y AND MADSEN, H., 2004. A fast Evolutionary-based Meta-Modelling Approach for the Calibration of a Rainfall-Runoff Model, 2nd International Congress on Environmental Modelling and Software - Osnabrück, Germany - June 2004.

KINGSTON, G., ROBINSON, D. AND GOULDBY, B., 2008. Reliable prediction of wave overtopping volumes using Bayesian neural networks", Proc. of FLOODrisk 2008, 30 Sep - 2 Oct, Oxford, UK In: Samuels and others (eds). Flood Risk Management: Research and Practice. Taylor & Francis Group London.

KINGSTON, G., RAJABALINEJAD, M., GOULDBY, B., VAN GELDER, P., 2011. Computational intelligence methods for the efficient reliability analysis of complex flood defence structures, Structural Safety, Volume 33, Issue 1, 2011, Pages 64-73, ISSN 0167-4730, https://doi.org/10.1016/j.strusafe.2010.08.002.

LIU, P.L.-F., LIN, P., HSU, T., CHANG, K., LOSADA, I.J., VIDAL, C., SAKAKIYAMA, T., 2000. A Reynolds averaged Navier-Stokes equation model for non-linear water wave and structure interactions. Proc. Coastal Structures '99, 169-174.

LIU, Y. AND PENDER, G., 2015. A flood inundation modelling using v-support vector machine regression model, Engineering Applications of Artificial Intelligence, Volume 46, Part A, 2015, Pages 223-231, ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2015.09.014.

MALDE, S., WYNCOLL, D., OAKLEY, J., TOZER, N. AND GOULDBY, B. Applying emulators for improved flood risk analysis, FLOODrisk 2016 - 3rd European Conference on Flood Risk Management.

MARTÍN-MARTÍN, A., ORDUNA-MALEA, E., THELWALL, M. DELGADO LÓPEZ-CÓZAR, E., 2018. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories, Journal of Informetrics, Volume 12, Issue 4, 2018, Pages 1160-1177, ISSN 1751-1577, https://doi.org/10.1016/j.joi.2018.09.002.

MATTIS, S.A., KEES, C.E., WEI, M.V., DIMAKOPOULOS, A. AND DAWSON, C.N., 2018. Computational Model for Wave Attenuation by Flexible Vegetation. Journal of Waterway, Port, Coastal, and Ocean Engineering, 145(1), p.04018033.

MCMILLIAN AND BRASINGTON, 2008. End-to-end flood risk assessment: A coupled model cascade with uncertainty estimation, Water Resource. Res., 44, W03419, doi:10.1029/2007WR005995.

MASE, H., SAKAMOTO, M. AND SAKAI, T., 1995. Neural Network for stability analysis of rubble mound breakwaters. Journal of Waterway Port Coastal and Ocean Engineering, November.

MELCHERS, 1999. Structural Reliability Analysis and prediction, Second edition, Wiley.

MYERS, R. Response Surface Methodology, Allyn and Bacon Inc., Boston 1971.

OFFICE OF SCIENCE AND TECHNOLOGY (OST), 2004. Foresight Future Flooding, Flood and coastal defence project of the Foresight programme.

PULLEN, T., LIU, Y., OTINAR-MORILLAS, P., WYNCOLL, D., MALDE, S., GOULDBY, B., 2018. BAYONET GPE: A generic and practical overtopping model that includes uncertainty. Proc. Inst. Civ. Eng. Mar. Eng. https://doi.org/10.1680/jmaen.2017.31.

PENNING-ROWSELL, E.C. AND OTHERS., 2013. Flood and Coastal Erosion Risk Management: A manual for economic appraisal. https://www.mcm-online.co.uk/manual/.

PEREZ, J., MENENDEZ, M., CAMUS, P., MENDEZ, F.J., LOSADA, I.J., 2015. Statistical multi-model climate projections of surface ocean waves in Europe, Ocean Modelling, doi:10.1016/j.ocemod.2015.06.001.

PRUDHOMME, C., CROOKS, S., KAY, A.L., REYNARD, N.S., 2013. Climate change and river flooding: Part 1 Classifying the sensitivity of British catchments. Climatic Change, 119(3-4), 933-948, doi: 10.1007/s10584-013-0748-x.

ROUSTANT, O., GINSBOURGER, D., AND DEVILLE, Y., 2012. DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization. Journal of Statistical Software, 51(1), 1 - 55. doi:http://dx.doi.org/10.18637/jss.v051.i01.

RUEDA, A., HEGERMILLER, C.A., ANTOLINEZ, J.A.A., CAMUS., P., VITOUSEK, S., RUGGIERO, P., BARNARD, P., ERIKSON, L.H., TOMÁS, A., MÉNDEZ, F.J., 2017. Multiscale Climate Emulator of Multimodal Wave Spectra: MUSCLE-spectra. Journal of Geophysical Research, 10.1002/2016JC011957.

SAYERS, P., HORRITT, M., PENNING-ROWSELL, E., MCKENZIE, A., 2017. Climate Change Risk Assessment 2017 Projections of future flood risk in the UK, Project A: Report prepared for the Committee on Climate Change, UK.

VAN-GENT, M., VAN DEN BOOGAARD, H., POZUETA, B., MEDINA, J. Neural network modelling of wave overtopping at coastal structures. Coastal Engineering 54 (2007) 586 – 593.

WOODWARD, M., KAPELAN, Z. AND GOULDBY, B., 2014. Adaptive Flood Risk Management Under Climate Change Uncertainty Using Real Options and Optimization. Risk Analysis, 34: 75-92. https://doi.org/10.1111/risa.12088.

Acknowledgements

The authors would like to acknowledge the support of the Environment Agency project management team, Ian Perkins, Sue Manson, Jenny Kirwan-Evans and Sally Brown. They would also like to acknowledge the contributions from the project steering group: Jerzy Banasik (Environment Agency), Mike Vaughan (Environment Agency), Martin Borthwick (Environment Agency), Jeremy Oakley (University of Sheffield), Mark McLaughlin (SEPA), Fiona McLay (SEPA) and JP Gosling (JBA Risk Management).

List of abbreviations

AEP	Annual exceedance probability
ANN	Artificial neural networks
CCRA	Climate Change Risk Assessments
EAD	Expected annual damages
FCERM	Flood and coastal erosion risk management
GA	Genetic algorithm
GPE	Gaussian process emulators
LSE	Limit state equation
MDA	Maximum dissimilarity algorithm
NaFRA	National Flood Risk Assessment
OST	Office of Science and Technology
PICO	Population, intervention, control and outcome
REA	Rapid evidence assessment
RMSE	Root mean square error
RBF	Radial basis functions
SOP	Standard of protection
SPR	Source, pathway, receptor
WMDA	Weighted maximum dissimilarity algorithm
WOS	Web of Science

Appendix A: Search strings

The 8 search strings used in the REA were:

- ((emulator OR emulation) OR (‘meta-model’ OR ‘meta model’ OR ‘metamodel’) OR ‘surrogate model’ OR ‘response surface methodology’ OR ‘design selection’) AND (‘flood risk’ OR ‘flood hazard’)
- ((emulator OR emulation) OR (‘meta-model’ OR ‘meta model’ OR ‘metamodel’) OR ‘surrogate model’ OR ‘response surface methodology’ OR ‘design selection’) AND (‘flood risk’ OR ‘flood hazard’) AND ‘present and future flood’
- ((emulator OR emulation) OR (‘meta-model’ OR ‘meta model’ OR ‘metamodel’) OR ‘surrogate model’ OR ‘response surface methodology’ OR ‘design selection’) AND (‘flood risk’ OR ‘flood hazard’) AND (‘coastal flood’ OR ‘wave height’ OR surge OR tide OR tidal OR ‘sea level’)
- ((emulator OR emulation) OR (‘meta-model’ OR ‘meta model’ OR ‘metamodel’) OR ‘surrogate model’ OR ‘response surface methodology’ OR ‘design selection’) AND (‘flood risk’ OR ‘flood hazard’) AND (‘pluvial flood’ OR rainfall OR precipitation)
- ((emulator OR emulation) OR (‘meta-model’ OR ‘meta model’ OR ‘metamodel’) OR ‘surrogate model’ OR ‘response surface methodology’ OR ‘design selection’) AND (‘flood risk’ OR ‘flood hazard’)) AND (‘fluvial flood’ OR ‘river flow’)
- ((emulator OR emulation) OR (‘meta-model’ OR ‘meta model’ OR ‘metamodel’) OR ‘surrogate model’ OR ‘response surface methodology’ OR ‘design selection’) AND (‘flood risk’ OR ‘flood hazard’) AND (‘inundation’ OR ‘flood depth’ OR ‘flood velocity’)
- ((emulator OR emulation) OR (‘meta-model’ OR ‘meta model’ OR ‘metamodel’) OR ‘surrogate model’ OR ‘response surface methodology’ OR ‘design selection’) AND (‘flood risk’ OR ‘flood hazard’) AND (‘defence breach’ OR ‘defence failure’)
- ((emulator OR emulation) OR (‘meta-model’ OR ‘meta model’ OR ‘metamodel’) OR ‘surrogate model’ OR ‘response surface methodology’ OR ‘design selection’) AND (‘flood risk’ OR ‘flood hazard’) AND (‘hydrol*’)

Appendix B: Literature database evidence list

Table 2: Evidence list

Paper ID key: SG=Steering Group contribution, GS=Google Scholar, WOS=Web of Science

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
SG_1_3	Roy, Pamphile T., El Moayd, Nabil., Ricci, Sophie., Jouhaud, Jean Christophe., Goutal, Nicole., De Lozzo, Matthias., Rochoux, Melanie C.	2018	Comparison of polynomial chaos and Gaussian process surrogates for uncertainty quantification and correlation estimation of spatially distributed open-channel steady flows	https://link.springer.com/article/10.1007/s00477-017-1470-4
SG_1_4	Gainza, June., Rueda, Ana., Camus, Paula., Tomas, Antonio., Mendez, Fernando J., Sano, Marcello., Tomlinson, Rodger.	2018	A meta-modelling approach for estimating long-term wave run-up and total water level on beaches	https://www.researchgate.net/publication/323621767_A_Meta-Modelling_Approach_for_Estimating_Long-Term_Wave_Run-Up_and_Total_Water_Level_on_Beaches

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
SG_1_5	Kingston, G.B., Rajabalinejad, M., Gouldby, B.P., Van Gelder, P.H.A.J.M.	2011	Computational intelligence methods for the efficient reliability analysis of complex flood defence structures	https://www.sciencedirect.com/science/article/pii/S0167473010000767
SG_1_6	Rueda, Ana., Cagigal, Laura., Pearson, Stuart., Antolinez, Jose A.A., Storlazzi, Curt., van Dongeren, AP., Camus, Paula., Mendez, Fernando J.	2019	HyCReWW: a hybrid coral reef wave and water level metamodel	https://doi.org/10.1016/j.cageo.2019.03.004
SG_1_7	Rueda, Ana; Hegermiller, Christie A; Antolinez, Jose Antonio A.; Camus, Paula; Vitousek, Sean; Ruggerio, Peter; Barnard, Patrick L.; Erikson, Tomas; Mendez, Fernando J.	2016	Multiscale climate emulator of multimodal wave spectra: muscle-spectra	https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016JC011957
SG_1_8	Anderson, D., Rueda, A.	2019	Time varying emulator for short and long term analysis of coastal flood hazard potential journal of geophysical research: oceans	https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JC015312

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
SG_1_10	Gouldby, B., Wyncoll, D., Panzeri, M., Franklin, M., Hunt, T., Hames, D., Tozer, N., Hawkes, P., Dornbusch, U., Pullen, T.	2017	multivariate extreme value modelling of sea conditions around the coast of England	https://www.icevirtuallibrary.com/doi/full/10.1680/jmaen.2016.16
SG_1_11	Hall, Jim W., Manning, Lucy J., Hankin, Robin K.S.	2011	Bayesian calibration of a flood inundation model using spatial data	https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2009WR008541
SG_1_12	Malde, S., Wyncoll, D., Oakley, J., Tozer, N., Gouldby, B.	2016	applying emulators for improved flood risk analysis	https://www.e3s-conferences.org/articles/e3sconf/abs/2016/02/e3sconf_flood2016_04002/e3sconf_flood2016_04002.html
SG_1_13	Camus, Paula., Mendez, Fernando J., Medina, Raul.	2011	A hybrid efficient method to downscale wave climate to coastal areas	http://dx.doi.org/10.1016/j.coastaleng.2011.05.007
SG_1_14	Young, P.C., Ratto, Marco.	2009	A unified approach to environmental systems modeling	https://www.researchgate.net/publication/225747621_A_unified_approach_to_environmental_systems_modeling

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
SG_1_15	Camus, Paula., Mendez, Fernando J., Medina, Raul., Cofio, Antonio S.	2011	Analysis of clustering and selection algorithms for the study of multivariate wave climate	https://www.sciencedirect.com/science/article/pii/S0378383911000354
SG_1_16	Antolnez, Jose Antonio A., Murray, A. Brad., Mendez, Fernando J., Moore, Laura J., Farley, Graham., Wood, James.	2018	Downscaling changing coastlines in a changing climate: the hybrid approach	https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017JF004367
SG_1_18	Harari, O., Bingham, D., Dean, A., Higdon, D.	2018	Computer experiments: prediction accuracy, sample size and model	https://www.jstor.org/stable/44841930?seq=1#metadata_info_tab_contents
SG_1_19	Morris, Max D., Mitchell, Toby J.	2008	Exploratory designs for Computational experiments	https://www.sciencedirect.com/science/article/pii/037837589400035T
SG_1_21	Young, P.C., Ratto, M.	2011	Statistical emulation of large linear dynamic models	https://www.tandfonline.com/doi/abs/10.1198/TECH.2010.07151
SG_1_22	Young, Peter C., Leedal, David., Beven, Keith J., Szczypta, Camille.	2009	Reduced order emulation of Distributed hydraulic simulation models	https://www.sciencedirect.com/science/article/pii/S1474667016389078

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
SG_1_23	Tych, W., Young, P.C.	2012	A MATLAB software framework for dynamic model emulation	https://www.sciencedirect.com/science/article/pii/S1364815211001915
SG_1_24	Gouldby, B and Hawkes, P.	2002	Joint probability of sea levels and fluvial flows in the Clyde estuary	https://www.estuary-guide.net/pdfs/FD2308_3429_TRP.pdf
SG_1_25	Pullen, Tim., Liu, Ye., Morillas, Pedro Otinar., Wyncoll, David., Malde, Sajni., Gouldby, Ben.	2018	A generic and practical wave overtopping model that includes uncertainty	https://www.icevirtuallibrary.com/doi/10.1680/jmaen.2017.31
SG_1_26	Ramsbottom, D., Sayers, P. and Panzeri, M.	2012	Climate Change Risk Assessment for the Floods and Coastal Erosion Sector	N/A
SG_1_27	Prudhomme. C., Crooks, S., Kay, A.L., Reynard, N.S.	2013	Climate change and river flooding: Part 1 Classifying the sensitivity of British catchments	http://nora.nerc.ac.uk/id/eprint/501334/2/N501334PP.pdf

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
SG_1_28	Kay, A.L., Crooks, S.M., Davies, H.N., Prudhomme, C., Reynard.	2014	Probabilistic impacts of climate change on flood frequency using response surfaces I: England and Wales	http://nora.nerc.ac.uk/id/eprint/502207/1/N502207PP.pdf
GS_1_1	Anderson, D., Rueda, A., Cagigal, L.	2019	Time varying emulator for short and long term analysis of coastal flood hazard potential	https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JC015312
GS_1_101	Liu, Y., Pender, G.	2015	A flood inundation modelling using v-support vector machine regression model	https://www.sciencedirect.com/science/article/pii/S0952197615002146
GS_1_104	Rohmer, J., Idier, D., Picheny, V.	2012	A meta-modelling strategy to identify the critical offshore conditions for coastal flooding.	https://www.researchgate.net/profile/Deborah_Idier2/publication/235932802_A_meta-modelling_strategy_to_identify_the_critical_offshore_conditions_for_coastal_flooding/links/0046352cea01c99726000000.pdf

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_1_11	Rohmer, K., Louisor, J., Idier, D.	2017	Boosting probabilistic coastal flood hazard assessment by combining extreme value analysis, full-process based models & metamodels	https://hal.archives-ouvertes.fr/hal-01546349/
GS_1_113	Bermdez, M., Cea, L., Puertas, J.	2019	A rapid flood inundation model for hazard mapping based on least squares support vector machine regression	https://onlinelibrary.wiley.com/doi/abs/10.1111/jfr3.12522
GS_1_136	Rueda Zamora, A.	2016	Methodologies for coastal flooding risk analysis	https://repositorio.unican.es/xmlui/handle/10902/9733
GS_1_156	Brown, S., Nicholls, R., Lazar, A., Hornby, D.D.	2018	What are the implications of sea-level rise for a 1.5, 2 and 3° c rise in global mean temperatures in the Ganges-Brahmaputra-Meghna and other vulnerable	https://link.springer.com/article/10.1007/s10113-018-1311-0

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_1_188	Richards, J., Mokrech, M., Berry, P., Nicholls, R.	2008	Regional assessment of climate change impacts on coastal and fluvial ecosystems and the scope for adaptation	https://link.springer.com/content/pdf/10.1007/s10584-008-9451-8.pdf
GS_1_195	Lopez-Lopera, A., Idier, D., Rohmer, J.	2020	Multi-output Gaussian processes with functional data: a study on coastal flood hazard assessment	https://arxiv.org/abs/2007.14052
GS_1_2	Gouldby, B., Mendez, F., Guanche, Y., Rueda, A.	2014	A methodology for deriving extreme nearshore sea conditions for structural design and flood risk analysis	https://www.sciencedirect.com/science/article/pii/S0378383914000210
GS_1_203	Rueda, A., Vitousek, A., Camus, P., Tomas, A., Espejo, A.	2017	A global classification of coastal flood hazard climates associated with large-scale oceanographic forcing	https://www.nature.com/articles/s41598-017-05090-w

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_1_227	Wong, T., Bakker, A., Keller, K.	2017	Impacts of Antarctic fast dynamics on sea-level projections and coastal flood defense	https://link.springer.com/article/10.1007/s10584-017-2039-4
GS_1_238	Rueda, A., Camus, P., Tomas, A., Vitousek, A., Mendez, F.	2016	A multivariate extreme wave and storm surge climate emulator based on weather patterns	https://www.sciencedirect.com/science/article/pii/S1463500316300592
GS_1_250	Sopelana, J., Cea, L., Ruano, S.	2018	A continuous simulation approach for the estimation of extreme flood Inundation in coastal river reaches affected by meso-and macrotides	https://link.springer.com/article/10.1007/s11069-018-3360-6
GS_1_257	Zhang, J., Taflanidis, A., Nadal-Caraballo, N., Melby, J.	2018	Advances in surrogate modeling for storm surge prediction: storm selection and addressing characteristics related to climate change	https://link.springer.com/article/10.1007/s11069-018-3470-1

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_1_283	Bermdez, M., Ntegeka, V., Wolfs, V., Willems, P.	2018	Development and comparison of two fast surrogate models for urban pluvial flood simulations	https://link.springer.com/article/10.1007/s11269-018-1959-8
GS_1_33	Chu, H., Wu, W., Wang, Q., Nathan, R., Wei, J.	2020	An ANN-based emulation modelling framework for flood inundation modelling: application, challenges and future directions	https://www.sciencedirect.com/science/article/pii/S1364815219307601
GS_1_38	Su, H. Cheung, S., Lo, E.	2018	Multi-objective optimal design for flood risk management with resilience objectives	https://link.springer.com/article/10.1007/s00477-017-1508-7
GS_1_398	Thrys, C., Arnbjerg-Nielsen, K., Borup, M.	2019	Identifying fit-for-purpose lumped surrogate models for large urban drainage systems using glue	https://www.sciencedirect.com/science/article/pii/S0022169418308606

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_1_4	Betancourt, J., Bachoc, F., Klein,T., Idier, D.	2020	Gaussian process metamodeling of functional-input code for coastal flood hazard assessment	https://www.sciencedirect.com/science/article/pii/S0951832019301693
GS_1_422	Woodward, M., Kapelan, Z., Gouldby, B.	2014	Adaptive flood risk management under climate change uncertainty using real options and optimization	https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.12088
GS_1_425	Betancourt, J.	2020	Functional-input metamodeling: an application to coastal flood early warning	https://tel.archives-ouvertes.fr/tel-02917021/
GS_1_432	Li, X., Willems, P.	2020	A hybrid model for fast and probabilistic urban pluvial flood prediction	https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019WR025128
GS_1_58	Tran, V., Kim, J.	2019	Quantification of predictive uncertainty with a metamodel: toward more efficient hydrologic simulations	https://link.springer.com/article/10.1007/s00477-019-01703-0

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_1_6	Bass, B., Bedient, P.	2018	Surrogate modeling of joint flood risk across coastal watersheds	https://www.sciencedirect.com/science/article/pii/S0022169418300143
GS_1_84	Yu, J., Qin, X., Larsen, O.	2015	Uncertainty analysis of flood inundation modelling using glue with surrogate models in stochastic sampling	https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.10249
GS_1_9	Rueda, A., Gouldby, B., Mendez, F.	2016	The use of wave propagation and reduced complexity inundation models and metamodels for coastal flood risk assessment	https://onlinelibrary.wiley.com/doi/abs/10.1111/jfr3.12204
GS_3_13	Jia, G., Taflanidis, A., Nadal-Caraballo, N., Melby, J.	2016	Surrogate modeling for peak or time-dependent storm surge prediction over an extended coastal region using an existing database of synthetic storms	https://link.springer.com/article/10.1007/s11069-015-2111-1

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_3_14	Jia, G., Taflanidis, A.	2013	Kriging metamodeling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment	https://www.sciencedirect.com/science/article/pii/S0045782513000637
GS_3_2	Kim, S., Melby, J., Nadal-Caraballo, N., Ratcliff, J.	2015	A time-dependent surrogate model for storm surge prediction based on an artificial neural network using high-fidelity synthetic hurricane modeling	https://link.springer.com/content/pdf/10.1007/s11069-014-1508-6.pdf
GS_3_69	Han, Z., Lu, W., Lin, J.	2020	Uncertainty analysis for precipitation and sea-level rise of a variable-density groundwater simulation model based on surrogate models	https://link.springer.com/content/pdf/10.1007/s11356-020-09177-2.pdf

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_4_118	Yang, J., Jakeman, A., Fang, G., Chen, X.	2018	Uncertainty analysis of a semi-distributed hydrologic model based on a Gaussian process emulator	https://www.sciencedirect.com/science/article/pii/S1364815217301378
GS_4_131	Song, X., Zhan, C., Xia, J.	2012	Integration of a statistical emulator approach with the SCE-UA method for parameter optimization of a hydrological model	https://link.springer.com/article/10.1007/s11434-012-5305-x
GS_4_155	Reichert, P., White, G., Bayarri, M., Pitman, E.	2011	Mechanism-based emulation of dynamic simulation models: concept and application in hydrology	https://www.sciencedirect.com/science/article/pii/S0167947310003993
GS_4_2	Abrahart, R., See, L.	2007	Neural network emulation of a rainfall-runoff model	https://hal.archives-ouvertes.fr/hal-00298815/

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_4_3	Moreno-Rodenas, A., Bellos, V., Langeveld, J.	2018	A dynamic emulator for physically based flow simulators under varying rainfall and parametric conditions	https://www.sciencedirect.com/science/article/pii/S004313541830455X
GS_4_42	Schnorbus, M., Cannon, A.	2014	Statistical emulation of streamflow projections from a distributed hydrological model: application to cmip3 and cmip5 climate projections for British Columbia	https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014WR015279
GS_4_427	Venkatesh, B., Mahindrakar, A.	NA	Improving the hydrological model efficiency using emulator based optimization (EBO)	https://www.researchgate.net/publication/331112177_Improving_the_hydrological_model_efficiency_using_emulator_based_optimization_EBO
GS_4_45	Machac, D., Reichert, P., Rieckermann, J.	2016	Fast mechanism-based emulator of a slow urban hydrodynamic drainage simulator	https://www.sciencedirect.com/science/article/pii/S1364815215301213

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_4_707	Carbajal, J., Leito, J., Albert, C.	2017	Appraisal of data-driven and mechanistic emulators of nonlinear simulators: the case of hydrodynamic urban drainage models	https://www.sciencedirect.com/science/article/pii/S1364815216307964
GS_4_8	Khu, S., Savic, D., Liu, Y., Madsen, H.	2004	A fast evolutionary-based meta-modelling approach for the calibration of a rainfall-runoff model	https://scholarsarchive.byu.edu/iemssconference/2004/all/81/
GS_4_97	Mahmoodian, M., Carbajal, J., Bellos, V.	2018	A hybrid surrogate modelling strategy for simplification of detailed urban drainage simulators	https://link.springer.com/content/pdf/10.1007/s11269-018-2157-4.pdf
GS_5_2	Kosti, S., Stojkov, M., Prohaska, S.	2016	Modeling of river flow rate as a function of rainfall and temperature using response surface methodology based on historical time series	https://iwaponline.com/jh/article-abstract/18/4/651/30062

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_5_310	Wolfs, V., Willems, P.	2014	Development of discharge-stage curves affected by hysteresis using time varying models, model trees and neural networks	https://www.sciencedirect.com/science/article/pii/S1364815214000310
GS_5_389	Abrahart, R., See, L., Dawson, C.	2010	Nearly two decades of neural network hydrologic modeling	https://www.worldscientific.com/doi/abs/10.1142/9789814307987_0006
GS_5_424	Singh, P., Deo, M.	2007	Suitability of different neural networks in daily flow forecasting	https://www.sciencedirect.com/science/article/pii/S1568494606000536
GS_5_57	Kamali, M., Ponnambalam, K.	2007	Integration of surrogate optimization and PCA for calibration of hydrologic models, a WATCLASS case study	https://ieeexplore.ieee.org/abstract/document/4414073/

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
GS_5_591	Kingston, G., Dandy, G., Maier, H.	2008	Review of artificial intelligence techniques and their applications to hydrological modeling and water resources management: optimization	https://www.researchgate.net/publication/277005048_Review_of_Artificial_Intelligence_Techniques_and_their_Applications_to_Hydrological_Modeling_and_Water_Resources_Management_Part_1_-_Simulation
GS_5_651	Tsoukalas, I., Makropoulos, C.	2015	Multiobjective optimisation on a budget: exploring surrogate modelling for robust multi-reservoir rules generation under hydrological uncertainty	https://www.sciencedirect.com/science/article/pii/S1364815214002874
GS_6_59	Christelis, V., Hughes, A.G.	2018	Metamodel-assisted analysis of an integrated model composition: an example using linked surface water and groundwater models	https://www.sciencedirect.com/science/article/pii/S1364815217313129

Paper_ID	Authors	Year	Title	Hyperlink (last accessed May 2025)
WOS_3_70	Mun, J., Kim, SW., Melby, J.A.	2018	A surrogate model for wave prediction based on an artificial neural network using high-fidelity synthetic hurricane modelling	https://www.researchgate.net/publication/326955224_A_Surrogate_Model_for_Wave_Prediction_Based_on_an_Artificial_Neural_Network_Using_High-fidelity_Synthetic_Hurricane_Modelling
WOS_4_79	Nagel, J.B., Rieckermann, J., Sudret, B.	2020	Principal component analysis and sparse polynomial chaos expansions for global sensitivity analysis and model calibration: application to urban drainage simulation	https://www.sciencedirect.com/science/article/pii/S0951832019301747

Would you like to find out more about us or 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 **(24 hours)**

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

Environment first

Are you viewing this onscreen? Please consider the environment and only print if absolutely necessary. If you are reading a paper copy, please don't forget to reuse and recycle.