Systematic Analysis of Uncertainty in Flood Inundation Modelling

Thomas David Michael Willis

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
Institute of Resilient Infrastructure
School of Civil Engineering

April 2014

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

The author would like to thank;

Professor Nigel Wright for supervision and patience whilst teaching a geography graduate about maths, and appreciating figures with an array of colours.

Dr Andy Sleigh for supervision and for patience whilst teaching a geography graduate about computer programming.

Dr Richard Dun for assistance with the Coventry test case including all the data and use of the hydrograph calculator, Dr Rob Lamb for the Mexborough data set and assistance. Dr Ignacio Villanueva for help in using and understanding LISFLOOD-Roe, Professor Paul Bates for giving insight into the research, Dr Jeff Neal and Dr Tim Fewtrell for help with LISFLOOD-FP.

Thanks to Leeds University and EPSRC for funding through the Flood Risk Management Research Consortium.

Thanks to my family: my parents and my brothers. Thanks to my friends, especially to Sam, Mark and Lim for being good office mates.

Lizzy Moyce for everything.

Abstract

Recent evaluations of 2D models have analysed uncertainty in data inputs into flood models, but have treated the model code as a black box. In this work, the influence of the numerical representation of the model on the results is evaluated. The purpose is not only to understand the significance of the physical scheme in the model on results, but also the importance of this in respect to other known sources of uncertainty, in particular boundary conditions, calibrated parameters such as Manning's friction values, DEM accuracy and other more subjective forms of uncertainty associated with the choices used by modellers in constructing models, such as building representation.

To further explore the impact that the level of physical representation has on model output, models were also analysed using risk and exposure based measures. The methods included vulnerability weighted measures and the use of damage curves from the Multi Coloured Manual.

A series of Monte Carlo tests were undertaken for a range of parameters over 3 test cases using the LISFLOOD-FP code. The LISFLOOD-FP code was chosen as it has several formulations for solving 2D floodplain flow within its framework, each with different level of physical representation.

The test cases included two urban events, a culvert overtopping event in Glasgow and canal embankment failure Coventry, and a river overtopping in Mexborough, Yorkshire a rural urban domain. The test cases provided a wider range of hydraulic conditions and are reflected events typically assessed with inundation models to ensure the effect of model bias was removed from the results.

The results for the test cases indicated that the choice of physical representation was the most critical in affecting model results, particularly for the urban test case. However, the interaction between factors and parameters also indicated that for certain scenarios, this becomes less critical to model results. The use of risk based methods also identified areas of variations between parameters sets and numerical schemes that are not identified with traditional model evaluation techniques.

Table of Contents

List of Tables

List of Figures

1. INTRODUCTION

Flood inundation models are an essential part of analysing risk associated with floods. The use of flood models as a component of risk analysis has increased in recent years with advances in the computational codes, computational power, and an increase in the coverage and resolution of data sets, in particular topographic data sets and data which can be used to calibrate computer models, such as satellite and airborne images. These developments have allowed 2-dimensional modelling to become mainstream, moving away from the traditional 1-dimensional modelling based on river cross sections. A number of different 2D model codes, have been developed to take advantage of these advances, based on different numerical approaches and levels of physical representation. These codes range from commercial software (ISIS, TuFlow), through to research based codes, such as LISFLOOD-FP (Bates and De Roo 2000), BreZo (Sanders et al 2000) and SFV (Horritt 2004). The codes are used in a wide variety of tasks and for different purposes from research to consultancy. Consequently a number of different approaches are undertaken in the solution of the governing equations to take into account computational costs, model output and the level of detail of input data. As these model codes have been developed, these models have become an increasingly critical component in the development of infrastructure and in risk analysis. Consequently the uncertainty in the model and the impact of it on results becomes increasingly significant.

The creation of a model code for a natural phenomenon is based on the following processes defined in principle by Gupta et al (2012) and Beven et al (2001), where each step provides a level of discretization of the natural process that will eventually form the computational code. A basic outline of this process is defined in Figure 1.1

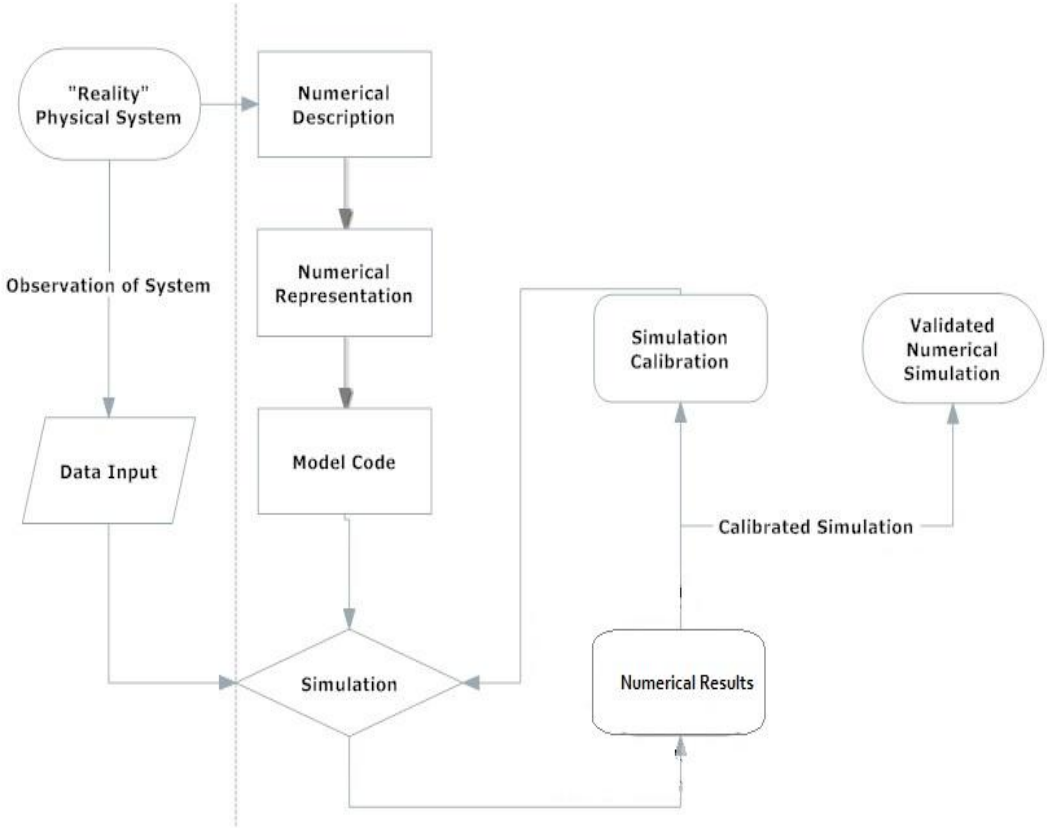

Figure 1.1: Conceptual Model of Model code development

The initial step in the formation of a computational model of a environmental process is to form a numerical description of the physical process, based on the governing equations of the process. The numerical description, also referred to as the conceptual model (Gupta et al 2009.) represents the complete mathematical structure of the process. For flood inundation problems, it is the Shallow Water Equations, (SWE) a special form of the Navier-Stokes equation, which represents the conservation of mass and momentum for 2-dimensional problems, as explained in Section 2. Development of the conceptual model will involve determining which aspects of the numerical description, the governing equations, processes and level of physics is required to model the phenomena. This choice is critical to the further development of the model code, and these decisions should be based on a sound understanding of the physics involved and in determining which aspects may be considered unimportant in the environmental process being modelled. This is based on determining the spatial and temporal scale of the problem and by removing components from the governing equations, a level of physical representation for the final model code is created. The level of physical representation is defined as the total number of equations and terms in the numerical code, and is defined in this research as the number of terms used from the conservation of momentum component of the shallow water equations. The numerical description forms the basis of the mathematical model (the numerical representation of the conceptual

model). This aspect of model development is concerned with producing an accurate and computationally tractable solution that is robust in terms of its output and solution issues. It will also include the precise numerical structure of the code. From this a computational code can be realised into a simulation of the environmental problem. In order to do this data collection must be made of the appropriate inputs to the model, as well as observed events of the problem, in this case flood extent, depth and velocities, in order to calibrate and validate the model.

Each stage of this process will introduce uncertainty into the model and model results, either through simplification of the initial problem, or in data capturing methods. Uncertainty in modelling is comprised of two components, aletoray and epistemic, where aletoray is the uncertainty associated with the random nature of the process. Epistemic uncertainty relates to the inability to accurately capture the process and maybe improved upon by refining the data capture process. Both manifest themselves in flood inundation modelling, either through the inability to provide enough data capture points, the longevity of records relating to this and the data capture methods.

The issue of uncertainty is a critical contemporary area of research in inundation modelling (Hall et al 2005, 2009, Pappenberger et al 2009, 2007, 2006, Aronica et al 2002, and 2012, Bates et al 2004, Dottori and Todini et al 2013, Leedal et al 2010, Tusubaki et al 2013, Stephens et al 2012) and a considerable effort is now being made to understand the sources of uncertainty, the impact of it on model results and methods to reduce and communicate issues surrounding it. The sources of uncertainty are varied, from uncertain input parameters such as friction value, and determining appropriate grid resolution through to inflow hydrographs, and random errors associated with input topography data sets. The input uncertainties have been identified as a contributing factor to overall model uncertainty. An area that has not typically been assessed within an uncertainty context is the level of physical representation in the numerical code. Previous efforts have used direct comparison and benchmarking approaches to determine the impact of this choice (Hunter et al 2008, EA Benchmarking study 2010 and 2012), but the impact of this in a wider uncertainty context has not been considered before. This is in part related to how to differentiate between different codes that use different spatial girds and numerical methods, an issue raised by Hunter et al (2008). In these benchmarking studies the level of physical representation has been shown to be a significant contributor to variations in simulations. It has also shown that the use of simplified approaches, codes which exclude terms from the shallow water equations will produce different results from the codes that use all terms based on the hydraulic conditions of the test, in particular where transcritical flows

are fundamental to the modelling problem. Differences also occurred between the different categories of models, in particular between different full SWE models. This indicates that not only is the level of physical representation critical to model results, but also the precise method of solving the full SWE.

Uncertainty analysis is a critical area of research in flood inundation modelling. With has focused on more quantitative methods, and the use of sensitivity analysis has allowed the direct impact of an uncertain input to be determined. Hall et al (2005) has shown how the use of sensitivity analysis can help to further model development by diagnosing inputs that contribute significantly to model variations across the distributions of model inputs. The use of sensitivity analysis to progress model development and to form an essential part of the model code development process has been indicated in wider research (Saltelli et al 2000). The number of different approaches has helped further understanding of uncertainty in modelling (Pappenberger et al 2008, Hall et al 2009), and demonstrate the need for further use of sensitivity analysis in modelling problems.

A key aspect in model development is to use it in the context to which it has been created. Flood inundation models are used to assess the damage related to floods on both economic and human scales, through the modelled output. Inundation models are typically assessed in terms of comparison with observed flood data. Uncertainty is present in this process through a number of areas including the evaluation technique, typically undertaken with methods such as goodness of fit (F2) or time history comparison (Nash Sutcliffe), the data used to calibrate the model (such as satellite data) and issues of equifinality, where multiple models with different parameter sets perform equally well for a single model function. Issues with these approaches have been considered previously (Stephens et al 2012, Mason et al 2009, Aronica et al 2002, Beven et al 2006), but methods based on damage may provide further insight into modelling uncertainty. Furthermore, estimating damage from floods is a main area of interest in evaluating flood risk with these codes. Subjective calibration methods based on vulnerability have been pioneered by Pappenberger et al (2007). These approaches include vulnerability weighted methods, and regional analysis. Both methods use information, either based on decision makers' knowledge or underlying data to enhance the model calibration process towards areas at critical risk from flooding. The concept of damage as a means of evaluating models could further be refined by using depth damage curves to estimate the cost of the event in terms of direct damage. The uncertainty with these approaches has been considered by Apel et al (2009), but demonstrates that these approaches could be used to provide information about model performance, and in data permitting exceptional cases, be used to calibrate

models. The use of these approaches could provide the means for communicating model uncertainty, refining the calibration process and provide direct damage estimations.

1.1. Aims and Objectives

The aim of this research is to provide insight into how significant the level of physical representation is to model results, and to determine which inputs should be considered by modellers when undertaking a modelling exercise.

1. To determine the effect of level of physical representation (the numerical representation) on model output and results.
2. To compare this effect with the uncertainty associated with other inputs typically considered in flood inundation modelling.
3. To determine the impact of model evaluation techniques (objective functions), and how the significance of the model inputs changes with different evaluation technique.
4. Determining how financial consequence can be used as a model output, to refine the modelling process either through subjective means or cost based means, and to determine the impact of module choice on these evaluation techniques.
5. Using sensitivity analysis to quantify which inputs are most critical to model output.

To meet these objectives two approaches are used. First, a systematic approach is used which uses a Monte Carlo style approach to testing the parameters. Each parameter is divided into discrete intervals based on the range of uncertainty of the input. This is to reflect values and approaches which might be considered by a modeller when undertaking a modelling exercise. Each discrete interval of an input is compared to every input interval to provide a model result for each part of the parameter space. Each model will be analysed with an evaluation technique, or objective function – a quantification of the model results based on how the simulation has compared to an observed data set. Each technique is then applied to the results to provide a qualitative overview. A second approach, based on sensitivity analysis techniques is used to explore the findings of the first section. This will be based on a screening method, the Morris Method, and a variance based analysis which uses the BACCO GEM-SA software to undertake the analysis (Oakley and O Hagan 2002).

In order to remove the potential of model bias from results, which is the impact of using test cases which may favour one model code, three test cases with different hydraulic conditions are used; an urban test case based in Glasgow, a canal embankment failure based on a historic event in Coventry, and a river overtopping event that occurred in Mexborough in summer 2007. The events when considered individually and as a combined data set will provide insight into the impact of physical representation, the contribution of other inputs to modelling and the use of damage and vulnerability based measures to refine the modelling process.

Within this work a number of working terms are used. In reference to the input parameters of the uncertainty study, all factors including parameters, coefficients and modules (defined as being different solutions to the SWE in the modelling framework, to distinguish between the model code (LISFLOOD) and the different numerical models) are referred to collectively as factors. This is in keeping with the terms used by the Sensitivity Analysis community (Saltelli et al 2000). The level of physical representation refers to the general problem of the number of terms in the model code from the Shallow Water Equations (SWE). This is a relative term, explained in more detail in Chapter 2. The LISFLOOD code is divided into separate computational sections to solve the numerous approaches to floodplain flow. Each section contains different levels of physical representation which are referred to as modules. An overview of the literature is provided followed by the methodology used to undertake the research. Three test cases are used, Glasgow, Coventry and Mexborough, each of which contains a number of different analysis to investigating uncertainty in flood inundation models. This is followed by conclusions and points of further work.

2. LITERATURE REVIEW

This section provides an overview of the background to this research. It is comprised of two sections; the first looks at the flood inundation models, the governing equations, creation of the computer model and related issues and introduces the LISFLOOD-FP code, which is used in this research. The second part, starting at section 2.4, looks in detail at research into issues of uncertainty in modelling, the use of exposure methods, and the use of sensitivity analysis. Finally, a review of this section identifies what areas of research remain outstanding and how the current research fits into this knowledge gap.

2.1. Modelling Background

The creation of a flood inundation model is dependent on a number of options available to the modeller and code developer. These options occur in both the development of the computer code and in the choice of inputs to that model and each introduce uncertainty into model results. In the creation of the computer code, a hierarchy of options is formed from the initial choice of level of physical representation. From this further choices including discretization methods, boundary condition representation, and numerical solvers must also be made. The level of physical representation is a significant one that can led to variations in the model results (Hunter et al 2008). The reasons and for this are examined below, as well as descriptions of the options in creating a model code. The LISFLOOD code is then described in detail to outline part of the research and examine the implications of the choice of physical representation.

2.1.1. The Shallow Water Equations

The most common equations in computational flood models, and other environmental problems involving free surface flow, are the Shallow Water Equations (SWE). These are based on two fundamental principles; the conservation of mass and the conservation of momentum. These can be derived from the Navier-Stokes fluid dynamic equations. When certain assumptions are made in the SWE, the main assumption being that the vertical acceleration is negligible and that the pressure distribution is hydrostatic and acceleration is not affected in the vertical direction. Therefore, the water is incompressible, and vertical acceleration can be ignored (Toro 2001). A further assumption is that the rate of change in flow and height elevation along the length of the problem is less than this length, leading to the term shallow water equations (Chow, 1988). These

assumption essentially creates a two dimensional (in plan) form of the equations, that are time dependent and form a quasi linear hyperbolic system (Toro, 2001). The SWE in 2D form contain a conservation of mass term and two momentum equations in the x and y direction;

$$\frac{\partial \boldsymbol{U}}{\partial t} + \frac{\partial \boldsymbol{F}}{\partial x}(\boldsymbol{U}) + \frac{\partial \boldsymbol{G}}{\partial y}(\boldsymbol{U}) = \boldsymbol{S}(x, y, \boldsymbol{U}) \qquad (2.1)$$

Where

$$\boldsymbol{U} = (h, q_x, q_y)^T \qquad (2.2a)$$

$$\boldsymbol{F} = \left(q_x, \frac{q_x^2}{h} + \frac{gh^2}{2}, \frac{q_x q_y}{h}\right)^T \qquad (2.2b)$$

$$\boldsymbol{G} = \left(q_y, \frac{q_x q_y}{h}, \frac{q_y^2}{h} + \frac{gh^2}{2}\right)^T \qquad (2.2c)$$

$$\boldsymbol{S} = \left(\mathrm{S_o}, \mathrm{S_f}\right) \qquad (2.2d)$$

Where $\boldsymbol{F}, \boldsymbol{G}$ are the flux terms in the x and y directions, h is the depth of water, u and v are the velocity components in the x and y direction, $q_x = uh$, and $q_y = vh$, are the discharge per unit width in the x and y direction respectively (where u, v are the depth averaged velocities in x and y direction), g is gravitational acceleration and $\boldsymbol{S}$ represents the sources terms of bed slope, and free surface slope term, and losses due to friction, here defined in terms of friction loss due to Manning's coefficient (2.2e).

$$S_o = -\frac{\partial z}{\partial x} \qquad (2.2e)$$

$$S_o = -\frac{Q|Q|n^2}{A^2 R^{4/3}} \qquad (2.2f)$$

Where $, \partial z$ is the bed slope elevation, and A is the cross section area, R is the cross section hydraulic radius, defined in rectangular channels as A/P, where P is the cross section perimeter. These terms are defined in Figure 2.1;

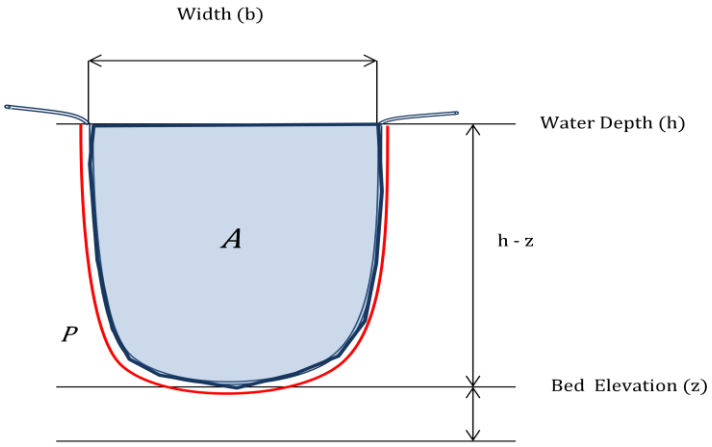

Figure 2.1: Schematic channel cross section . Cross section area (A) is represented by the blue area, cross section perimeter (P) by the red line

2.2. THE CREATION OF THE MODEL CODE

In developing a flood inundation model, the governing equations must be reformulated into a form from which the desired variables, in this case water depth and velocity, can be calculated. The most significant steps involve determining the level of physical representation, the numerical solution, temporal and spatial discretization, and the boundary conditions and formulations. Each step contains a number of approaches that have implications for the solutions provided by the code.

2.2.1 Classification of Models

Using the notation of Chow et al (1988) the momentum equation for the x direction consists of five force components (the same equation is used in the y direction, and the complete system requires the coupling of these equations), defined in conservative differential form below, in equation 2.3.

$$\frac{1}{A}\frac{\partial}{\partial x}\left(\frac{Q^2}{A}\right) + \frac{1}{A}\frac{\partial Q}{\partial t} + g\frac{\partial y}{\partial x} - g(S_o - S_f) = 0$$

Convective Acceleration	Local acceleration term	Pressure Force term	Gravity Source	Friction Source

Kinematic Wave model ←

Diffusion Wave model ←

Acceleration Wave model ←

Dynamic Wave model ←

The local acceleration describes the change in momentum relative to the change of velocity in a volume of water over time; the convective acceleration term describes the change of momentum relative to the change of velocity in the surrounding water. Collectively these two terms represent the effect of transferred momentum on the water and account for unbalanced forces terms acting within and against a control volume. The pressure force term describes the change in momentum as a result of variation in water depth along the channel. The gravity and friction force terms are proportional to the bed slope and the friction slope, and describe the rate of change of momentum as a result of the bed slope and energy losses due to friction.

By including or excluding terms from Equation 2.3. various forms of the equation can be obtained which can be used to model the physics to a greater or lesser degree of completeness. Using the above notation again to define these physical models, the full momentum equation, comprising all five terms, is alternatively known as the dynamic wave model and full SWE model (as in EA 2010 Benchmarking report). The first level of reduced physical representation models is created by removing the convective acceleration term from the mathematical model. This creates the acceleration formulation of LISFLOOD (Bates et al 2010), first used in an inundation model by Aronica et al (1998). As can be seen from Equation 2.3, exclusion of the inertial terms creates the diffusion wave model, used in the original LISFLOOD formulation (Bates and De Roo, 2000). The simplest model is the kinematic wave model, where the force of motion is caused by changes in the bed slope, which excludes the mass and velocity of the water mass from the momentum equation. The choice of model is based on modeller experience and from determining the dominant hydraulic features to be modelled. Each model has some justification for being used in a computational code. The kinematic approach, whilst ignoring many of the terms, captures the main properties of a large scale flood wave (Cunge 1976). A flood wave will be dominated by changes in the channel and floodplain slope, and at larger scales additional hydraulic effects which cannot be modelled by this approach, maybe negligible. This may be appropriate in modelling large scale events, at the continental scale, such as the Amazon basin. The diffusive wave approach, improves on this model by including the pressure term which allows for backwater effects to be modelled. The impact of backwater effects, where the impact of changes in water level will propagate back upstream, are significant in subcritical flows which are typical flow conditions in rivers and floods (Chadwick, 2004). The inclusion of one of inertia terms has also been suggested (Bates et al., 2010), Aronica 1998). In these

models, only the convective term is excluded, which is considered unimportant in floodplain flow (Hunter et al., 2007), and allows computational codes to use larger time steps. This model can then be used in for similar studies as the diffusion wave model, but at a higher resolution and for more detailed studies.

These methods represent reduced physics approaches, or simplified computational models. A review of the use of simplified approaches was undertaken by Hunter et al (2007), which determined that the use of simplified approaches was justified when the length scale of the problem was significantly greater than the water depth. At larger spatial scales the use of fully dynamic model codes became less critical as local acceleration terms reduce in significance in coarse resolution models. The authors also concluded that these simplified codes have a tendency to be sensitive to the topography of the model (and therefore the cell size of the model) and time integration and discretization methods. Application of these methods to flood inundation problems needs to be evaluated by the requirements of the model results. To date this has still not been quantified for 2D problems.

The full dynamic wave model however includes all these impacts and the inclusion of the inertial terms will account for local variations in topography and over time, through the local and convective acceleration terms respectively. These terms then capture super critical flow and the impact of transcritical flows, as well as including shocks and discontinuities. Consequently in urban environments, where supercritical, low depth high velocity flows occur in complex pathways (Froude numbers greater than 1), this approach has been proved in test cases to capture the hydraulic conditions of the observed event (Hunter et al., 2008, Schubert et al 2008). However, significant computational costs are associated with the implementation of full SWE models (Bates and De Roo 2000), as well as an increased number of additional code development options, including the choice of numerical solver.

2.2.2. Discretization of the Mathematical Model

In order to produce a numerical solution, the mathematical model, the product of determining the level of physical representation, is formulated to solve for the desired unknown variables. Direct numerical methods are commonly used due to the ease of implementation and the speed of the solution when applied to simulation of real events. In this approach, a computational grid is formed where the spatial model domain is divided into sections or cells which represent control volumes within which the unknown flow variables are calculated. The governing equations are re formulated into a discrete formulation based on variables in the

cells or sections of a grid, where the approximate solution to the governing equations can be determined at points in the computational grid. These approaches produce approximate answers that are highly dependent on the computational grid for determining the value of the unknown variables (Toro 2001). Within these methods a number of additional modelling discretization factors must be considered, such as spatial discretization, and are described in greater detail in the following sections.

Three methods are common in flood inundation modelling, finite volume (FV), finite difference (FD) and finite element (FE). The most straight forward method is the finite difference method that is based on the differential form of the governing equations, using peicewise approximations of the variable to be solved at the cells. The original LISFLOOD formulation used this approach (Bates and De Roo 2000). FV methods are based on using integral formulation of governing equations, rather than differential form used in finite difference methods. The variables in integral form are updated as a cell average in each grid volume, which allows more complex numerical solutions to be made. The Finite Element method is broadly based on dividing the governing set of equations to a series of sub sets (the grid cells) where a local approximation to the main equations is calculated. The FE approach uses trial and error methods to estimate the local coefficients. Consequently, the method has a better approximation to the global equations although numerical shocks are difficult to capture with this approach. By using a linear interpolation method, FV provides a more suitable framework for incorporating shocks into the solution, which is a critical component of fluid dynamic modelling (Horritt et al., 2007).

The significance of this choice in affecting inundation results was explored in a paper by Horritt et al (2007) by comparing a Finite Volume Model (SFV, (Horritt and Bates, 2002), and a Finite Element model, TELEMAC 2D. Using an event from the River Severn in November 2000 as a test case, and comparing the model with validation data, in this case 4 aerial images of the flood, a series of test were undertaken with the two models based on an identical parameter space. The paper shows that in calibrating to an image associated with the low flow situations both models were sensitive to the friction parameters, compared to the higher flow images. Both models produce similar model evaluation functions in the calibration of the models, but variations were noted in the predictive ability of the models, where the SFV model produced a higher, and more sensitive, model evaluation function. The TELEMAC code was sensitive to the mesh size, noting that in finer resolution the model was unlikely to over predict flows in meanders. In the

predictive test, the greater sensitivity of the SFV code to channel friction and TELEMAC to floodplain friction values was noted. The findings demonstrate that the importance of the numerical solution is difficult to determine, and that each approach has a different level of sensitivity to model inputs. The varied response of each model to the input space indicates that the significance of uncertainty in inputs will be affected by the numerical solution. A further implication is that whilst the discretization approach produces variations in model results, method of evaluating the results will determine the level to which that impact is significant, for example comparison of simulated extent.

2.2.2.1. Further Code Development Considerations

The model is discretized spatially and in time. In Temporal discretization, two main methods exist, implicit and explicit. The implicit approach is based on a stencil approach, where the unknown variables for a cell are updated using information from adjacent cells and adjacent time steps. An advantage of this approach is the greater stability afforded by the scheme. With variables being based on multiple finite difference points, the time step can be increased, relative to the computational cheaper explicit approach. Examples of this approach include DIVAST and ISIS 2D ADI, but as computational power has increased this method has decreased in use, partly as explicit methods are more computationally feasible, and the comprise required in the numerical formulation of the implicit approach. Further reasons include the increased complication required from the code and additional steps required to solve initial and boundary conditions. The scheme is stable at time steps larger than explicit methods can allow, but is less accurate than explicit approaches, and work best at similar timesteps as explicit approaches, but with increased computational cost. The explicit approach updates values at a cell based on values from the known values of the previous time steps. The time step must be sufficiently small, however, to ensure that wave celerity (or speed of information) is captured within each cell, and is necessary, but not a guarantee of model stability. The explicit approach is common and is used across models with different levels of physical representation (LISFLOOD, JFLOW). In cases where these two approaches have been compared, results indicate that this choice is less significant than the choice of physical representation (Hunter et al 2008).

The Spatial discretization involves dividing the model domain into computational cells, which represent control volumes with which to calculate the unknown variables. The two main approaches are to represent the spatial domain as a regular grid, where computational cells are represented with regular sized volumes,

and an irregular grid where the cells are represented as triangles. This approach is further divided into structured and unstructured approaches, where cells are not based on arrays, but require connectivity calculations to determine adjacent cells. Grid cell shape can be square, such as in the LISFLOOD-FP (Bates and De Roo, 2000) code or triangle such as BreZo (Begnudelli and Sanders, 2006), the variation between these two approaches being that whilst a square affords a distinct advantage in terms of reduced complexity of the code, the triangle grid allows a complex environment to be modelled with fewer cells, as the grid can be refined in a way that allows an accurate representation with a limited increase in cell count. Further refinement with the triangle method is created by using a unstructured gridded method, where computational cell size can be reduced in complex flow areas, in particular around buildings (Schubert et al., 2008).The choice of spatial discretization method represents a key aspect of model development, as the advantage afforded by a simpler code can be overcome by the use of more effective grid meshes to represent complex urban areas (Fewtrell et al., 2008).

Horritt and Bates, (2001) explored the impact of the two approaches, using the TELEMAC model, which uses triangular elements and LISFLOOD model which uses regular square gridded data. Comparing the two approaches for a test case in the upper reaches of the River Thames, with Synthetic Apeature Radar (SAR) flood outline data to calibrate the results, the paper found that both methods replicated the data equally well, noting that due to the resolution of the data it would be difficult to suggest if there was a difference between the two approaches. The use of using two different physics models, in TELEMAC and LISFLOOD would also make it difficult to compare the two discretization approaches directly.

Further choices involved in the construction of the flood inundation model include the methodology for dealing with the wet/dry boundary. Flow in this region is very shallow, and becomes critically close to failing the depth integral assumptions of the governing equations causing unrealistically high velocity values, or negative depth values in the solution. Approaches include the use of water depth thresholds, where the mass and momentum fluxes are set to zero below threshold depth values (Bates and De Roo 2000, Sleigh et al., 1998), adding a small quantity of water in dry cells adjacent to wet cells, which ensure model stability but include a small mass error in the model, and dynamic adaptive meshes at the wet/dry boundary, that will evolve over the course of a model realisation (Liang and Borthwick, 2008) ensuring the computational effort is focused at the more critical modelling points. The wide variety of approaches to this problem are determined in part by the choice of physical representation in the numerical model, with simplified models not

requiring the same precision to the boundary as a full SWE model. The full SWE which will include some semblance of shock capturing or absorbing methodology, will potential fail at this point. The simplified model equations do not include shocks, and consequently are more robust to these conditions. A number of studies have explored this option. The EA benchmarking study (2010) contained a test which examined wet dry boundary solutions across a variety of models. This test found that fully dynamic models could produce similar flood wave propagation to within a few percentage points of each other, but simplified models were more varied in their outputs. Horritt (2002) used three different approaches to solving the boundary with a finite element model. Each approach reproduced different aspects of a beach wetting and drying problem better than the other formulations, indicating that a compromise is required in determining the solution to this boundary.

A common approach in flood inundation modelling is to represent the channel as a separate domain from the floodplain. This approach has a number of advantages, as it allows existing river channel survey data to be utilised, whilst reducing the computational cost of the simulation. A number of approaches have been undertaken to solve the flow in this location, from masking river cells in the 2D domain, allowing overtopping water to enter the 2D domain directly, conserving momentum as well as mass, as in LISFLOOD (Horritt 2001, Bates and De Roo 2000), to interpolated water levels based at the model boundary of the channel floodplain where momentum and mass can be conserved, as in the ISIS code. The most comprehensive comparison of linking methods was undertaken in the EA 2D model benchmarking report (2010), explained in further detail below. All models capable of this approach were used in a valley filling exercise, and produced results that varied in modelled water velocities and depths. The valley filling exercise allowed for a number of approaches to dealing with the linking method and representation of the channel, banks and restrictive flow features, such as bridge crossing and culverts. Each approach produced acceptable with no method appearing to be superior to another.

2.2.3 Comparison of Flood Inundation Model Codes

The wide variety of flood inundation models has led to a number of benchmarking studies aimed at highlighting the relative performance of each method. The most comprehensive is the 2010 EA Benchmarking study of 2D hydraulic codes (EA 2010). Eight test cases were devised, with different hydraulic properties to test specific facets of the models. A number of codes were evaluated in the study ranging from simplified physics models Rapid Flood Spreading Model (RSFM)

through to full dynamic wave models with higher order solutions and shock capturing methods, such as TUFLOW and ISIS. The test cases provided a wide range of methods for examining the code's performance, including a number of idealised experimental cases (i.e. not based on real scenarios or data) involving filling of depressions, speed of flooding over an extended floodplain, and conservation of momentum over an obstacle. Other test cases were based on more realistic applications, such as valley filling and dam failure exercises. These test cases were for the most part based on idealised situations and not real flood events. The main conclusions of the review were based on evaluating the variations between and within full SWE models and simplified models. Full SWE models were determined to be appropriate in most flood risk evaluation techniques, providing consistently realistic solutions, which compared well with other SWE models and highlighted shortcomings with the simplified models. However, the report noted that the use of simplified models has advantages in other areas, in particular in estimating final water level values, which were comparable to the full SWE models. The main area in which the simplified models produced poorer results than the dynamic models was in the velocities and conservation of momentum. For example, in the conservation of momentum test, simplified models could not capture the overtopping effect, but in valley filling exercises, the methods proved to be just as effective as full SWE models. This difference was significant in the dam failing exercise, where accurate reproduction of the state of flow just after failure is important in capturing the risk associated with this type of event. Furthermore the paper highlighted that the use of a simplified approach did not translate into a reduction in computational cost, a key factor in the use and development of these (Bates and De Roo, 2000). A significant saving is offered by approaches such as the RFSM, which provides a final level value for the water across the model domain, but at the cost of time varying information. This limits the applicability of these models to more detailed problems. The approach also requires a number of pre processing steps to determine the flow paths *a prior* to ensure that they are explicitly represented in the model.

Other test cases of particular interest were 1D/2D linking methods and valley spreading methods, which provided significant variations between all code types. Whilst all models capable of this approach produced acceptable results, the variations between them were not limited to the mathematical approach adopted beforehand. A significant point was the variation between full SWE models with a shock capturing method (such as ISIS 2D TVD) compared to full SWE models without (such as MIKE FLOOD). In test cases where transcritical flow was present

and accounted for a significant component of the hydraulics of the location, such as the dam break failure, the requirements to accurately depict the shock in the flood wave led to a variation in the depth and velocity.

A further analysis of comparison of model approaches was a benchmarking paper undertaken by Hunter et al (2008). This paper focused not only on a comparison of different model codes, it also looked at the impact of parameters which are typically considered in both a modelling exercise and in wider research. The benchmark modelling was performed for the Greenfield area of Glasgow. The inflow boundary conditions were a point source, which represented an overtopping culvert. The paper highlighted a key variation between simplified and full SWE models, in the extents reproduced by the two model types. The area is a residential urban area which is comprised of several roads, which act to constrict the flow, in a west-east direction, and buildings. The hydrograph input point is an overtopped culvert at the top right of the model domain, that creates runoff along the main street, and along aback street where the water pools at the lowest elevation point of the model domain. A hydrograph input was devised, as well as a number of distributed friction value surfaces, based on underlying topography. A further factor in the test case considered the error associated with the LiDAR derived DEM. The benchmarking study used a wide variety of codes including DIVAST, TRENT, LISFLOOD, JFLOW, and TUFLOW, which represent both simplified approaches, and full SWE models.. One of the main variations was between simplified approach models and the dynamic models, in reproducing a flood extent. The full dynamic models were able to produce a wider flood extent as the result of being able to conserve momentum down the main flow path slope and overcoming some of the obstacles in the domain. These obstacles were unable to be overtopped with the simpler model. The depth of this water in the wider flood extent was only 5cm, and similar variations of depth at control points throughout the domain was noted, suggesting that the change in the vulnerability to flood risks was relatively low. Indeed in the more critical areas of the model, where water depths and velocities are high, such as down the main flow path, all models produce similar depths. These results indicate the importance of the inertia terms in urban areas, where complex topographies and flow paths dominate, but that the overall effect on vulnerability measures may not be as significant. It is noted however, that the large slopes present in the domain may impact the model results, which may not be typical to other modelling exercises or case studies.

From this analysis, the need to determine the impact is essential in reducing computational uncertainty associated with model code, and in particular the level of

physical representation, which has potentially the most significant impact on model results.

2.3. LISFLOOD-FP

In order to evaluate the impact of the physics model within a hydraulic computational code there is a requirement to chose one code which isolates these factors in the model, whist retaining other common features, such as computational boundaries, and DTM treatment. Therefore a code was required which provided a framework which within it had a number of different codes. The LISFLOOD code, developed at the University of Bristol was chosen (Bates and De Roo 2000). The LISFLOOD code is a 1D/2D code that represents the channel flow as a nested 1D model, and the floodplain as a regular grid 2D domain. The floodplain is represented as a series of interlinked regular squares, where the elevation is estimated as a piecewise value at the cell-centre. This was designed to take advantage of remote sensed digital-elevation data and present a more feasible methodology for computing flood inundation. In this section the various modules within LISFLOOD-FP responsible for solving the floodplain flow are described.

2.3.1 LISFLOOD-ATS

The initial formulation for LISFLOOD was first developed in 2000 (Bates and De Roo 2000) and was based on an analytical method for the diffusion wave. The concept behind this model was the use of a simplified physics model in order to reduce the computational cost compared to a full shallow water equations approach. The diffusion wave approach was shown to be a reasonable approach to modelling flood inundation at the reach scale (Hunter et al 2007). The use of simplified physics models consequently became more popular as increases in data allowed more extensive use of 2D approaches. The LISFLOOD formulation is however strictly an analytical approach, and does not utilise a numerical approach which has been adapted into later diffusion wave based codes, such as the MAST code (Arico et al., 2011).

The equation system for the model is

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} + q = 0 \qquad (2.4)$$

For the continuity equation, where Q is discharge, x [m] is cell size, A[m²] is area, t[s] is time and q[m²/s] is lateral inflow. The momentum equation is reduced to

$$S_o = S_f \qquad (2.5)$$

Where S_o is the bed slope term and S_f is the friction slope term, the concept of this system is that the friction and bed slope balance, and that further acceleration terms can be neglected in the numerical code. The equation system is then implemented in the floodplain as an explicit finite difference scheme. Channel flow is solved using a 1D Newton Raphson solver. The 2D flood plain cell flow is solved using flow values from neighbouring cells, using a finite difference discretization of the time derivative term;

$$\frac{\partial h^{i,j}}{\partial t} = \frac{Q_x^{i-1,j} - Q_x^{i,j} + Q_y^{i,j-1} - Q_y^{i,j}}{\Delta x \Delta y} \qquad (2.6)$$

Where $\partial h^{i,j}$ represents the water free surface height in a cell, at a point in time. This equation relates the variable of depth to the intercell fluxes $Q_x^{i,j}$, to complete the continuity equation of 2.4. Intercell fluxes are calculated using the Manning's uniform flow equation.

$$Q_x^{i,j} = \frac{h_{flow}^{5/3}}{n}\left(\frac{h^{i-1,j} - h^{i,j}}{\Delta x}\right)^{1/2} \Delta y \qquad (2.7)$$

Where Q = discharge (m/s³), *n* is the Manning's friction value, *i,j* are cell index values and Δ*x* Δ*y* represent cell size. The amount of water available between cells, h_{flow} is defined as being the difference between the highest elevation of the two adjacent cells and the difference in water surface elevations.

The initial code suffered from numerical instabilities and the creation of a 'checkerboard' failure, where water would move rapidly between adjacent cells over timesteps. This was due to the lack of acceleration terms in the numerical model which caused the model to overestimate the interval flux (Bates et al., 2010). In order to counter this, a number of solutions were tried. Flux limiters were used to ensure that the discharge between cells did not exceed a defined value to ensure a reasonable flux value (Hunter et al., 2005). This, however, created model results that were insensitive to friction parameters but dependent on cellsize and timestep, and consequently produced results with limited physical relevance to flood inundation. The code and model were analysed to determine a suitable time step relevant to diffusion wave models (Hunter et al., 2005). The new formulation for estimating the time step and ensuring numerical stability is based on von Neumann analysis of diffusion based systems, which defines the stable time step as a quadratic of the cell size. As a consequence, cell size has a significant impact on model run time. The time step is defined as;

$$\Delta t = \frac{\Delta x}{4} \min\left(\frac{2n}{h_{flow}^{5/3}}\left|\frac{\partial h}{\partial x}\right|^{1/2}, \frac{2n}{h_{flow}^{5/3}}\left|\frac{\partial h}{\partial y}\right|^{1/2}\right) \qquad (2.8)$$

The optimal timestep with this method ensured model stability, and was implemented in a dynamic way ensuring that the optimal size timestep was

determined for each iteration over the physical model domain. It is noted, however that the model still displays a general insensitivity to floodplain friction parameters, compared to full SWE models. Whilst this method ensured numerical stability and retained an efficient solution of flood model problems at large spatial scales(~20m), it becomes extremely restrictive in urban areas, where complex topography requires a small cell size (~1m) to represent key urban features. As data sets and computational power advanced, the ability to model complex urban environments at high spatial resolution became more applicable, and reduced the ability to use the ATS formulation as a multipurpose flood model.

2.3.2 LISFLOOD-ACC

In order to improve the computational costs of the diffusion code, a new form of the equations to be used in flood modelling was formulated, which still continued the concept of reduced physical representation, but had an approximation to inertia terms to allow for a larger time step. The ACC code, which stands for Acceleration, was formulated on the removal of the advection acceleration term from the governing equation and the inclusion of the local acceleration term from the Saint Venant equations. The derivation of the numerical model is based on Bates et al 2010, and is formulated here in a 1D format.

Beginning with equation 2.3, the convective term is dropped. Dividing through by hydraulic radius *R* leaves the following inertial equation system

$$\frac{\partial h}{\partial t} + \frac{\partial q}{\partial x} = 0 \qquad (2.9)$$

$$\frac{\partial q}{\partial t} + gh\frac{\partial (h+z)}{\partial x} + \frac{gn^2|q_x|q_x}{h^{7/3}} = 0 \qquad (2.10)$$

Equation (2.9) represents the simplified 1D continuity equation, whilst equation (2.10) represents the momentum equation. The system is implemented into a 2D model through a simple 2D finite difference method, using a similar storage cell methodology as was implemented with the initial LISFLOOD model, where intercell momentum flux is calculated with equation (2.11)

$$q_{i-1/2}^{t+1} = \frac{q_{i-1/2}^{t} - gh_f\frac{\Delta t}{\Delta x}(y_i^t - y_{i-1}^t)}{1+g\Delta t n^2\left|q_{i-1/2}^{n}\right|/h_f^{7/3}} \qquad (2.11)$$

In order to ensure numerical stability, a semi implicit solution was devised, where the momentum flux was updated with values from the previous time step and current time step for the upstream cell. This formulation is first order in space. The acceleration formulation has therefore the advantage of using the CFL condition to determine the timestep.

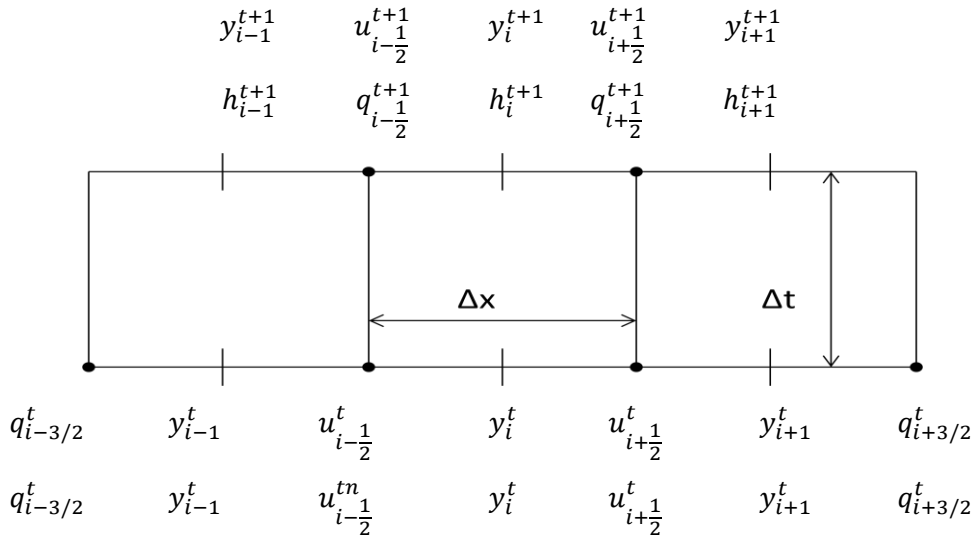

Figure 2.2: The mathematical structure of the Acceleration model. The variable to update (y_i^{n+1}) is cell centred with variables estimated from intercell fluxes. The figure demonstrates the concept of the semi-implicit approach (from de Almeida et al 2012)

The initial formulation produced numerical instabilities at low friction values (Bates et al 2010). Additional work to the formulation was provided by (de Almeida et al., 2012) which looked at the numerical framework of the model. This work determined that the instabilities, which arise from the non linearity of the mathematical model, are unable to be diffused at low friction values due to the semi implicit approach adopted to ensure numerical stability. These instabilities are naturally diffused by the model at higher friction values, but additional terms and approaches would be required to ensure that instabilities are dealt with in the numerical model. De Almeida et al introduced two methods, a *q upwind* method that works in a similar fashion to a finite volume upwind method, and a *q centred* method, which used a similar approach to the original semi implicit approach, but took it further to include information from the neighbouring cells. The explicit formulation of the two methods appears as below;

$$q_{i-\frac{1}{2}}^{t+1} = \left[\theta q_{i-\frac{1}{2}}^{t} + (1-\theta) q_{i-\frac{3}{2}}^{t}\right] - g h_f \Delta t \left[\frac{(y_i^t - y_{i-1}^t)}{\Delta x} + \frac{n^2 \left|q_{i-1/2}^t\right| q_{i-1/2}^{t+1}}{h_f^{10/3}}\right] \qquad (2.12)$$

$$q_{i-\frac{1}{2}}^{t+1} = \left[\theta q_{i-\frac{1}{2}}^{t} + \frac{(1-\theta)}{2}\left(q_{i-3/2}^{t} + q_{i+1/2}^{t}\right)\right] - g h_f \Delta t \left[\frac{(y_i^t - y_{i-1}^t)}{\Delta x} + \frac{n^2 \left|q_{i-1/2}^t\right| q_{i-1/2}^{t+1}}{h_f^{10/3}}\right] \qquad (2.13)$$

Where equation (2.12) is the discretized form of the momentum equation for *q upwind* and where θ represents a weighted linear relationship value with the intention of creating numerical diffusion to increase stability in areas of sharp discontinuities. Equation (2.13) is the discretized form of the momentum equation for the *q centred* method, where value of q are updated with a 3 point weighted

relationship with data from upwind and centre values at different time steps. This is similar to the Lax Wendroff method, and if θ is set to zero, this equation becomes the Lax method. A comparison of the methods is provided in de Almeida et al (2012), for a range of test cases, and notes that whilst both methods provide a stable and consistent solution to the ACC formulation compared to the original approach, there is an over dependency on the value of theta in the upwind method. With the *q centred* approach however, the ability to adjust the amount of diffusion improves the applicability of the solution and improves its performance, relative to the *q upwind* solution in low friction solutions. In spite of the additional terms and extension to the original ACC methodology, the resultant model still provides an efficient compromise between full physical representation and computational efficiency. Comparisons between the ATS model and ACC in a high spatial resolution urban area model showed improved runtime but with similar depth output, and a small variation in the water depth velocity (Fewtrell et al 2011). In this study a high resolution DEM was created for a pluvial flood event model, and demonstrated both the consistency between the formulations and the computational improvement of the ACC method. The ACC formulation therefore has an advantage not only of reduced run time compared to both the diffusion wave approximation and full shallow water models, which allows for modelling of wider spatial areas, particularly in urban areas.

2.3.3 Godunov Method

Leveque (2002) provides an overview of the Godunov method, defining it in a three step process, which describes the basic computations required but allows the detail of the solution to be increased at each step, depending on the governing equations used. The basic algorithm involves reconstructing the function $\tilde{q}^n(x, t_n)$ based on cell averages, evolving of the governing equation to determine the function at the next time step, and averaging the function over the each grid cell to obtain a new average. The Godunov method is based on FV methods, where the spatial domain is discretized into individual volumes, or grid cells, where an average state for a conserved variable component is maintained as an initial value problem in a cell, which results in piece wise distribution of the data. The conserved variable is updated with the following conservation formula

$$U_i^{t+1} = U_i^t - \frac{\Delta t}{\Delta x}\left[F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}\right] \qquad (2.14)$$

Where $F_{i+\frac{1}{2}}$ is the intercell flux value for the boundary between the cells at i and at $i+\frac{1}{2}$. This presents a Riemann problem, where two constant states (the conserved variable in the adjacent cells) are separated by a discontinuity. The application of this method with the LISFLOOD module is based on utilising a regular square grid,

with a upwind source term to deal with topography. The method uses a number of solutions for the intercell flux.

2.3.3.1. The LISFLOOD-Roe model

This example of the Roe solver (Roe, 1981) is based on a conservative non linear system of equations, which have similar properties to the SWE, but with fewer terms to allow for easier explanation of the solver. This form of the governing equations is based on the 1D Toro (2001) example. The 2D form used in LISFLOOD-Roe is based on the same formulation, but with an additional term in the discrete form of the equation to account for multiple dimensions.

The Roe solver first approximates the non linear system,

$$U_t + F(U)_x \equiv U_t + AU_x = 0 \qquad (2.15)$$

Where *U* is the conserved variable, $F(U)_x$ is the flux related to the variable and *A* is the Jacobian matrix, with the linear system

$$U_t + \tilde{A}U_x = 0 \qquad (2.16)$$

Where, $\tilde{A}$ is the Roe constant coefficient matrix, which approximates the Jacobian matrix *A* in the previous equation. This matrix is based on values U_L and U_R, products of the Riemann problem developed at the intercell flux. Average values must then be determined for the variables based on this information. The discrete upwind form of this system is;

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{\Delta x}\left[F_{i+1/2}^* - F_{i-1/2}^*\right] + \Delta t H_i^n \qquad (2.17)$$

Where $F_{i+/-1/2}^*$ represents the intercell flux for the Roe solver;

$$F_{i+1/2}^* = \frac{1}{2}(F_i^n + F_{i+1}^n) - \frac{1}{2}\sum_{j=1}^{3} \tilde{\alpha}_j \left|\widetilde{\lambda}_j\right| \tilde{e}_j \qquad (2.18)$$

Where $\tilde{e}_j$ represent the eignvalues of the approximate Jacobian matrix, λ_j represent the eignvectors, and $\tilde{\alpha}_j$ represent intercell variables. These components represent the linear variables used in the Roe solver.

The time stepping solution is based on the Courant-Freidrich-Lewy CFL condition;

$$c = \frac{\lambda \Delta t}{\Delta x} = \frac{\lambda}{\Delta x / \Delta t} \qquad (2.19)$$

Where c is a dimensionless number that needs to be less than 1 for stability, and λ is equal to $\sqrt{gh}$, the celerity of the wave. This allows the timestep to be determined by the equation;

$$\Delta t_{max} = \alpha \frac{\Delta x}{\sqrt{g h_t}} \qquad (2.20)$$

This methodology was originally the TRENT flood inundation programme (Villanueva and Wright, 2006), and was compared with the storage cell method

used in LISFLOOD. This TRENT model had been used in an extensive benchmarking study (Hunter et al 2008). The LISFLOOD module has been further used in comparison work with the other LISFLOOD modules, comparing results from a range of EA Benchmarking test cases (Neal et al., 2011). The benchmarking studies have shown that not only is the Roe solver comparable in terms of model output to other full SWE based models, it has also produced model results that vary from both the diffusion wave and ACC LISFLOOD approaches.

2.3.3.2 LISFLOOD-Rusanov

The Roe solver forms part of this module, which as described above provides a robust, linearization method, but other solver provide simpler solutions, with the disadvantage of a coarser numerical solution to the intercell flux. The main module that will be used in this research is the Rusanov module, which uses an upwind method with

$$\mathbf{F}_{i+1/2}^{\mathrm{Rus}} = \frac{1}{2}(\mathrm{F}_i^n + F_{i+1}^n) - \frac{1}{2}\mathrm{S}^+(U_{i+1}^n - U_i^n) \quad (2.21)$$

Representing the intercell flux. In this S^+ represents the wave speed and it is the estimation of this value that provides a significant reducing in computational costs compared to the Roe solver. For the Rusanov flux the simplest method, according to Toro (2001), is to evaluate the eigenvalues to determine the max wave speed at the interface. As the 2D Shallow Water Equations are a hyperbolic system, the initial conditions, and therefore the eigenvalues must be known, in order to converge to a solution, and to ensure an appropriate time step is determined. A simpler method is to determine maximum wave speed is determined from the time step

$$S^+ = C_{cfl}\frac{\Delta x}{\Delta t} \quad (2.22)$$

According to Toro (2001), this method provides a significant reduction in computational time as S^+ is part of the eignvalue system that must be known beforehand to know if the mathematical structure is hyperbolic in nature. In order to ensure numerical stability the timestep Δt must be related to this value

$$\Delta t = \frac{C_{cfl}\,\Delta x}{S_{max}^n} \quad (2.23)$$

For each time step there will be a unique value of S_{max}^n , which ensures upwind local information for each calculation of the flux. If the CFL condition is 1 then this reduces to the Lax Friedrich method. This method has significantly reduced costs compared to methods that determine eignvalues and appropriate estimations of correct wave speeds that are found with other higher order approximations.

Another notation utilises the present variables

$$S^+ = \max\left(|U_L| + c_L, |U_R| + c_R\right) \quad (2.24)$$

Where ;

$$c_L = \sqrt{gh} \quad (2.25)$$

Is the wave celerity for the left and right waves. This wave speed choice is more robust than the above notation, and includes estimations taken from relevant grid cells. This method can be simplified by removing U left and right values and simply taking the left face values. This value is derived locally and ensures smearing is not excessive in certain areas, particularly at wet/dry fronts.

The Rusanov flux has been used in other flood inundation models. Simões (Simoes, 2011) introduced a finite volume model, STORM, which uses a unique switching technique to solve the flux equation. Using a finite volume framework and an irregular grid formation the inviscid flux calculation, formed by the formation of the Riemann problem at intercell boundaries, use the Rusanov flux in areas of low variation, and an upwinding method developed by Alcrudo and Garcia Navarro(1993), which uses a combination of Harten and Roe solutions to solving the one dimensional Riemann problem. The solution switches to the Rusanov flux when the variation between cell values falls below 0.01%. The advantage of this splitting method is a reduced computational effort, as in areas of low water surface gradient, without the risk that excessive diffusion of the numerical will adversely affect results. Conclusions from a paper presenting this approach (Simoes 2011), indicates that, the use of the Rusanov solver in conjunction with a higher order Riemann solver, provides a more computationally efficient model whilst still providing consistent solutions. It is noted that the choice of switching value is critical in maintaining an accurate solution, without increasing runtime. The Rusanov flux represents a simplistic approach to solving Riemann problem whilst retaining the higher level of physics of the full SWE model. The use of such approaches may allow a distinct reduction in computational costs without a corresponding reduction in modelling complex hydraulic conditions. Determining the relative impact of this choice could provide a distinct insight into potential development of models, as well as furthering information on the appropriate numerical structure and complexity of the computational code.

2.3.4 LISFLOOD Comparison Work

The LISFLOOD framework provides an ideal and unique opportunity to test both physical representation and numerical solutions to the importance of model results. Individually the codes have been used extensively in research, but Neal et al (2011)

explored the impact of the three levels of code across a range of test cases used in the Environment Agency 2D benchmarking study. Across the range of test cases, it could be determined that in gradually varying flows, the diffusive wave formulation produced similar results to the dynamic wave approach. The variation between approaches was much more apparent in test cases with complex topography, rapid flow changes, and slope transition points and in transitional flows. For coarser resolution models, such as the 50m model used in part of a valley filling test case, the modules produced similar results, confirming previous results (Fewtrell et al., 2008) and indicating the reduction in importance of inertial terms at high cell sizes and the interaction of parameters. For a test case involving flow over a bump, which directly tests model momentum conservation principles, the ATS model was unable to clear the obstacle, unlike the ACC and Roe formulations. For the higher resolution model in the valley filling exercise (10m resolution) the variation between the SWE model and the ACC at critical points in the domain was as great as 1.6m. Other variations included the Roe solver producing later arrival times for gradually varying flood water, and slower increases in depths. The implication in a vulnerability context is not clear from these results, but the overall differences would appear to suggest that whilst simpler models can replicate similar results, the application of them is case specific and should be carefully considered. Another broader point is the approach required in benchmarking studies, where arbitrary choices, such as the frequency of recording depths and the control point locations can impact significantly on model results, as much as 9%. Determining this factor is crucial to producing robust results.

Each paper highlights not only the applicability of each module to inundation problems, but of the advantages of using an intergrated modelling system within which modules of varying levels of physical representation. Analysis can then be focused on the immediate variations between model types and can be directly attributed to the physical representation within it.

2.4. UNCERTAINTY ANALYSIS

Computational models contain a level of uncertainty associated with the process of creating a conceptual model and writing a computer code. In flood inundation modelling a number of other factors contribute to the overall model uncertainty. The next section discusses in detail the sources of this uncertainty and the methods used to identify, quantify and reduce it in model results. Contemporary research of uncertainty analysis in flood inundation modelling is comprised of two main subjects, understanding and quantifying uncertainty, and reducing or explaining uncertainty. The main aim is to determine the sources, account for them and improve the modelling process, or include them in model results. Broadly, uncertainty in computational modelling is described through two basic categories. The first category describes the inherent inability to fully capture a natural system and the inability to full describe the system we wish to model, due to the number of process or variables present in the system. This, however, can be overcome by additional research, measurement and so represents a difficult but not insurmountable obstacle to be overcome. This type of uncertainty is referred to as *Epistemic Uncertainty,* and relates to inaccuracies with measurement devices and processes. The second category relates to the timing of natural events, and the inability to determine natural variations within the system. This uncertainty is *Aleatory Uncertainty*, or stochastic uncertainty, that must be accounted for, but may not be measured. It consequently represents 'unknown unknowns' in the system that may not be overcome by simply improving or extending measurement processes. Walker et al (2003) defined the areas of uncertainty in three terms; dimensionality, which refers to the type and nature of the uncertainty; the regions of it, which refer to scale; and location of it in the model and types of uncertainty, which refer to the epistemic and aleatory terms above.

In order to quantify the impact of uncertainty, a statistical approach to analysing the data must be made, where the distribution and range of the uncertainty associated with all relevant inputs is known or can be accounted for. From this the uncertainty can be propagated through model realisations in order to create detailed results that represent all possible outcomes for the simulations. A number of studies and approaches have been created and utilised in flood inundation modelling in order to perform this type of analysis. Typically a calibration process is adopted, where model realisations based on multiple parameters are evaluated to observed data. Approaches include Monte Carlo sampling where a range of factors are tested based on probability distribution (Aronica et al 2002). A number of more sophisticated methods have been developed. The Generalised Likelihood Uncertainty Estimation (GLUE) methodology, developed by Beven and Bingley

(Beven and Binley, 1992), has been popular in hydrology and hydrodynamic modelling. It relies on an element of subjectivity in the analysis of results to ensure that only realistic model results are considered in the final stage of analysis. Broadly, the method relies on some estimation of probability distribution for each parameter. This is then used to develop a series of tests that are used in a Monte Carlo style but with each model realisation being described with a likelihood weighted value, based on the probabilistic values associated with each parameter. The results are then evaluated to some model performance criteria, and then accepted as being a realistic model or rejected based on this value and a weighting applied. A threshold value is required in order to evaluate the realism of model, which has been argued as reducing the method's robustness in providing a method to evaluate uncertainty (Mantovan and Todini, 2006). A distinct advantage of this approach is that the concept of equifinality is incorporated into the results, which is the principle that numerous parameter sets (combination of parameters required for a single model realisation) can produce similar model outputs. GLUE has been used extensively in research (Aronica 2002, Hunter et al 2005, Bates et al., 2004, Jung et al., 2012, Brandimarte and Woldeyes, 2013, Nott et al., 2012), due in part to the Bayesian principles within the methodology that allow for the progressive improvement of modelling results and the calibration process. The use of subjective threshold values can impact results however. Other methods include Bayesian approaches by (Hall et al., 2011), the Hydrological Uncertainty Processor developed by Krzysztofowicz (2000), and systematic approaches, such as used in benchmarking studies (Hunter et al, 2008). Each approach is based on assumptions that must be relevant when considering the overall objectives of the research.

Each approach requires knowledge of the dimensionality and location of uncertainty. Using the model construction framework outlined by Gupta et al (2012), the key areas of uncertainty in modelling can be defined as Input Uncertainty, Computational Uncertainty, and Output Uncertainty. The following section will outline the main aspects of these areas and subsequent examples, before addressing research that has been done to either explore uncertainty through analytical methods and communicating it through model results. A number of terms are defined here which will be used throughout this and following sections. The term *factors* will be used to describe model inputs, parameters and boundary conditions in a broad sense. This is in keeping with the terminology defined in sensitivity analysis and accounts for the number of different inputs required in

constructing any computational model, but particularly environmental models (Saltelli 2000).

2.4.1. Input Uncertainty

This category can be split into 3 specific sub sections; hydrographic uncertainty (also boundary condition uncertainty), topographic uncertainty (which relates to both measurement data uncertainty and feature representation) and parameter uncertainty. Each subsection has received considerable research with varying levels of detail and approaches. A few studies have explored all of these factors in a single framework, in particular Hunter et al (2008) which will be referred to throughout this section.

2.4.2. Hydrographic Uncertainty

A critical component of a flood inundation model is the inflow boundary conditions, which can be described in terms of discharge, water level and in some chained models, rainfall. Each boundary condition provides the forcing data with which to run the model and to propagate water through the channel and routed across the floodplain. As such it represents an essential factor in modelling, and the uncertainty in the data is critical to ensuring accurate model results. It has been speculated that the uncertainty in the boundary condition is the most significant in all inundation modelling (Hunter et al 2008), but describing the uncertainty in a sufficient detail with which to proceed to uncertainty analysis can be difficult. Hydrographic uncertainty concerns many areas, from direct water level measurements through to conceptual modelling issues, an overview by Di Baldassarre and Montanari (Di Baldassarre et al., 2009) provides a detailed analysis of the sources of uncertainty in river discharge estimation.

2.4.2.1. River Gauge Discharge

This is the input boundary conditions for most flood inundation models. Typically river discharge estimations are determined by recording measurements in a river using the velocity area equation to determine the discharge;

$$Q(x,t) = A(x,t) \times v(x,t) \qquad (2.26)$$

Where A is the cross section and v is the recorded mean velocity. This data is recorded over a period of time in order to determine a relationship between river stage and river discharge. This forms a number of points that allow a rating curve to be developed, where a stage value relates to a discharge value, which can then be

extrapolated to determine a wider range of flow conditions. The formation of the rating curve is based on analytical functions, with the power law relationship being the most common. This process is common to practical methods of monitoring water levels (Pappenberger et al., 2006), but contains a number of uncertainties. As mentioned by Di Baldassarre and Montanari (2009), both the measurement of A and v will introduce uncertainty. Research has estimated these uncertainties at about 5-6%, of the estimated value, although higher estimates have been made (8-20% by Pelletier et al (1987)). In the research by Di Baldassarre and Montanari (2009), the greatest source of uncertainty is related to the rating curve uncertainty, which is a product of interpolation of the rating curve beyond the measured points, the presence of unsteady flow conditions during the measurement phase, and seasonal changes in the roughness parameters, which will affect the rate of flow. Using a case study in the River Po, and the HEC-RAS model, the combined effect of the rating curve uncertainty is 21.2%, with a range of 6% and 38%. The total uncertainty range of river flow data is calculated at 25%. Further assessment of rating curve assessment and how it affect inundation modelling was undertaken by Domeneghetti et al (Domeneghetti et al., 2012). This research used a synthetic rating curve approach to test model results and then propagating the uncertainty to the 2D flood model. It is again noted that the contribution of the rating curve function to model uncertainty, and that this uncertainty has a strong influence over the choice of Manning's friction values in inundation modelling.

2.4.2.2. Hydrological Modelling Uncertainty

Hydrographic models are used to provide input in to models in scenarios where observed data is limited. In this scenario rainfall is used to derive a hydrograph through a rainfall/runoff model. The use of these models also contains a number of uncertainties from conceptual to simplifying assumptions that will impact on results (Romanowicz and Beven, 2003).

Aronica et al (2012) studied the impact of the influence of rainfall models on inundation model results. The study was based on the Glasgow test case first used by Hunter et al (2008). Rainfall data was used from a gauge outside of Glasgow, which provided a large enough temporal resolution with which to determine the magnitude of rainfall for a duration period. This data was analysed to determine the probabilistic occurrence of maximum annual rainfall events, the duration and associated rainfall levels. This then provided input to a rainfall/runoff model, from which a number of synthetic hydrographs were then created and run through a 2D hydraulic model, first used in Aronica et al 1998 (Aronica et al., 1998) which ignores

the convective term from the shallow water equations. It is therefore analogous to LISFLOOD-ACC. A modelling ensemble of 200 realisations with unique hydrograph inputs was analysed by using control points and probabilistic hazard extents determined by using water depth and velocity functions. The results indicate that the rainfall uncertainty has a spatial aspect; where water accumulates the uncertainty in the rainfall is more apparent than for control points located in flow paths. In these regions parameter choice is more apparent. This highlights the complexity of the problem of determining significant parameters, in that a strong interaction between parameters exists both spatially and temporally. The research also highlights the uncertainty from rainfall/runoff modelling, how it propagates and its impact on model results.

2.4.3. Topographic Uncertainty

The advent of remote sensing techniques to capture topographic data has allowed significant advancements to be made in the field of flood inundation modelling, allowing extensive 2D modelling to be undertaken over large spatial scales (Bates and De Roo 2000, Marks and Bates, 2000, Sampson 2012). Previous approaches to 2D modelling relied on digitised OS data and interpolated 1D model results, which were based on surveyed field data. The use of Digital Elevation Models (DEM) derived from a number of methods, but in particular the Light and Radar Detection method (LiDAR), has allowed detailed modelling to be undertaken (Hunter et al 2008,Schubert et al 2008, Fewtrell et al 2008, Wang and Liang, 2011). Other remote sensing techniques have been used to create terrain data, such as the SRTM (Shuttle Radar Topographic Mission undertaken by NASA) satellite data, which provides low resolution data, but at a cheaper cost and with a greater coverage of the earth surface. This data set has proved to be useful at large continental scales, such as the Amazon as reported by Trigg (Trigg et al., 2009) and Wilson et al. (2007).

Whilst the use of this data has greatly improved modelling, it has distinctive uncertainties. Cobby et al. (2001) provides a useful overview and critique of the use of LiDAR. In this research, two significant areas of the use of LiDAR are highlighted, the introduction of a random error, estimated by the authors as a root mean square error (RMSE) of ±15cm for every elevation calculation point made, and the interpretation of LiDAR data where vegetation has been encountered. A standard algorithm has been created which eliminates the areas of vegetation and urban features from the original, raw data to leave a more accurate representation of the ground surface (Mason et al., 2007). These digital surface models (DSM) provide

the terrain data sets for a large number of flood simulation models in the UK. As a result of the improvements in data acquisition and the wide availability of it modelling of regions with complex terrain is possible, such as urban areas. As the grid resolution has improved, so has the relevance of the LiDAR error to affecting model results (Hunter et al 2008).

Tsubaki and Kawahara (Tsubaki and Kawahara, 2013) research the impact of LiDAR RMSE on hydraulic modelling. The research reviewed the effect of DEM error on model results in a idealised urban area, and the effects on a rural domain model using comparative model runs of a range of DEMs with varying error, cell size and representation. For 3 test cases (an urban, rural and plain topography cases), a standard DEM and the same DEM degraded by a randomly generated error DEM, based on the RMSE of LiDAR. The impact of error was found to be considerably less in the urban area, where the complexity of the flow path becomes more significant. In this scenario it was determined that the random error had little impact on dominate topographic features that controlled flow paths. This was further confirmed when exploring the resolution of the cell size, where increasing cell size reduces the predictive performance of the flood model results. This paper highlights the importance of underlying topographical representation as an important aspect of inundation modelling, while also highlighting that in model domains with reduced complexity, the importance of the LiDAR error becomes more significant. LiDAR error was also studied as part of the Hunter et al (2008) benchmarking paper. In this scenario, the error was determined to be a significant aspect of affecting model results. This was in part due to the full SWE models producing a wider flood extent for the Glasgow benchmark study, but which was only 5cm deep. It was highlighted that this was lower than the DEM error, and that this could be a contributing factor to the variations shown between model codes.

Apart from the random error in the data, the issue of topographic feature representation is also another area of uncertainty in topographic data The importance of representation in topographic data is highlighted in two papers by Fewtrell et al (2011) and Sampson et al (2012). Using terrestrial LiDAR Cloud Scanners, a high resolution DEM for a location in Alcester, Warwickshire was created. This region had experienced flooding during the 2007 UK summer floods. Cloud Scanning techniques create millions of geo-referenced data points to create highly detailed surveys of locations, with a point measurement RMSE accuracy of ±5cm, depending on the surface being scanned. It should be noted that the level of accuracy will increase dramatically on glass and metal surfaces so its application to

highly built up inner urban areas is yet to be determined. The accuracy of the position of the data point is dependent on the control network, determined before the use of the cloud scanner. In the case of the papers mentioned above, the UK Rinex GPS network was used with SmartNET automatic referencing system, which allows for rapid data collection, with a reasonable RMSE of 90mm. The data sets were then processed into a 10cm DEM, which could resolve key urban topography features such as kerbs, drains and crucially the centre ridge of higher elevation present at the centre of roads. This is crucial in the movement of water, and is critical in the design of roads and drains. The authors highlight that in representing this, a significant alteration is made to the flow of water at the initial stage of flood water routing compared to a standard 2m resolution DEM model constructed from LiDAR derived data. Four DEMs were created from this high resolution data at 50cm, 1m, 2m and 5m. Using LISFLOOD-ACC and ATS, several model realisations were undertaken. The results indicated the importance and influence of resolving elevation at critical points in the road network, and at a sufficient scale. Water flow is directed down one side of the main roads in the region, which act as main flow paths in the model. Over the course of the model run this creates a small variation of +-15cm at some of the highlighted control points. Further analysis of control points revealed that in comparison to the higher resolution DEM, the coarse models failed to flood certain areas, produced lower depths or different velocity factors. In comparison at this scale, the choice of model type was relatively unimportant, although the authors note that the use of two simplified approaches may produce bias in the results and that a full dynamic model would be required to produce significant conclusions. The significance of the cell size is similar to Fewtrell et al (2008), in that it must be at a scale to sufficiently resolve the features present. The paper by Sampson et al (2012) furthers this, using a range of inflow hydrographs in addition to a high resolution DEM (at 10cm) and coarse DEMs (at 1m) from both airborne and terrestrial sources in the same case study location. This work also compared two different approaches to inundation modelling, the LISFLOOD model and ISIS FAST, a formulation based on pre processing of the DEM into sub basin units and solving simplified physics intercell flux and mass calculation between them. As with the previous work, it was found that model sensitivity was high in relation to the underlying topography, particularly with vertical height elevation in the road. An important aspect raised, however is the relative computational cost of this approach which is significantly higher, and would not create a significant change to the vulnerability associated with a flood.

The representation of buildings is a important aspect in flood inundation modelling as it impacts the choice of cell size and grid formation. In urban based domains, this problem becomes particularly acute, where high resolution spatial domains are required to solve flow through complex flow paths and highly variable topography. A paper by Schubert et al (2008), explored the different methods of representing buildings and the variations in the grid cell structure on model results. Four main categories of building representation were devised and tested for the same Glasgow test case area as had been used by Hunter et al (2008). Each category attempted to reflect the range of possible choices of boundary conditions devised and used across typical hydraulic models. Building Blocks method (BB) represented buildings as physical blocks located within the footprint of OS derived building footprints. The block elevation is set at 6m above the Digital surface model (DSM). Precise treatment of the buildings is non essential as the depth of flow and quantity of water is below 6m of water at any point. A building hole (BH) method was also created using the DSM model, and within the footprint of the building, a blank hole in the DSM. Using the BreZo code (Sanders 2000), two different boundary conditions were invoked by these methods, a slip boundary, and a dry wall boundary. In this way, the two physical representation methods provide an insight into the mathematical formulations behind the different representation methods. A representation method based on using high friction values in the cells that fall in the footprint of buildings was also tested. This method known as the Building Resistance method has been used previously (Liang et al 2008), and provides a computationally simple solution to representing buildings, as well as providing some level of representation for water flowing through the footprint of the building. These approaches were then tested for different flow conditions, and with different grid meshes. It was also tested alongside a range of friction parameters and computational models. This benchmarking approach identified that each building representation approach did not produce a significant variation in output, but did represent a significant trade off between computational costs and pre processing options. The additional factors benchmarked suggest small variations, but a significant conclusion was that for coarser grid resolution, model types converge in terms of flood extent and depths. This indicates that parameter interactions are a significant aspect in assessing model types and factors, as well as the fact that cell size will form a key part of that.

Further exploration of this work was conducted in a later paper by Schubert et al (2012), in which different topographic representation techniques were compared in a test case based on the Baldwin Hills Reservoir Dam Failure event that occurred in

California, USA in 1962 (Schubert and Sanders, 2012). An additional methodology of building representation based on a concept of representing the buildings as porous materials and using porosity equations to solve the free surface equations in these areas, was used (Sanders et al., 2008). This Building Porosity method (BP) has the advantage of representing part of the flow that moves through the buildings, as well as not requiring a high resolution mesh, like the BB method. An additional term parameter is added to the model code, φ, which represents the amount of space available to the water in the porous material. Sanders et al defined porosity in three separate terms, a convective porosity term a capacity porosity term, which represented the amount of space available at cell interfaces for the flux, and in the cell. The final term defines the cell based frontal area, which contains a drag coefficient. These terms are combined to create an additional term in the computational model that is then estimated on a cellular basis. In this test case, local variations between the methods did not produce a significant change to the estimated extent of the outflow. Local changes were restricted to arrival time of the wave and some of the flow paths that resulted from the failure. The test case had some observed material with which to evaluate model results. This highlights the applicability of the methods to a number of approaches, but also highlights that methods to analyse the models that contain local information in the evaluation technique may produce a different set of conclusion.

The issue of grid size in urban areas was studied in detail by Fewtrell (2008). Using the Glasgow test site used in Hunter et al (2008) paper, a number of models were created using variations in cell size and resampling methods, and compared to a high resolution grid model. The results indicated that a grid scale that is equivalent to the main building structures and road features is sufficient to produce reliable model results. Models run at a coarse 16m and 8m resolution created blockages to flow paths that were critical in model performance, where as the 2 and 4m models were more accurate. The choices of resampling techniques were less critical than the cell size, although methods to resample may have been affected by the steep nature of the site. In this respect the grid resolution is a more significant factor than friction values, which were also varied in this test, but were less of a sensitive factor in affecting model results. An important conclusion is that the choice of cell size is strongly location dependent, and must be assessed relative to the buildings and topographic features present in the domain. This must be assessed before model runs can be undertaken.

The significance of river conveyance in affecting flood inundation modelling has been noted in several papers (Bates and De Roo 2000, Trigg et al 2009, Wilson et al 2007). The values of friction and the physical dimensions of the channel model have been shown to be among the most significant factors in modelling (Pappenberger et al., 2007b, Pappenberger et al., 2008, Hall et al., 2005). A detailed analysis of this factor and the impact on model results has been undertaken by Fewtrell et al (2011). Using 3 different methods to represent river geometry and channel representation and 2 variations of hydraulic structures in the channel, a detailed analysis of the effect of channel representation was undertaken for the river Eden in Carlisle, using post event wrack marks surveyed with differential GPS to evaluate the model. The river channels were based on increasing complexity in representing the longitudinal profile and the cross section profile, and two additional models each with simplified representation of bridges. In the case of the bridge models, one channel was represented with the measured bridge geometry and utilised a head loss function to calculate the impact of the bridge on flow and discharge values, whilst the second was a simplified flow constriction cross section. Both LISFLOOD and ISIS TUFLOW models were used in this test case, with both LISFLOOD and TUFLOW used for the simplified approaches, and TUFLOW used for the complex river models. The models and river geometry were then compared with a range of constant friction parameters in both the river and the floodplain. The authors concluded that the ability to represent the channel capacity was a critical aspect in flood inundation modelling, and the performance of the model relative level of topographic representation, whilst critical to impacting model results, was within the uncertainty boundaries of the observed data set. The results indicated that the backwater effect was as great as 10cm affected for a great a distance as over 2km from the flow constriction point. In high flow conditions, however, this impact may prove to be a critical to the flood inundation extent. A clear point from this research is the relative influence of each of the factors. Simplified numerical models can produce acceptable results, but the full SWE models are more robust and less dependent on calibration processes. Channel representation becomes increasingly critical in high flow conditions, and that the choice of bridge representation will also impact on downstream and upstream flow.

The issue of representation of hydraulic structure has also been considered in an uncertainty analysis framework by Brandimarte et al (2013). In this work, the effect of bridge representation on the backwater effect, which is the increase in water surface level behind a flow constricting object in sub critical flow typical of rivers, is

studied. The reason behind the research was to explore the impact of using a deterministic approach to modelling this effect in an engineering context. Using a pragmatic approach to the problem, and predetermining the key factors to use in the uncertainty analysis, a systematic study of the impact of friction value choice and hydrographic input was undertaken. The main results focused on a comparison of deterministic approaches vs. an uncertainty approach, and highlighted the need to carefully consider the method of uncertainty analysis and the potential drawbacks of a deterministic approach. Using a test case from the Tallhalla Creek and the HEC-RAS model, an uncertainly analysis was undertaken using a GLUE approach to determining the input parameters. 100 inflow hydrographs were created based on return period values with a suitable range. It was determined from this that a significant cause of uncertainty was from friction values. The research concluded that the uncertainty analysis provided a more robust means of evaluating downstream flow, but that practical considerations must be accounted for in the undertaking of this analysis.

2.4.4. Parameter Uncertainty

This sub category represents the most extensively researched area in flood inundation modelling uncertainty, in combination with other factors. Here an outline of some of the methods and the principles of this issue are covered. Parametric uncertainty covers the choice of values of parameters, and in particular the choice of roughness coefficient. This type of uncertainty has particular relevance to modern approaches to flood inundation modelling and with this research as the advancement of simplified codes such as LISFLOOD-ATS and other diffusion wave based codes, the roughness parameter is an effective parameter. This is where the value taken by the modeller can account for simplifications made in the conceptual model, and theoretically produce a model results that replicates full SWE model results. Consequently, the value of the parameter may differ dramatically from the value related to the land coverage. However increased complexity of the model does not reduce the main issues of uncertainty associated with the choice of friction value, as the ‘real’ value will vary in both spatial and temporal scales. Furthermore the friction value cannot be determined everywhere or measured at sufficient detail for modelling purposes. Determining a single value for a model run may represent a significant simplification of reality (Aronica et al 2002, Hunter et al., 2007), although practical methods for overcoming this sort of issue are almost impossible, due to the changing nature of this value over space and time (Beven, 2007, di Baddalesse 2012).

Aronica (2002) undertook a detailed exploration of this issue, using Generalised Likelihood Uncertainty Estimation method (GLUE), a Bayesian approach to determining the influence of factors on model results(Aronica et al., 2002). The work was based on a 1D/2D LISFLOOD model for the Imera river, in Sicily and the Thames, UK. The parameters varied are friction parameter for both the floodplain and the channel. The work highlighted the significance of the choice of friction parameter in influencing model results, using both model performance gradients over parameter space and probabilistic flood inundation extents to display data. A key point in this research is the sensitivity of model results to the friction parameter and the need for careful consideration of these values in modelling.

An evaluation of the spatial variability associated with the parameter value was undertaken by Hall et al (2005). This research divided a river and floodplain into 3 sections and explored the relationship between the values of friction parameters for the floodplain and the channel, as well as input boundary conditions and channel width. Using Sobol' Indices to quantify which factor contributes most to the variation in model results, it was determined that channel friction values and channel parameters had the greatest influence over the model results. The work also provided further insight into the spatial influence of parameters, noting that the inflow had a greater influence at the top of the model, close to the boundary inflow point, and decreasing influence further down the model. A significant point, and one that has been noted in previous work is the relative lack of importance that can be attributed to the floodplain friction parameter, with the model results indicating that this factor is of low importance across in any of the floodplain sections.

2.4.5. Output Uncertainty – Methods of Communication

Output uncertainty covers both model results and evaluation methods. The uncertainty in model results refers more to methods devised to communicating uncertainty as this is an important aspect of any modelling exercise, particularly where modelling is used to inform decision makers (Beven 2010).

The issue of communicating uncertainty in hazard estimation has been studied by Aronica et al (2012). As well as studying the importance of model input to modelling results, model results were displayed in terms of a probabilistic hazard classification. Traditional hazard results are displayed simply as an extent, without further knowledge relating to the modelled velocity or the combined impact of water depth and velocity. This approach not only combines these aspects it also provides further information concerning the modelling results. Aronica et al (2012) note that

in situations where flood extent is insensitive to inflow boundary conditions, the use of hazard as a model output provides a more robust and reliable method for communicating risk, as well as highlight specific locations in the flood domain where the risk is greater. Work concerning expert insight into communication has been researched (Pappenberger et al., 2013, Warmink et al., 2011). In this work, the issue of how to present uncertain data is explored. Both case studies identify the need to use ranges of data where at certain thresholds an action or hazard can be expected. An important aspect is simply the provision of information, rather than a simple binary option. By using probabilistic approaches, more information is provided to decision makers concerning the ensemble approach of modelling and the potential range of events that could occur. It was felt that this was sufficient to allow decisions to be made with greater confidence.

An important aspect of communication is the use of simple metrics to describe uncertain parameters. Sensitivity analysis provides a methodology that allows for analysis of the parameters, as well as evaluation criteria, such as importance measures, with which to describe model uncertainty.

2.5. Model Evaluation Techniques

The need to calibrate models also introduces uncertainty related to the observed data and the model evaluation technique used to assess models.

Satellite and airborne image data is an essential component in determining flood extents, and providing data with which to evaluate models. Work by Horritt (Horritt, 2006), Mason et al (2007) and (Mason et al., 2010) have explored statistical methods by which flood extents can be outlined. By using this data in conjunction with LiDAR data, depths can also be extracted to enhance the calibration process (Mason et al 2012). However, given the uncertain nature of the extraction methods, calibration methods may require some knowledge of this uncertainty. Pappenberger et al (2007) introduce a fuzzy set methodology, which allows models to be evaluated with information about the uncertainty in satellite imagery incorporated into the calibration process. The uncertainty related to satellite data has also been explored by Stephens et al (2012). Using an example from the River Dee in the UK and using ERS-2 SAR image of a 2006 flood, the spatial dependency and sources of error associated with not only with the data, but also the evaluation technique. The study notes that not only is ‘noise’ associated with satellite imagery a significant source of uncertainty, but using binary measures such as F^2 is also significant in not identifying suitable parameter spaces. The use of F^2 has been

questioned in other studies (Prestininzi et al., 2011, Pappenberger et al 2008), and its use is often questionable in steep sided valleys.

Other model evaluation techniques have also been studied for effectiveness, such as the Nash-Sutcliffe coefficient. The NS coefficient has long been used as a means of model communication, although it has tendencies to award models that replicate shapes of curves, rather than peak values (Pappenberger et al., 2005). An evaluation of the technique was researched by Gupta et al (2009). In this study, the NS coefficient was decomposed into components. It was determined that NS consists of three components, representing correlation, bias and variability. In order to improve the method, an evaluation technique that incorporates these 3 components individually is suggested, by using the Euclidian distance of the three components. This research highlights not only the imperfect nature of the approach, but that methods are required to assess how to improve these approaches, rather than to consistently use methods that contain uncertainty, but are well established.

In the presence of imperfect evaluation techniques, multiple calibration techniques present a methodology to evaluate uncertain models. A case study by Dung et al (2011) for the Mekong Delta, Vietnam, used both the Nash-Sutcliffe and F^2 method to evaluate the model, and then explored the Pareto front, created from the multiple model evaluations (Dung et al., 2011). The results indicated that models with high values for one of the functions rarely performed as well for the other function. This compromise in parameter still allowed a set of parameters to be correctly identified, and highlights model deficiencies. A lack of representation in the dike systems was identifiable through this approach. Although this approach is unlikely to be widely applied as it requires data that is often not available for all occasions, the uses of comparing models across multiple functions is critical to identifying model parameters.

2.6. Sensitivity Analysis

As outlined above, uncertainty analysis investigates the potential of inputs to affect the output. Sensitivity analysis is described as the inverse of this problem that is to determine which parameter has contributed significantly to the uncertain model outputs (Hall et al., 2009). It represents a key level in determining important and influencing parameters to model results. The overall objective of sensitivity analysis is to determine which input contributes most to the overall output uncertainty, and there are 2 main approaches for dealing with this, local sensitivity analysis (LSA) and Global Sensitivity Analysis (GSA). Local Sensitivity Analysis deals with

determining the time and spatial derivatives of model results at specific location. In this sense it represents a more detailed investigative tool, but lacks the ability to deal with the complete model being investigated, due to its analytical nature (Saltelli et al 2000). Its use is limited in flood modelling setting, as the scale of the modelling problems are often greater and more significant than localised modelling issues. The second method, Global Sensitivity Analysis provides the detailed analysis required of most environmental models, to determine the impact of inputs over the entire model run. The use of these approaches in flood modelling is limited, however, due to the computational demand required to perform some of the standard sensitivity analysis methods. A key factor in using GSA methods is knowledge of the relationships between, and the distributions, of inputs.

2.6.1 Approaches to Global Sensitivity Analysis

A number of sensitivity analysis approaches have been developed. An outline of the main methods is provided here.

2.6.1.1 Variance Based Methods

The most prominent approach for sensitivity analysis is variance based approaches, as variance provides a natural indicator of output variation as a product of input distribution, and is used in standard statistical analysis methods such as Spearman Ranks, Student T tests and Pearson coefficients (Saltelli et al 2000). Broadly, variance based methods, use information in the model output to determine the importance of model inputs. The outputs from these approaches are correlation ratio/importance measures. The importance of an input factor *x* can be assessed by considering the conditional probability distribution of y (model output) with respect to *x,* to determine the importance measure of x to output as in the equation;

$$p_Y(y) = \int P_{Y|x}(y|x) PX(x) dx \qquad (2.27)$$

Where $p_Y(y)$is the conditional distribution of output y, with respect to the conditional probability of PX, the distribution of the input parameter set. This essentially links the output from a set of models to the distribution of the input and forms the basis for the variance based approaches. Examples of this approach include the Fourier amplitude sensitivity test (FAST), (Cuiker et al 1970), the Ishigami function (Ishigami and Homma, 1990) and Sobol Indices (Sobol 1990). Of these methods, the Sobol method has been used widely in research, as it provides robust results with a reduced computational cost compared to the other methods (Hall et al 2005, Hall 2009, Pappenberger 2008). The Sobol method is a Monte Carlo approach to

solving the integrals of the variance calculation. The total variance is decomposed into integrals of increasing dimensionality, from single parameters to combined parameters. The contribution of a individual parameter to the total variance of model output, D which is based on the total distribution of model results, is calculated with the following equation

First order $S_i = \frac{D_i}{D}$ (2.28)

And

Total $S_{Ti} = \frac{D \sim i}{D}$ (2.29)

Where D_i is the variance attributed to parameter i, $D \sim i$ is the variance attributed to all parameters except i.The first order term therefore represents the direct contribution of that parameter to influencing the variance, whilst the total is the combination of the parameter and all its higher order interactions. Higher order interactions occur when 2 or more parameters begin to influence the results as a result of their interactions. Results with low first order, but high total indices indicate a model with highly interactive parameters.

The FAST method differs from this approach by converting multiple dimensional integrals into a single integral that is based on a search curve that explores each individual factor with a different frequency in order to define the parameter space (Saltelli et al 2000). The FAST method has the advantage of being computationally cheaper than Sobol in calculating total indices. Both approaches have the advantage of using indices that rank from 0 to 1, based on the contribution of a parameter to output variance, making interpretation of results simple. An additional advantage is the model independence of the approach which allow the inclusion of dependant parameters that won't impact model results (Saltelli et al 2000). These methods are, however, computationally expensive. The original method for determining the total and first order indices for the Sobol method with a parameter space of size *k*, required *n x (2n +1)* model evaluations. For a case with 5 parameters, that can extend to nearly 2000 model runs. Sampling approaches, such as Latin hypercube, and winding stairs, which aim to replicate the parameter space with fewer model runs, but whilst retaining the same factor as an output.

The use of variance as the mathematical moment to describe is challenged by Borgonovo (2006), who created a moment independence approach to SA. This does not rely on using descriptive statistics to determine variation of output relative to input, but measures the variations in surface area of the response pdf to determine a measure of sensitivity, δ (Borgonov 2006). This method has the

advantage of not requiring detailed information about each factors properties and a measure of contribution from 0 -1. Use of this approach in an environmental context highlights the advantages of reducing the assumptions associated with inputs (Hartmann et al., 2013).

2.6.1.2 Screening Methods

The computational cost of the variance based method mean that for models that have either a high computational cost, or models that contain large number of factors, this approach become intractable. Screening methods represent a efficient approach to ranking factors, with a reduced computation cost compared to the Sobol methods. The principle behind the approach is to short list important factors, which control output uncertainty, whilst recurring computational cost through limiting the number of model realisations. A number of approaches have been created include One At a Time designs (OAT), group screening factorial fraction designs, such as the Andreas Iterated Factorial Fraction Design (IFFD) (Andreas and Hajas 1993) through to supersaturated designs, which use less model runs than factors. Examples of the OAT approach include Morris method (Morris 1991), where an initial sample of the parameter space is determined to create an initial model run. A factor is randomly chosen to be moved through its parameter space by a predetermined amount. The model is run again to create an output for the trajectory point in the parameter space, hence the OAT terminology applied to each method. Once all parameters have been sampled the elementary effect is determined for a parameter with the following equation

$$d_i(X) = \frac{y(x_i,\ldots..x_{i-1},x_i+\Delta,x_{i+1},\ldots..x_k)-y(X)}{\Delta} \qquad (2.30)$$

Where x is any value in the parameters space, Δ is the predetermined value by which the parameter value is evolved, and y is the model output.

The process is repeated n times, until the parameter space has been sufficiently well covered. Research indicates that convergence of the results for a 5 parameter space can occur with as few as 10 runs (Saltelli 2002). From the total model runs the mean elementary effect can be determined. This is equivalent to the first order effect of a variance approach, whilst the standard deviation of the elementary effect represents higher order terms. Analysis of these two factors by graphical means can identify significant factors, and identify interactive factors. This approach has been used previously in modelling studies, and has been shown to identify significant parameters and provide similar ranking values to a variance approach (Zhan et al., 2013)

The Andreas IFFD approach is an example of a group screening design approach, a carefully determined trajectory pattern that will determine main interaction effects and some higher order effects. The approach is grouping parameters, determine high, low and intermediate values and evaluate the models in groups. Results are evaluated based on the Stepwise regression which is used to evaluate the models. The disadvantage of the method is that in using grouped approaches unimportant parameters that coincide with important factors may appear more important. This is overcome by using Latin hypercube approaches (Andreas 1997).
The advantages of a screening approach, with respect to variance based methods have been highlighted (Herman et al., 2013), and present a computationally efficient means to undertaken sensitivity analysis, albeit qualitative.

2.6.1.4 Bayesian Methods

Where knowledge about factors is uncertain, Bayesian analysis can provide a framework with which to evaluate the inputs of a model. Bayesian approaches have been successfully implemented into uncertainty analysis frameworks (GLUE, Beven and Binley 1992). Broadly, a Bayesian approach involves describing prior assumption concerning input factors. This approach relies on using prior beliefs over the system, inputs and models to determine outputs. These beliefs are represented as distribution functions, which can be updated based on information with Bayes formula. This approach has the advantage of not requiring absolute distribution functions and knowledge for the inputs, as these can be updated. The resultant outcome is a likelihood of occurrence, which in modelling terms would be that a model represents the physical system. These applications are consequently extremely useful in determining uncertainty.

2.6.1.5 Additional Methods

Other SA techniques provide alternative insights into model performance, based on assessing local rather than global, variations. Derivative or local sensitivity analysis methods occur less frequently in analysis of models. The method involves varying inputs around a nominal value to determine the local impact on model results. The sensitivity factor is then a product of the effects of altering the input on the output. This approach is less popular in modelling however, as it tends to explore single factor impact rather than combined parameters. An example of a local SA method being used in a hydrological modelling exercise was undertaken by Tang et al (2007) in a multiple comparison research. This research used the Parameter estimation software (PEST, Doherty 2003), to determine the linear effects of factors on model outputs. This was compared with GSA methods, and was found to provide a less detailed level of analysis, as a result of the lack of information

concerning parameter interaction. Given the complexity of the models used, this is an essential component in SA of environmental models.

Generalised Sensitivity Analysis is broader statistical analysis. These techniques include correlation measures such as Spearman's and Pearson's, regression analysis and sample tests such as Mann Witney and Student t tests. These methods tend to be used in initial stages of analysis, as they suffer from a lack of robustness and some broad assumptions, such as the requirement for linear behaviour to describe model parameter relations in the student t- test or the requirement for parametric inputs in Spearman's ranking. Examples in research include the use of regression coefficients by Pappenberger et al (2006), which included using information about the behavioural aspects of the model to improve the robustness of the results associated with the Spearman ranking coefficients.

2.6.2 Applied use of SA with Flood Inundation Models

Due to the computational cost, the use of sensitivity analysis techniques to explore parameter importance in flood inundation modelling is limited. Hall et al (2005) undertook the variance based Sobol Indices method to investigate the significance of parameters in model calibration for a typical flood inundation modelling exercise. The LISFLOOD-ATS code was used to model a section of the River Thames, evaluating the model by using the F^2 method. A 6 level factor space was determined for the test case, including the inflow peak value, land surface elevation, channel width, channel elevation, floodplain friction and channel friction. Information for the distribution of the uncertainty about each factor was determined from previous estimates, however for friction values a set of initial model runs which formed the basis of calculating the distributions, using a model performance over parameter space figure as a analogue for a continuous distribution. A replicated Latin hypercube was then used to sample the factor space, and to calculate both first order and total sensitivity indexes a total of 1,700 model runs was undertaken. Further analysis came in the dividing the domain into 4 floodplain regions, and undertaking analysis relative to model performance in these regions. The results indicate that over the whole domain the channel Manning's friction value is the most significant parameter, with first order indices of 0.46 and a total indices value of 0.086. Further factors relating to the channel also have significant values, 0.21 and 0.26 total order values for channel width and elevation respectively. This indicates that the majority of variation in the F^2 value can be attributed to factors relating to channel conveyance. The relative insensitivity of floodplain factors can be attributed not only to physical properties of overland flow (Knight 2002), but also the

insensitivity of the ATS code to variations in floodplain friction (Bates 2010). Each subdivision also provides insight into parameter importance and model performance with 4 equally sized domains running downstream. There is a noted improvement in F^2 value in the downstream direction, although variance remains the same across all the sub domains. The Sobol Indices produce the same order of important factors, but with a significant interaction between channel elevation and channel Manning's value in the third sub domain, and an decrease in the relative rank of the input further down the model domain. This highlights a spatial structure This research details the advantages of these approaches in model development, and the advantage of using the Sobol method to quantify importance of modelling parameters. The cell size of the domain is relatively coarse (25m), allowing a large number of model realisations to be completed at a relatively low computational cost. This advantage may not be afforded in urban locations, where high computational costs would be expected, thus limiting the applicability of the method.

The number of GSA methods available creates uncertainty in determining the best approach. Direct comparison of GSA methods in hydraulic and hydrological modelling have revealed certain advantages of methods, and highlighted where methods can complement other results, and be used in conjunction with other approaches. A detailed study that used 5 methods was undertaken by Pappenberger et al (2008). In this study the Sobol indices, Morris method, Regression analysis, regional analysis and Killback-Leibler entropy method (a measure of how the model varies from a prior normal distribution), are used. These methods represent a range of possible approaches to SA. The case study is a 10km stretch of the River Alzette in Luxemburg, for a flood event that occurred in January 2003 with a peak discharge value of $67m^3/s$. The event had a recurrence probability of 1 in 5 with water depth data available to evaluate model realisations using the Nash Sutcliffe efficiency. As with the Hall 2005 paper described above, the model domain is divided in sub sectors, this time to coincide with water level data. An initial set of model realisation were used in a GLUE uncertainty analysis, to narrow the parameter range for the sensitivity analysis. This created a 6 level factor parameter space which included Manning's friction values for channel and floodplain, a model weighting value theta, and 20 hydrographs. Downstream boundary conditions also formed part of the parameter space, which are a friction value and a downstream slope value. The factors were assumed to be independent and from a uniform distribution, and for the more computationally complex methods a number of bootstrap sample were undertaken to evaluate a smaller number of model runs. The results of important factors varied across the different sensitivity

analysis. Evaluating the combined ranking results of Sobol and the Killback-Leibler entropy method show few occasions of agreement. Both methods rank the inflow as the second most significant impact greater, and initial slope as a factor of less significance, but agreement with the order and significant of other parameters is poor. This highlights the uncertainty associated with the choice of sensitivity analysis, and in the definition of sensitivity. Each method makes a different assumption about sensitivity, which will impact on the final results and choosing appropriate methods is vital in making conclusions. This issue is overcome in the paper by creating an average rank for each parameter across the different methods and for the each region of the model domain. In Hall et al (2005), this analysis revealed a spatial structure to the sensitivity analysis results, with input boundary conditions being significant at the top of the model, with downstream boundary conditions having a greater impact for the lower sections. An important aspect in the research is the importance of interactions, in that parameters have higher total effects than individual effects. This not only complicates the analysis, it also introduces significant uncertainty about how to analysis parameter sensitivity. A significant conclusion is that in determining appropriate SA questions more robust conclusions can be made. This last point highlights the difficulties in applying these methods to flood inundation modelling.

The use of Multiple Global sensitivity analysis was explored by (Cloke and Pappenberger, 2009), and found similar results. The paper uses Sobol, Morris and Killback-Leibler entropy method to investigate the sensitivity of the ESTEL-2D finite element model for a stretch of the river Severn. The 11 level factor space was applied to the 3 SA methods, and found that Upslope pressure and Air pressure featured as a sensitive parameter in all approaches. However, this upslope parameter had only first order significance in the Sobol method, whilst higher order significance was found with the entropy method. Again, this research highlights the careful justification required in determining appropriate SA methods, and that due to the different definitions of sensitivity in the different methods, the ranking of important factors will also vary compared to other SA approaches.

A further study into the impacts of different sensitivity analysis methods was undertaken for a lumped watershed model by Tang et al (2007). Four methods were used, local sensitivity analysis, regional sensitivity analysis, ANOVA decomposition, based on model runs sampled with a IFFD approach, and Sobol indices, based on bootstraps. Two linked models were used, the SNOW 17 conceptual temperature block index snow melt model, and the Sacramento rainfall runoff model which created a 20 factor level design, including infiltration rates,

storage factors, rainfall melt factors and depletion values. The ANOVA and Sobol Indices produced similar results, as a result of both approaches being variance based methodologies. Both methods also proved more robust than local SA and regional SA methods, and could also describe higher order interactions in the results.

The use of multiple methods in an integrated form was undertaken for the DTVGM hydrological model in the Huahie River, in China by Zhan et al (2013). The model is a time variant hydrological model that is comprised of rainfall runoff model, flood routing and evapotranspiration models. As a result, it contains a large number of parameters and input factors. In order to create a well posed sensitivity analysis, the authors created an integrated approach, using the efficiency of a screening method, the Morris method, in order to inform a Sobol SA method. Further computational savings were made by using the results from the Morris method to create a response surface of the parameters, and use the response surface as the input to the Sobol analysis. The response surface could be adapted to fit the data availability creating a certain amount of flexibility in the approach to future research. In this paper the MSM approach was used. The analysis showed that 6 of the 14 input factors were considered significant, and that the approach only required 24hours of computer time, and 2000 model runs. The classic approach may contain as many as 100,000 model runs at a cost of 1 month of computer processing time. The advantages of using combined approaches are highlighted by this value alone. However, in applying this method, the inputs were considered to be of uniform distribution, which may limit its extension to other modelling problems. This assumption is critical in the use of emulators and many SA techniques.

The research and methods mentioned above have assumed knowledge of the structure of the uncertainty in the modelling inputs. Where this knowledge is itself uncertain, Bayesian approaches provide the means to evaluate parameters and undertake sensitivity analysis. Hall and Manning (2011) adapted a Bayesian approach to a calibration process. This is based around choosing appropriate values for parameters based on assessing a range of values against observed data to update information. In this way it essentially represents a factor fixing approach to sensitivity analysis, i.e. what factors should be fixed to ensure reduced variability in model outputs (Saltelli et al 2004). In order to reduce the cost of calculating the integrals of the posterior probability, emulators and Monte Carlo methods (MC) were used to replicate the model response surface. A 5 stage evaluation process was created, that involved determining prior estimations to factors and parameters, calculating regression coefficients, and determining a model inadequacy function. A

likelihood value was also required to determine behavioural and non behavioural models, a similar concept employed in the GLUE approach of Beven and Bingley. The calibration process is extended to include reconstruction of the model response by means of a Bayesian emulator. A worked example is undertaken with the LISFLOOD model for a stretch of the Thames, using a SAR dataset to evaluate model output with the classic F^2 approach and data from the Buscot weir to act as hydrological input. The results, demonstrate the advantage of this technique. Water depths and surface profiles were replicated well, and with uncertainty boundaries provided as a result of the modelling approach. In assessing the model, prior assumptions were also updated, with the friction value assumptions being made more precise after calibration. The paper highlighted several key points. First, that Bayes theory provides a sound framework with which to discuss and evaluate modelling uncertainty. By extending it to include a measure of model structure error, all the major sources of uncertainty in the modelling process are included. Secondly, the results are robust, even with a large number of uncertain factors and models. This point is an important aspect with future development of determining uncertainty in modelling, where determining multiple sources of uncertainty will be essential to reducing uncertainty. Furthermore, this method applies a more stringent probabilistic basis to derived results. Other methods, such as the cell inundation frequency diagrams used in Aronica (2002), and the GLUE methodology present results with a informal probability range (Manning and Hall 2011). This method improves on this by using formally defined probabilistic parameters to create results. Additional advantages also include less sensitivity to threshold values for defining behavioural models. Manning and Hall note that results are more sensitive to prior distribution values, but reflect that this is a natural aspect of using expert knowledge than creating arbitrary values for determining models. The ability to account for all sources of uncertainty gives a Bayesian approach a distinct advantage over the other SA methods.

2.7. Risk and Exposure Analysis

In the previous sections the main methods of model evaluation are based on direct model output, such as extent, depth or water surfaces, although determining vulnerability and risk associated with a flood event is a key aspect of inundation modelling. Many methods have been devised with which to estimate the financial risk or risk to life associated with these natural hazards, which are used to determine the financial benefit associated with improving flood defences and to estimate the reduction in vulnerability associated with new infrastructure. The

method is based on applying mathematical models of the relationship between depth of water and the damage associated with the level of water, such as depth/damage cost curves. The complete assessment therefore requires knowledge of water levels, and land coverage/building data for the region.

In some models, like the UK based Multi Coloured Manual (Penning-Rowsell 2010), a number of approaches are provided based on the amount of data and the level of detail required or available to decision maker. The assumption is that the data provided is the best available to conduct the risk analysis, and that the main source of uncertainty lies within estimations for the depth damage curve. In this sense the Risk Analysis techniques are deterministic, using a single value with uncertainty bounds based on the costing of materials rather than on the modelling process. An objective of this research is to explore how financial risk estimation methods can be used to evaluate models and to communicate risk. In this way, a more direct decision maker based tool can be developed to communicate modelling uncertainty and to allow non technical users of flood inundation models a more immediate concept with which to understand uncertainty. A number of methods for determining vulnerability and financial risk of flood have been developed. These reflect the number of definitions of risk and vulnerability, as well as the different requirements from decision makers to deal with the consequences of flooding (EA, 2010). For example methods range from highly detailed financial models which account for individual building material, hydrodynamic characteristics and warning times (Penning-Rowsell 2010, Merz 2004, Schwarz and Maiwald 2007, 2008), to more direct evaluations of simulated water depths and velocities to determine the hazard associated with an event (Defra 2010)

$$Flood\ Hazard = h * (v + 0.5) + df \quad (2.31)$$

Where h is water depth, v is velocity df is a debris factor, which increases in value in increasing depth, to represent the hazard caused by debris entrained in the flood waters.

2.7.1 Methods for determining vulnerability and risk

Methods for assessing flood vulnerability are based on economic and financial losses due to flooding. The losses may cover a range of direct and indirect costs (such as immediate loss of property and stock to the disruption of life outside the flood limit) as well as tangible and intangible consequences (loss of personal possessions through to loss of trust in authorities) (Merz et al., 2010). In assessing the consequence of a flood as a hazard with a risk outcome, vulnerability and loss must be defined and estimated. A number of approaches have been developed in

the literature and in wider research in order to estimate the cost of flooding, and the vulnerability of a local community.

One of the most definitive approaches provided is the MCM method developed by Penning-Rowsell et al. that provides a comprehensive approach to determining the financial implications of flooding. This methodology developed originally in 1970 uses a combination of synthetic data and recorded data to derive a number of damage depth curves which can be discretized to the level of individual units. This approach is similar to the German HYDROTECH and MUR methods (Merz et al 2004). The approach is flexible, and scale independent allowing the method to be used in a wide variety of approaches, depending on the user's needs and data availability. Broadly, each depth/damage function is specific to certain classes of building, and can be extended to be based on social classes. It represents a formal approach to defining economical vulnerability, that is the combined financial impact (the direct tangible consequence of flooding) and the longer economic issues associated with major events (indirect and intangible consequences).

The use of more subjective approaches to evaluating vulnerability has also been suggested. The advantage of a subjective approach are not to explore the financial and economic impact, but to determine key structures in the study area that are of key value to the local community or are determined to be more significant than other components of the study area by decision makers. The method, developed by (Pappenberger et al., 2007a)involved two approaches, a weighted cell approach and a localised model evaluation technique, to improve the process of model calibration. The weighted cell involves assigning cells with a weight value based on the objects within the cell i, j, and using this value to enhance the model evaluation technique, using the equation;

$$L(v) = \frac{\sum_{i,j=1} v_{i,j} * S_{i,j}}{n} \qquad (2.32)$$

Where $v_{i,j}$ is the vulnerability weight, $S_{i,j}$ the similarity factor and n the models in the complete ensemble of behavioural models. The vulnerability wieght is determined before, and can be altered to reflect temporal considerations (such as increasing the weighting value on roads for events that occur during rush hour, or decreasing weighting values for offices and schools during evening times) as well as reflect valuable assets in the study region.

A second method developed is using the global model evaluation method extent F^2 for local sub domains. As with the previous method it requires careful consideration of the sub domains to be chosen. In the paper, these approaches identified models

that had produced high global values that had then failed to provide accurate results in the areas of interest, for a case study from the River Alzette in Luxemburg. The weighting method identified regions in the model, where the estimation of risk changed considerably as a result of the vulnerability measures compared to the traditional global method approach. Such techniques provide the means to incorporate vulnerability without the requirement for additional information relating to population and local building type.

2.7.2. Applied Risk Modelling in Flood Inundation modelling

A number of papers have explored not only methods and models for describing risk, but also the relationship between these methods. Apel et al (2008) explored the major sources of uncertainty in a flood risk assessment, in both the hazard aspect of the modelling and the assessment of vulnerability and economic risk. Using a linked probabilistic model comprised of process orientated modules, where each module represented a physical element of hydrodynamic process and vulnerability, uncertainty analysis was undertaken for a section of the Rhine. The chained model comprised an inflow hydrological model, using gauged flow information from Cologne to determine flood return probabilities, a tributary input value, and a depth discharge relationship module. A dike breach failure module was also considered. The estimation of vulnerability was undertaken with a depth damage relationship module, for residential buildings that occurred with postmark defined parcels of land. The damage was determined with the following equation

$$y = (2h^2 + 2h)/100 \qquad (2.33)$$

Where, *h* is the depth of water. The research only considered residential buildings, and based the damage cost on 3 different depth/damage curves, including a square root method (HYDROTECH 2001), and a linear depth damage rate (MURL). These values were combined with the residential asset value to determine cost. The process modules contained uncertain inputs or processes with estimations of the uncertainty determined by statistical inference, as in the case of the hydrological input, where a GEV approach was used to derive the inflow, or estimated based on the nature of the model, for example the use of 3 different damage estimation models. Using a Monte Carlo approach to evaluate each component of the model generated 105 model realisations with which to determine the significant sources of uncertainty. The research identified that uncertainty for the entire model chain was greatest with extreme events. When analysing the depth damage risk curves, it found that uncertainty in depth damage curves was greater at extreme value occurrence events, as a result of the large range of values represented by the 3

curves, and also determined that the depth damage curve was the greatest source of uncertainty in the modelling chain when considering the annual maximum series as an input. This was determined to be a product of the fact that small variations in depth will produce a wide range of potential results based on the depth damage curves. This sensitivity to depth provides a unique aspect with respect to evaluating uncertain models.

This paper highlights not only the importance of the choice of model in determining risk, but also its importance in an uncertainty context, i.e. related to all sources of uncertainty in a modelling context. Importantly, the conclusions indicated that the module uncertainty was not additive, and that in reducing model uncertainty will also require a reduction in all sources of uncertainty.

This work was extended by a paper by the authors by investigating the level of detail required in flood risk analysis (Apel et al., 2009). 3 approaches of varying physical and process scale were used to determine the hazard of the flood event and the vulnerability of the surrounding area, with the aims of the paper being to determine the appropriate level of hazard model and vulnerability model required to match flood risk estimates, using a comparative approach. The hazard models were a linear interpolation model, using information from river gauges over a spatial region to create a water slope, which was then intersected with a DEM to create a water depth surface. The second level was the LISFLOOD1D -2D model, followed by the finite volume 2D method of Aronica et al (1998) the vulnerability models were determined by the scale at which they could resolve vulnerability. The most basic model were a set of meso-scale models, using the linear interpolation methods from the previous paper. Again as with the 2008 paper, this resolved depth damage curves in parcels of land. The next level in the vulnerability model was again based at the meso-scale but used a model based on surveyed data from private houses affected by flooding to create a Flood Loss Estimation model (FLEMops) model to resolve depth damage curve. The model contained data on 5 classes of depth, 3 building types, and 2 building qualities. A mean damage ratio was then calculated for the region based on percentages of building types and quality in land parcel regions using the following equation;

$$DR_{mean} = 0.31 * DR_{OFH} + 0.25 * DR_{SDH} + 0.44 * DR_{MFH} \qquad (2.34)$$

Where DR_{OFH} represent the damage ratio for family homes of average quality, DR_{SDH} is the ratio for semi detached homes, and DR_{MFH} is the damage ratio for multiple family homes. An additional term was added to include impacts of pollutants and contaminates. This model was extended to a micro scale, creating

the 3rd level in the vulnerability model. This involved using the macro scale model and applying it to individual buildings, which were assigned a mean property value. A second approach involved creating unique damage curves for each building type.

This research contained a unique dataset with which to evaluate the modelling approaches. For a flood that occurred in Eilenburg, Germany along the river Mulde, a maximum flood extent from both surveyed and satellite data were available. This was enhanced with flood depth data from over 380 locations. Datasets were available for flood loss estimates from the Saxonian Relief Bank, which dealt with the compensation claims for the event. It estimated a total cost of €77.12 million Euros, with additional information available concerning general costs for each claimant.

Overall, the choice of hazard model was less critical and produced less variability than the choice of vulnerability approach did. All hazard models produced a reasonable replication of the flood outline, where as a number of issues, including image resolution (to identify building types) and assumptions relating to building classification and representation. It was determined that at the meso-scale model was more than sufficient to create reasonable flood risk results. An important conclusion was that assessment of vulnerability models should be separate from hazard models, as issues of equifinality can occur, where poor hazard models can produce reasonable risk results, as a result of deficiencies in both approaches.

The use of Monte Carlo methods in flood risk assessment was investigated by Yu et al (2012). This research combined Monte Carlo methods and fuzzy probabilistic approaches to create fuzzy flood damage estimates and probabilistic damage contour maps. The objectives are similar to the paper by Apel et al (2008), in that it is a probabilistic approach to determining the main sources of uncertainty by using a modelling chain with different components and inputs to determine the overall variability, and attempt to identify the main source of uncertainty. The joint method is based on determining the uncertainty for factors relating to the hazard section of the model, which forms the Monte Carlo part of the model, and using fuzzy set analysis to derive first the water depth and membership function in a cell for all MC runs, and then using the depth data to determine damage ratio. A 3 level damage value was calculated, based on depth damage model developed by Benning et al (2000). A polynomial curve fitting was undertaken to derive a credible depth damage curve and a highest and lowest rated depth damage curve. In this way all aspects of uncertainty will be captured in the modelling process and propagate through to the final results, which are a total damage cost value, with 95% range

values and a probabilistic flood map. Using the LISFLOOD code as the hydrodynamic model, and a 14 level factor space for the analysis, a test case with 1,000 model realisations was undertaken. The results for the test case were a damage cost of $154,000 with upper and lower values of $266,200 and $104,900 respectively. The damage contour maps allow the identification of areas at greater risk, with relevant probabilistic values attached to them, although it is noted that the uncertainty related to the depth damage curve has considerable influence over this output. This method allows the uncertainty of the modelling approach to be considered in model outputs, and therefore be apparent to decision makers. It is noted by the authors that whilst the approach is broad in its process, the use of MC methods is potentially restrictive, but improves on the deterministic approach often used in economic risk analysis.

Merz et al (2010) provide a comprehensive analysis of the approaches available to determining risk, as well as the issues associated with certain methodologies. The paper outlines the main sources of uncertainty in the risk assessment modelling, noting that lack of process representation, temporal transformation (the change of risk to a region or element over hazard time scales and longer time periods) and a lack of validation data all contribute to uncertainty related to the models and the application of existing methods. A particular point noted is that small alterations with the hydrodynamic conditions may have a significant impact on damage calculations. This sensitivity to hydrodynamic conditions represents a counter point to issues with classical model evaluation techniques, such as the use of F^2 in steep sided valleys and in valley filling flood events, where the variation of modelled extent is small compared to the variation with depth.

2.8. Conclusions

A key area of research that has not been addressed is the assessment of physical representation as a source of uncertainty in flood inundation modelling. By using an integrated platform such as LISFLOOD this can be addressed in a systematic manner. Furthermore, by using this model, sensitivity analysis can also be used to evaluate the significance of this model. Another aspect detailed in the literature review is the methods by which the models are assessed in an uncertainty context. Uncertainty communication is an important area of research, and using methods normally used to describe financial loss, may provide an effective means of communication and a more In order to assess the research gap the following questions must be answered;

- How significant is the choice of physical representation in creating uncertainty, and how does it relate to other sources?
- How can model evaluation techniques be advanced to not only provide a more stringent analysis of model parameters, but also enhance uncertainty communication?
- How can sensitivity analysis be used to analysis these points in more detail?

The uncertainty associated with model code represents a significant component of modelling uncertainty that has yet to be addressed in an uncertainty analysis context. The construction of the mathematical model and specifically the level of physical representation represent the most significant aspect of the computational code. To date benchmarking studies (Hunter et al 2008, Defra 2010) and direct comparison studies (Fewtrell et al 2011, Horritt et al 2002) have explored the relative impact of each approach on case studies, but further work is required to explore it in a uncertainty context. In order to determine the overall impact of this choice, it must be considered in a structured format, comparing other factors considered important in modelling exercises and for which the uncertainties are better understood, but still to be resolved. Furthermore, in order to explore levels of physical representation a modelling framework within which individual modules of varying levels of physical representation is required, in order to isolate these aspects, and to remove the potential drawback of using codes with different discretization techniques, that may impact on model results (Neal et al 2011).

A current issue also highlighted within the literature review is the uncertainty associated with the methods used to assess model outputs, and to determine their performance at simulating real events. Typically model analysis uses binary extent comparison methods such as F^2 and Nash Sutcliffe which are favoured for assessing models, although they contain drawbacks. F^2 is based on extent which may not be the only appropriate measure of risk associated with floods (Stephens et al 2012). The calculation method is also restrictive in locations with steep valley sides. In uncertainty analysis and in calibration studies, this would potentially fail to provide an adequate description of model response to varying and uncertain parameters. In order to explore the appropriateness of these methods and to understand the implications of the choice of model evaluation technique, a number of methods should be assessed simultaneously. Where this has been undertaken, more insight is given to model performance (Dung et al 2011). Furthering this, the issues of uncertainty communication could be improved by using methods more relevant to decision makers (Pappenberger et al 2012). Depth damage curves and vulnerability measures could be used to communicate modelling uncertainty. These methods also have some of the advantages of including more implicit information

within them (such as water depths in depth damage curves) and being adjusted based on local requirements (Pappenberger et al 2008). The use of these approaches to determine appropriate levels of modelling has been considered (Apel et al 2010), but an extension of this would be to determine how appropriate the approaches are to calibration processes and model evaluation.

Sensitivity analysis forms a effective means of determining what model factors contributes to output variance (Saltelli et al 2000). The use of sensitivity analysis has also been identified as an essential part of model development (Hall 2009, Beven et al 2010), and has already been used to explore and identify key model parameters (Hall et al 2005) although the application of these methods is restricted by computational considerations. However, determining an appropriate SA approach would be essential to help determine questions concerning key model parameters.

3. METHODOLOGY

In order to achieve the objectives, the research was designed to compare each input in a systematic way and to use multiple objective functions to assess the model results. In order to remove model bias from the results, three test cases with different hydraulic properties were used. The test cases explore different aspects of the models and the significance of the level of physical representation. Two of the cases have observed data sets by which the model realisations can be evaluated. The main characteristics of each test are outlined in the preceding data chapters. This approach, of systematic evaluation of inputs is in keeping with the EA 2D benchmarking approach and the benchmarking approach of Hunter et al (2008), where each test examined a particular aspect of model performance, and by using real life scenarios, ensures the research results relevant to commercial modelling practices. Two approaches were used to meet the main objectives of the research;

- A Monte Carlo uncertainty analysis approach, which uses a systematic approach to determining and testing parameters. This approach is designed to use values of inputs that may typically used by modellers, based on the range of uncertainty of the input.
- A Global Sensitivity Analysis approach to quantify the impact of parameters and inputs on outputs.

For the first part, each test case was used as the basis of a Monte Carlo test whereby a range of parameters and inputs will be tested in a systematic way, with each level of factor (the value of the considered input based on how the input is discretized) being tested against the other input levels. This way the parameter space will be effectively mapped by comparing each level of parameter across the range of parameters. The parameter factors are ranged and discretized to reflect values that might typically be considered in a modelling exercise. This allows a greater freedom in selecting and creating the parameter space, but at the loss of providing a more detailed analysis, which methods such as the GLUE approach and direct probabilistic methods may use. This, however, has two advantages; first that precise nature of the distribution of the inputs is not required, which in the case of the level of physical representation would be difficult to define. The results of this research may help to define this. Secondly, the methodology can reflect the uncertainty of the inputs (based on the range of the input), whilst maintaining a flexible approach to undertaking the analysis for three test cases. The assumption

here is that taking into account the full range of the factor is enough to capture the uncertainty of the input. The results of this research should be considered in relation to this sampling approach. The models are then assessed using various objective functions, and assessed with exposure-based methods. The exposure methods include regional analysis, subjective vulnerability functions and damage cost based on Depth damage curves, which consider relative regional risk and the impact on global model results. This investigated the significance of the objective function in determining influential factors, and investigates the ability for exposure-based methods for communicating uncertainty.

The second stage will involve quantifying the importance of factors in influencing model results by using sensitivity analysis, which was undertaken using screening methods to provide a qualitative result and quantified using Sobol Indices. However in order to achieve a convergence of results for a case with 5 parameters would require over 1200 model realisations for a single test case, which with the computational cost of the ATS solver would be significant. In order to reduce this cost, a Gaussian emulator approach was adopted.

The main techniques and methods are outlined in the following section. Each test case will have specific elements in either the analysis techniques or parameters that will be explained in greater detail in the introduction to the data chapters.

3.1. Input factors

For each test case, the same principal factors are used in the parameter space. Friction values, hydrograph inputs, cell size, the root means square error (RMSE) value of the Digital Elevation Model (DEM) and building representation type are all considered. These inputs represent the most significant sources of uncertainty in inundation modelling, such as the hydrograph and friction value, or values that are typically considered in creating a model such as the cell size. The values typically used are set in Table 3.1. Each parameter is divided into discrete intervals, and described from that

<table>
<tr><th>Parameter Type</th><th>Range</th><th colspan="3">Levels</th><th>Notes</th></tr>
<tr><td>Cell Size</td><td>2m – 4m for urban tests, 20m-40m</td><td colspan="3">2 levels</td><td>Based on work by Fewtrell et al (2008)</td></tr>
<tr><td>Hydrograph</td><td>20% of calculated hydrograph</td><td colspan="3">5 levels (-20%,-10%,0,+10%,+20%)</td><td>Value based on Di Baldassarre et al (2012)</td></tr>
<tr><td>DEM error</td><td>0cm - 15cm</td><td colspan="3">2 levels - Original surface and 15cm degraded</td><td>LiDAR vertical RMSE</td></tr>
<tr><td>Building Representation</td><td>BR, BP, BB</td><td colspan="3">3 levels</td><td>Based on Schubert and Sanders (2012)</td></tr>
<tr><td rowspan="2">Friction Value</td><td rowspan="2">Low Friction - 0.008-0.020 (13 levels at 0.01 difference)

High Friction 0.015-0.075 (13 levels at 0.05 difference)</td><td>Number</td><td>Low</td><td>High</td><td rowspan="2">Manning’s n for low friction
Manning’s n for high friction
Uniform friction values</td></tr>
<tr><td>1.</td><td>0.008</td><td>0.015</td></tr>
<tr><td></td><td>2.</td><td>0.009</td><td>0.02</td><td></td></tr>
<tr><td></td><td>3.</td><td>0.010</td><td>0.025</td><td></td></tr>
<tr><td></td><td>4.</td><td>0.011</td><td>0.03</td><td></td></tr>
<tr><td></td><td>5.</td><td>0.012</td><td>0.035</td><td></td></tr>
<tr><td></td><td>6.</td><td>0.013</td><td>0.04</td><td></td></tr>
<tr><td></td><td>7.</td><td>0.014</td><td>0.045</td><td></td></tr>
<tr><td></td><td>8.</td><td>0.015</td><td>0.05</td><td></td></tr>
<tr><td></td><td>9.</td><td>0.016</td><td>0.055</td><td></td></tr>
<tr><td></td><td>10.</td><td>0.017</td><td>0.06</td><td></td></tr>
<tr><td></td><td>11.</td><td>0.018</td><td>0.065</td><td></td></tr>
<tr><td></td><td>12.</td><td>0.019</td><td>0.07</td><td></td></tr>
<tr><td></td><td>13.</td><td>0.02</td><td>0.075</td><td></td></tr>
<tr><td></td><td>Single Value 0.01-0.07</td><td colspan="3">7 Levels for a uniform friction value surface.</td><td></td></tr>
</table>

Table 3.1: Typical Parameter Space for the test cases. Each parameter is divided into discrete intervals (levels) which forms the parameter test space.

Cell Size; The value for cell size is determined by two factors, minimum feature representation and computational runtime. For the urban areas, this is required to be at least 2 meters. This is based on previous studies that have identified 2m as being sufficient to represent main flow paths in urban areas (Fewtrell 2008). With gaps between buildings being approximately 10m for Glasgow and 8m for Coventry, a minimum 2m resolution is more than sufficient to represent the main flow paths in the urban areas. A practical consideration is also taken, given that the time stepping procedure of the ATS code can significantly increase the runtime of a model by an order of magnitude compared to more complex models, a 1m

resolution would be restrictive for urban areas, even for the Glasgow test case where the grid domain is around 1km². The Mexborough case was modelled at 20m and 40m due to the rural nature of the model domain. It also represents a trade off between computational costs and topographic representation. Whilst building size is less of an issue in determining the cell size, for this case the river, which is around 40m in width, must be represented in the topographic domain.

Hydrographic input was varied between a range of ±20% of the estimated value (Domeneghetti et al 2012). This value represents an approximate value associated with the combined uncertainties of estimating input boundary conditions. Whilst this estimation does not take into account precise details about the measurement procedure, it is felt to be sufficiently wide to cover the range of potential approaches, including the use of rating curves, power laws and modelling assumptions that will be used to estimate inflow.

DEM Error. The DEM used in the models are all based on LiDAR derived data sets, which have a RMSE of the relative true ground rating of ±15cm for elevation and 5cm for lateral position (Cobby et al 2003). This is incorporated into the input parameter space in a similar way to the approach of Hunter et al (2008), by degrading the original DEM according to the RMSE value for height and lateral position. As knowledge of the precise nature of the true ground value is unknown, the degrading process will be random, based on the range of the two RMSE values. A two level factor is created, with the original and degraded surfaces representing the range of error associated with LiDAR data.

Building Representation is a key aspect of flood modelling in urban areas. The choice of approach may significantly impact both the model results and the run time (Sanders et al 2008). In these test cases, 3 approaches are used to represent buildings, based on the approaches first suggested in Sanders et al (2008), Building block method, (BB), Building resistance method (BR), and Building porosity method (BP). The position of the building is determined using OS Landline data. For the building block method, the elevation is raised by 6m for all cells that fall within the footprint of the building. The building resistance method contains no physical representation but a high friction value for all cells in the building footprint, to allow for some representation of the flow through the building. The Manning's coefficient is fixed at 0.1 in all test cases, which is high enough to create a significant reduction in water discharge levels through the building footprints. The Building Porosity method uses a fixed conveyance porosity value of 0.5 for cells in building footprints for all test cases. This value represents the amount of cell space available to be

occupied by the water, as well as the term used to reduce the intercell discharge volume. Additional building representation techniques such as the Building Hole method (BH, where the building is removed from the DEM) are not considered in the first part of the test case, as the modelled effect would be similar to the building block approach, but is incorporated into the Sensitivity Analysis section.

The Manning's roughness coefficient is determined by the underlying land coverage. Two approaches are taken to represent the friction value. First, a single value for the domain is used, from 0.01 to 0.07 at 0.01 intervals. Secondly, the value is spatially distributed based on OS Landline vector data, with roads, paths and other low friction surfaces being assigned values based on 0.008 to 0.02 at 0.001 intervals, and higher friction surfaces assigned values from 0.015 to 0.075 at 0.005 intervals. The larger parameter space, compared to the other inputs is justified, as the values represent the range of values that could be used in a modelling exercise. It also reflects that this value will most typically be considered in a modelling exercise, and has a large uncertainty associated with determining the appropriate value. In order to reduce the computational cost of the test, the spatial distributed friction values were not cross compared. This would create a model sample of 700 realisations comparing only spatially distributed friction values and level of physical representation. A more detailed analysis of friction uncertainty relevant to modelling exercises can be achieved by comparing spatially distributed values versus uniformly distributed values. For each test case, a uniform Manning's value was chosen from the range 0.01 to 0.07 at 0.01 intervals. This range represent a wide range of surfaces as well as providing comments to the questions; are spatially distributed friction value significant in model results compared to a uniform friction value, given the uncertainty associated with this value is both a temporal and spatial value (Hall et al 2005)? A potential issue on this research is that only two surfaces types are used to create spatially distributed values. A more detailed approach, providing several surfaces with appropriate friction surfaces may be required to investigate this fully.

This systematic approach of cross comparison of parameters at discrete levels provides a manageable approach to investigating uncertainty from a number of different sources. Other approaches, such as GLUE, may provide more detailed analysis. A more stringent statistical approach, which uses knowledge of the probability distribution functions to sample the parameter space, would provide a greater insight into the relative performance of each factor, such as the GLUE approach (Beven and Bingley 1992). Here, sampling from the extreme edges of each of the parameter range to define the factor space of each test case may over

estimate the uncertainty. However a probabilistic approach may lack the control of data that a systematic approach allows, where by each model run can be attributed to its original set of parameters, and allow more detailed analysis of model results.

3.2. Objective Functions and Model Evaluation

Model realisations were assessed in two ways; a standard approach which involves assessing direct model outputs such as water depth, and a exposure based approach which will involve using vulnerability, hazard and financial risk as model evaluation techniques (essentially methods that require secondary models or further analysis). The standard approach compares in two ways. Firstly, the model outputs directly, including water depths, velocities and time series profiles for locations within the model domain. This allows a direct and detailed comparison of variations between factors on model results. Secondly, it uses the objective functions to evaluate the models. Each test case will use the $F^{\langle 2 \rangle}$ evaluation technique (Bates and De Roo 2000), but each test case will use additional functions that are data dependent to provide further insight into evaluation techniques. For the Glasgow test case, a Nash Sutcliffe coefficient is used to evaluate the model, for Coventry a depth measurement is used, and for Mexborough a comparison of observed depths.

The problems of using the $F^{\langle 2 \rangle}$ function are well noted (Beven 2010, Stephens 2013). The $F^{\langle 2 \rangle}$ statistic is a binary performance measure that compares the observed extent with the modelled extent, and evaluates it according to equation 3.1

$$F^{\langle 2 \rangle} = \frac{\sum P_i^{D_1 M_1}}{\sum P_i^{D_1 M_1} + \sum P_i^{D_1 M_0} + \sum P_i^{D_0 M_1}} \qquad (3.1)$$

Where $\sum P_i^{D_1 M_1}$ represents the sum of pixels flooded in both modelled and observed data, $\sum P_i^{D_1 M_0}$ represent cells that are flooded in observed and not in the model where as $\sum P_i^{D_0 M_1}$ represents the opposite group of cells. The function ranges from 0 to 1, where a value of 1 represents a perfect fit between observed extent and modelled extent. This measure has been used extensively (Bates and De Roo 2000, Aronica et al 2002, Hall et al 2005,Schumann et al 2011, Prestininzi et al 2011), and provides a useful indicator of global model performance. The drawbacks of this approach include poor differentiation between models in steep sided topography, where significant increases in water elevation do not correspond to significant increases in extent. This is an issue in valley filling exercises. A number

of approaches have focused on extracting water depths from the extent to evaluate model performances (Mason et al 2009).

Each test case also contains a second evaluation method with which to evaluate how the significance of a parameter changes according to the method used to evaluate simulation outputs. For the Glasgow test case, a comparison of water depths over the time of the model domain is made using the Nash Sutcliffe coefficient. This method is regularly used to evaluate model outputs, and provides a similar evaluation value as the comparison of extent function. The coefficient is defined as;

$$NS = 1 - \frac{\Sigma_{t=1}^{T}(Q_o^t - Q_m^t)^2}{\Sigma_{t=1}^{T}(Q_o^t - \bar{Q}_o)^2} \quad (3.2)$$

Where Q_o^t is the observed discharge value at time t, and Q_m^t is the modelled discharge at the same time. $\bar{Q}_o$ is the mean observed depth. The coefficient ranges between 1 and -∞, where 1 represents a perfect fit. The method has been noted to fit peak values better than the overall model range, but still represents the most common evaluation of discharge/time series.

For the Coventry test case, an observed depth was used to compare model outputs. This was evaluated simply as the difference between the model and observed. An absolute value was not used – this allows insight into the variations attributed to the different inputs.

The second objective function used for the Mexborough test case is a root mean square error of the difference between observed water depths and modelled water depths. This method has been used in Stephens et al (2012) as an evaluation technique that provides greater detail into model output than an evaluation of extent. The method requires using The RMSE is defined as;

$$RMSE = \sqrt{\left(\frac{x_o^{i,j} - x_m^{i,j}}{n}\right)^2} \quad (3.3)$$

Where $x_o^{i,j}$ is the observed water depth in cell i,j and the $x_m^{i,j}$ is the modelled water depth in the same cell. A value of zero would represent a model which would be close to replicating the observed data exactly (there could be the potential for identical errors to cancel out over the model domain). Therefore lower RMSE values represent better performing models.

In addition to objective functions, a number of uncertainty methods are also considered to evaluate the models. Uncertainty extent plots (also refered to in the

literature as probabilistic extent plots), such as those used by Aronica 2002 and Hall 2005, provide important data about model behaviour over a parameter space. In this approach the value of a pixel is determined by the frequency of inundation occurrence across the parameter space, with a range of 0 to 1, where 1 is flooded in all model realisations. By plotting this across the model domain, these figures can demonstrate where the modelled extent is most uncertain. It should be noted that the output is better described as an uncertainty value due to the sampling nature of the inputs, and because the value represents the frequency of the pixel flooding across the model ensemble. The output is described as an uncertainty flood extent, with a measure of inundation frequency displayed (i.e. a value of between 0 where no flooding has occurred and 1 where the pixel has flooded in all test cases) for each pixel displayed. Further model evaluation will be using control points to compare, water depth, velocities and arrival times over the course of each model run.

3.3. Risk and Vulnerability

Several methods that represent the range of potential approaches to determining risk from model outputs are considered, in order to evaluate models according to exposure to risk. The first approach is based on the determining the financial consequence of floods using Depth damage functions and the approach described in the Multi Coloured Manual (MCM) (Penning-Rowsell et al 2010). This is undertaken by first determining the classification of building type based on OS data sets. Buildings are divided into groups based on this classification into residential units, light and heavy industrial units, workshops, schools, pubs, shops and local services (including libraries). Each category has a corresponding depth/damage cost curve, to determine the cost per square meter in relation to the depth of water in cells located near the buildings. A buffer around the perimeter of the cell defines the points where the cell depth value is used to evaluate the cost per square meter depth damage functions. This approach is used to allow a comparison between the different building representation methods, and ensure that the same cells are used for all model runs. This represents a subjective component of this approach and a potential source of uncertainty. Each method may cause a variation in how water accumulates against the edge of the building, which may impact results. The approach is hereafter referred to as the Cost of Damage method.

In order to derive a Cost of Damage, assumptions must be made about building type and susceptibility to flooding in order to identify the correct depth damage functions to apply. Here the assumption used is that the properties represent a

mean susceptibility range (i.e. the building do not fall in to extreme values of vulnerability – neither high or low susceptibility). Consequently there is a certain amount of uncertainty in this analysis, based on these assumptions. this can be avoided by using the range of potential depth damage curves for a building as a reflection of the uncertainty and incorporating it into the overall analysis. Due to computational cost this has not been considered but could potentially be included in future studies to identify how critical this choice is relative to other sources of uncertainty in model evaluation techniques and inputs. The depth damage curve was also altered to 2013 levels by applying the CPI index to the 2010 values supplied in the Monte Carlo Manual. For each building type, the depth damage curve represented the total economic loss to a property. Each test case used a short duration damage curve (less than 24 hour inundation duration), which may not hold for the Mexborough test case, but ensures consistency between each test case.

Subjective vulnerability approaches were suggested by Pappenberger et al (2008) as a means of identifying individual areas of interests and weighting model results towards these areas. The two approaches suggested in that research were used here and using those methods allows a comparison of qualitative and quantitative approaches to be made, if the Cost of Damage method is considered a quantitative approach. The first approach is a localised $F^{\langle 2 \rangle}$ method using the same approach as discussed above, but focusing on smaller sub domains within the model. The sub domains were determined by considering the total model domain, and dividing it into regions of economic importance, or vulnerable peoples, based on economic or census data. This approach allows the classic $F^{\langle 2 \rangle}$ approach to be considered in the context of important assets and removes the bias typically associated with that approach, in particular performance in steep terrain, and focuses it on areas of concern. This method therefore represents a standard function approach which is enhanced by knowledge of risk in the model domain.

The second approach is a weighted cell methodology. Regions in the model are assigned a risk value between 0 and 1, where 1 represents assets of greater value, depending on unique characteristics determined before the modelling process. The weighting is applied to the $F^{\langle 2 \rangle}$ function, and model results are then evaluated to the model function $F^{\langle 2 \rangle}VW$. This is similar to the $F^{\langle 3 \rangle}$ approach of Prestininzi et al (2011), but with the weighting being dependent upon subjective vulnerability values. The weights were spatially distributed across the model domain into regions surrounding the areas determined as vulnerable. The precise distribution is unique to each model domain, depending on the case, and is described in more detail in

the test case chapters. The previous use of this method concluded that it represented a robust method for assessing models and for determining the risk to peoples and buildings. By using both methods insight will be gained about the best approach to evaluate vulnerability using model results.

A spatial probabilistic approach to hazard mapping, similar to the probabilistic extent method suggested above and utilised by Aronica et al (2012) is also used. The method uses velocity data and water depth data for individual cells to determine a hazard value. The hazard value for the cell is then divided into 4 categories which are then assessed in the same way that the frequency of cell inundation is assessed above (i.e. a value of 1 represents cells that have flooded in all model realisations). This approach is modified here to account for the large number of models. First, the maximum depth and velocities of each is cell is determined. The hazard value is then determined from these factors, from which a hazard value based on Defra flood risk to people hazard level is used, as in Equation 2.31 and 3.4 (Defra 2010).

$$Flood\ Hazard = h * (v + 0.5) + df \qquad (3.4)$$

All cells with hazard values above 0.75 are then determined as hazardous. This value is the lowest threshold value in the Defra guidance, above which it represents a hazard to people. This assessment is undertaken for all model results individual before it is combined to create the final probabilistic outline. The method used here is conservative in comparison to the Aronica et al (2012) method. The inclusion of all hazard cells removes detail from the analysis and using maximum values of velocity and depth rather that the combined maximum means that the precise hydraulic nature of the flood is not captured. This however, represents a compromise between computational cost and analysis. By assessing risk in this way, velocity is included which plays a key factor in determining the risk to humans, as well as buildings. Moreover, it also allows insight into how the levels of physical representation define flow paths and critical flow regions in the model, as well as insight into the hydraulic outputs of the models. Furthermore, assessing this hazard uncertainty spatially allows regions of uncertainty to be identified as well as identifying where the highest level of hazard is likely to be determined. As with the extent outline, the term uncertainty is more appropriate to describe the output, with a measure of hazard frequency of 0 to 1 displayed, similar in concept to the inundation frequency introduced earlier. This is also similar to the probabilistic approaches that have been used to describe this output in other studies (Aronica et al 2012), although the terminology is altered here to reflect the difference in approach to the parameter sampling.

3.3. Sensitivity Analysis

In order to quantify the impact of factors, Global Sensitivity Analysis (GSA) is undertaken on model results. Two approaches were used, a screening method and a variance based method, which uses an emulator to replicate the model response surface, and then evaluating first order effects and total effects of key factors based on the emulator's response surface. The approach provided an insight into key factors in flood inundation uncertainty, whilst reducing the level of computational cost associated with this type of analysis.

3.3.1. Morris Method

The use of GSA methods is restricted by its computational cost. To compute the full range of Sobol Indices requires *N(d+2)* model runs, where *d* is the number of factors to be considered, and *N* represents the model runs determined by a convergence factor, where the parameter space is sufficiently covered to represent all parameter sets to create a model evaluation function. In large parameter test cases the number of model runs required can quickly become intractable. Even when the number of parameters is low, where complex computational models are used, such as environmental models, the computational effort may still be significant. In these scenarios the use of screening methods may still provide an insight into parameter behaviour. A popular, simple method to implicate is the 'One At a Time' (OAT) methodology, devised by Morris (1991), which is extremely efficient in terms of computational effort and is easily implemented as a result of non-restrictive assumptions. The Morris method can determine which factors are negligible, linear or additive or non linear and interactive, and was used in this research. The main methodology, described by Saltelli et al (2000) is to divide the *k* dimensional input vector **X** into discrete levels (where *k* is the number of parameters) which can assume integer values based on the set {0,1/(*p*-1),2/(*p*-1),...,1}, where *p* is the number of levels in the design space.

Using a random starting point from the vector **X**, each factor is sampled from its range of values, based on the number of levels in the case. This model is realised to create a base value. From this, one of the factors is then evolved by Δ, a predetermined multiple of 1/(p-1), that remains within the sampling region.

The elementary effect describes the impact of varying a factor on the previous model realisation or base value, and is defined as

$$d_i(X) = \frac{y(x_1,\ldots..x_{i-1},x_i+\Delta,x_{i+1},\ldots..x_k)-y(X)}{\Delta} \qquad (3.5)$$

Where Δ is a predetermined variable of *p.* **X** is a selected value in the region of experimentation, and the effect of delta describes the variation of factorx_i. This is repeated until all factors have been sampled, from the initial base value, to create the first trajectory. A new base value is sampled and realised to create a new trajectory through the modelling space. The number of trajectories required to sample the model space sufficiently is dependent on the number of model factors.

The mean, μ, and the standard deviation, σ, of the element effect of each factor over the total iterations provide the outputs. A large mean indicates a factor with significant first order influence. A high standard deviation indicates that the model has a high interactive and non linear influence on model results.

The efficiency of the method is that each parameter will only require one model run per trajectory. Previous studies have indicated that for 5 parameters, approximately 10 trajectories will be sufficient to determine the mean and standard deviations. In this research, this will be explored by using twenty trajectories for each test case to test for result convergence. In this research, the factors will be divided into 4 discrete levels, based on the range values described in table 3.1, producing an input space of 6 factors at 4 levels, which includes the level of physical representation. The number of levels has been determined in order to include these factors, and to include the building representation factor as well. The assumption is that the number of levels sufficiently covers the parameter space, and provides an adequate description of the inputs.

This method can only provide qualitative indicators of factor sensitivity, unlike variance based methods. It also requires parameter independence, which is assumed for the parameters used here. There are also indications that where the model parameter space is small the use of a screening method does not provide a robust measure of sensitivity (Pappenberger 2006). This however is countered by the fact that it requires significantly fewer model runs, and can still identify factors with significant contributions to model variations (Herman et al 2013).

3.3.2. BACCO Sensitivity Analysis

In order to provide quantitative results, moment based sensitivity analysis, such as variance decomposition, is required. Although the parameter space is small in this research, the computational cost associated with the use of flood inundation models means using a standard sampling approach for a GSA method, such as Monte Carlo sampling, may become prohibitive when dealing with the numerous runs associated with variance based methods. In this scenario, emulators provide a statistical approximation to the model that can be further assessed with GSA

methods, to provide the same output that a standard SA approach would take. The use of an emulator in a calibration study has been undertaken for a flood inundation model (Hall and Manning 2011).

The BACCO (Bayesian Analysis of Computer Code Output) methodology used here is based on the approach developed by Kennedy and O'Hagan (2001) and Oakley and O'Hagan (2002), in which an emulator is created based on Gaussian principles. The emulator is a statistical approximation to the computational mode, and is designed on training runs of the original model at a number of locations within the parameter space. From these training runs, the emulator is then constructed as below;

$$\eta(v,t) \sim \mathrm{N}\{m_t(v,t), c_n[(v,t),(v',t')]\} \qquad (3.6)$$

Where $\eta(v,t)$ is the response of the computational code at points v,t, where the training points have been conducted, $m_t(v,t)$ is the emulator mean, and $c_n[(v,t),(v',t')]$ the emulator variance, which are both solved numerically within the Gaussian framework.

Here, the assumption is that the model code can be represented as a random function that can be described in terms of normal distribution. The emulator is formally Gaussian because the marginal distribution of all factors, including the model function of the computer code, is multivariate normal. The advantage of a Gaussian process is replication of the training points in the emulator surface, ensuring that model uncertainty is reduced where sufficient knowledge of model performance is known. The process provides the mean and variance for points of the emulator away from the training assuming the principles of smoothly varying functions are maintained from training point to training point. This allows the uncertainty related to the emulator to also be defined. In general the uncertainty related to this can be reduced by improving the coverage of the parameters and increasing the number of training points. The sensitivity analysis undertaken is a variance based methodology. Broadly, to determine the first order impact of a parameter the variance of the output is decomposed to individual integrals and evaluated by:

$$n^2 = \frac{Var_{\boldsymbol{x}}[E(Y|\boldsymbol{x}]}{Var[Y]} \qquad (3.7)$$

Where $Var_{\boldsymbol{x}}$ is the variance conditional expectation of y dependent on $\boldsymbol{x}$, where $\boldsymbol{x}$ is a subset of the total parameter region used to calculate $Var[Y]$. The measure is therefore the contribution of the parameter where 0 is no contribution, 1 is responsible for full impact on model results. In order to estimate the impact of a

single parameter or factor, X is varied, whilst the other factors are fixed. In order to examine the total impact of a factor, X is fixed, whilst the other factors are varied. This methodology explores the contribution of the factor to the total output.

In this research, the GEM-SA software developed by Kennedy and O'Hagan (2001) was used to undertake the construction of the emulator and the variance based SA, using the principle described above. The input factors were based on the parameters outlined in Table 1, although both the building representation factor and level of physical representation were removed, as defining it in terms of a pdf function is difficult, as the factor represents only a 4 level input. Instead a comparison between a simplified approach (ACC) and a full SWE model (Roe) was used to compared the relative contribution of the remaining inputs from the parameter space to model output was undertaken. The implications of this are that only the relative effect of the input can be explored, and the impact of changing physical representation on the importance of inputs can be explored. A series of initial runs revealed that a LP-Tau sampling process (a quasi random sampling that covers the entire parameter range in uniform units, similar to a Sobol sample) provided a smoother and more accurate emulator approximation of the computational code, than a Latin Hypercube MC sampling procedure. The initial runs also showed a good approximation between the calculated variance and the observed variance based on the training runs. This is essential in determining the appropriateness of the method, as well as giving greater confidence in the results. For the Morris method, all the standard objective functions were used. However, for the BACCO GEM approach, only the $F^{\langle 2 \rangle}$ described above will be used to explore how parameter importance changes relative to the evaluation criteria. This was due to finding the $F^{\langle 2 \rangle}$ method a sufficiently smooth function to construct the emulator. It also allows a comparison between locations, an approach not possible with the other evaluation functions. The exposure based methods also compromised the smoothness of the emulator output.

3.4. Test Case Model Setup

For each test case a series of initial model runs were undertaken to determine which parameters to include in the input parameter space, which aspects of the model domain would require adjusting and to determine model run times and model convergence.

Initially, models were run with fixed parameter values to determine model run times, model stability and convergence, evaluation of the DTM, and to provide additional

data about the flooding process and hydraulic conditions which could be used to define an appropriate parameter space.

Model run times were critical in determining the number of parameters that could be considered. The restrictive nature of the ATS time step (Equation 2.14) meant that this would be a key factor in determining the parameter space. An indicative example of the model runtimes is provided in Table 3.2, which provides an insight into model runtimes across the whole model ensemble. All test cases were undertaken on the Leeds ARC1 – HPC system which is a Linux operating system operating on two quad core 2.8Ghz Intel Zeon processors with 12GB of memory.

Test Case	ATS	ACC	Rusanov	Roe
Glasgow	700mins	>1min	>1min	>1min
Coventry	890mins	3mins	10min	7min
Mexborough	228mins	4min	12mins	12mins

Table 3.2: Indicative model runtimes for each LISFLOOD module across the parameter space.

A restriction of this system was a limit of 48 hours computation time. In order to ensure each test case could be computed with limited user operation, model runtime was restricted to fall within this limit. This was important for both the

The mass balance of each test was also investigated from these initial runs. Mass balance was found to be with 5% error for the Roe and Rusanov solvers, whilst ATS and ACC were near 0% error. This was assumed to be an acceptable error range for the models. It is important to note in later results that this may impact the overall model results, but here it is assumed that this is low enough not to have a global impact (i.e. effecting model output and function types that are related to model domain wide aspects, such as F^2).

The initial tests also allowed convergence to be determined, and appropriate time stepping conditions to be determined for the LISFLOOD modules. It was determined that the Rusanov solver required a restrictive CFL condition to ensure stability for all test cases(>0.1). The Roe solver also used a low CFL condition (0.2) for all test cases. The ACC module used a fixed theta value of 0.8. This ensured stability across the whole friction range, although the sensitivity of this module to this value has been noted (de Alemida 2012). Essentially, these limits were chosen to ensure model completion. Optimisation of the time stepping was not undertaken and these values were fixed early on to allow for the large model ensemble to be completed. Focus on the model runtimes has also be excluded from discussion, as it is assumed that other numerical solutions to the level of physical representation

may be more effective. It also ensures that focus remains on the level of physical representation rather than on numerical solutions.

Evaluation of the DTM was undertaken to ensure that no unusual errors occurred in the DTM which may affect flowpaths. This was particularly critical for the Coventry test case, which was based on a 2008 DTM, but had to be adjusted to represent the 1978 topography. In addition to this a number of other factors were considered, such as boundary conditions and topographical representation. A pragmatic choice was made between how critical the results were to model results and the increased computational cost. For example, in the Glasgow test case the use of downstream boundary conditions was explored. This impact, whilst notable, was essentially local however, and was fixed for the test case model realisations.

From these results a number of parameters were either carried over, or fixed for the test case evaluation. For the urban test cases, an identical parameter space was chosen. This allowed a certain amount of direct comparison between the two case studies despite the different hydraulic conditions of the test. The Mexborough test case was originally to be built around the use of 1D channel models but consistent mass balance errors and instability with the full SWE models prevented this.

For the urban test cases, the complete model ensemble was 4800 model runs including all 4 LISFLOOD models. For Mexborough, this was reduced based on the longer model runtime to 2400 model runs. Again the time taken to run the ATS module proved critical to the number of parameters that could be evaluated. Each model realisation was then evaluated with the objective functions, which allowed for a quality control of the realisations to be undertaken based on the model results. Individual model runs could then be evaluated in more detail to determine cause of unusual results. In this case either additional runs could be performed or the results could be noted within the analysis.

A similar process was undertaken for the sensitivity analysis approach, with the focus being on convergence in the model results. This was achieved with a relatively low number for the Morris test (around 30 realisations) which fits in well with previous considerations of this approach (Saltelli et al 2000)

4. GLASGOW TEST CASE

This test case is a culvert blocking event that occurred in the Greenfields region of Glasgow, a highly developed urban environment consisting predominantly of residential buildings, with small businesses and a school also present. This test case has been used extensively in research; Hunter et al used it as the basis of their benchmarking case study (2008); Fewtrell et al used it to explore the impacts of cell size and grid meshing on model results (2008), Schubert et al used it to test building representation (2008) and Aronica et al used it to explore rainfall runoff uncertainty and probabilistic hazard mapping (2012). As outlined in the benchmarking paper, the nature of the topography, inflow boundary conditions and the low friction surface create transcritical flow conditions, with super-critical flow developing along the top of the main flow path in the early stages of the model. This, as well as the complex urban topography that includes roads, paths, and buildings, will provide a rigorous test of model capabilities. A test case with trans-critical flows will also provide insight into the levels of physical representation in the numerical scheme, as specific treatment or inclusion of hydraulic shocks may be essential in determining model output.

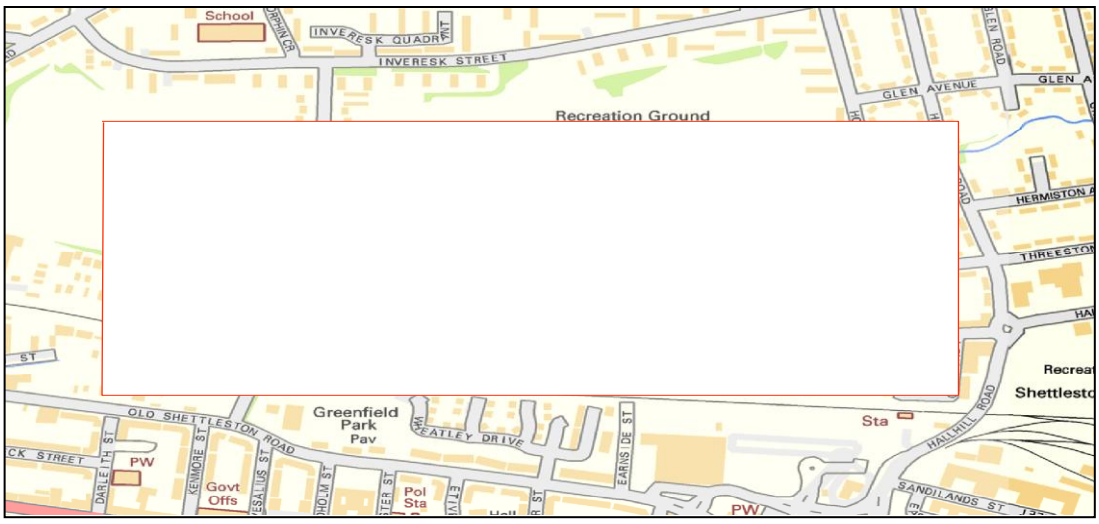

Figure 4.1: The location of the test case. The red box represents the model domain, the cross represents the culvert and the inflow point.

The modelled event is based on a observed event that occurred at the site. Flow from a small river that enters from the top right of the domain enters a culvert, identified by the red x in Figure 4.1. This culvert contains a trash screen, which when blocked leads to the flow exceeding the channel and over topping at this point. Flow moves from right to left across the model domain, in the general slope

direction leading to the formation of two main flow paths, the first along the main road, the second moves toward the bottom centre of the model domain. Given the high level of urban development and the presence of kerbs and roads, the flow paths are well defined and constrain the flow in several locations, including the main east west flow path, where water depths remain low. The flow paths converge at the far left side of the model, whilst water begins to pool at low elevation regions at the bottom centre. The main flow path runs down a steep slope (around 6% gradient), which creates super critical flow in this region. This represents part of the challenge of this test; the modelling of transcritical flows in low friction environments. The flow continues to be high velocity – low depth for the duration of the inflow hydrograph, as a result of the slope and the constrained flow path in this region.

In the absence of observed data for inundation extent or water depths, a high resolution benchmark model based on a LISFLOOD-Roe realisation provided proxy data to assess model output. The resolution was set at 10cm, with distributed surface friction values of 0.014 for roads and paths, and 0.045 for other areas. The hydrograph was based on the estimated hydrograph value from Hunter et al (2008). The output is then used as an observed data set with which to use the goodness of fit comparison method F^2 This approach has been used before (Fewtrell et al 2008), and whilst it would not be suitable in a validation study, it provides a means of describing model output that is common to other studies and allows the research to meet its objectives. Furthermore, as the spatial resolution tends to zero, the numerical model begins to converge on the original PDE equation system, and so has a level of robustness with which to justify this approach. An additional advantage is that the analysis can be extended to evaluate modelled data against ‘observed’ benchmark discharge data series, and use Nash Sutcliffe coefficient to evaluate the model. As with the analysis of extent, the purpose is to use a single global output of the model to evaluate in a wider analysis context rather than to validate the model.

The input factor space and files were then constructed from a number of data sets, as well as being defined by the main modelling characteristics of the site. The input factor space for this test is defined as in Table 4.1 in chapter 3. The DEM is taken from the Hunter Benchmark paper. This is based on a 1m resolution LiDAR derived Digital Surface Model (DSM), which is then combined with OS Master Map vector data, to define the location of the buildings, kerbs, which are set at 10cm raised elevation from the surface of the model, and roads. OS data is also used to

determine features to be considered in the building representation factor and the location of spatially distributed friction values, based on land surface codes. The building footprint location was used to construct building blocks by increasing the elevation of the cells inside the footprint by 6m to create a digital terrain model of the site (DTM). This data was also used to define the cells that would have increased spatial friction values, for the building resistance method (BR) or to represent porous media cells (BP). The surface was divided into roads and vegetation surfaces and was assigned the values in Table 4.1. The hydrograph was also based on the hydrograph determined in the Hunter et al paper which is an approximate outflow based on an event from July 2002 (Figure 4.3). Using an uncertainty range of 20% the input was sampled at the extremes of this range and at 10% intervals. Work by Aronica et al (2012) has confirmed that this shape is suitable for this test case.

Figure 4.2: DEM of the test case, with location of control points

A number of locations points are used to evaluate model outputs, and are displayed in Figure 4.2. These points are based on control points used in previous papers, and represent the key regions in the model that are affected by inundation and contain distinct hydraulic conditions relative to other regions in the model.

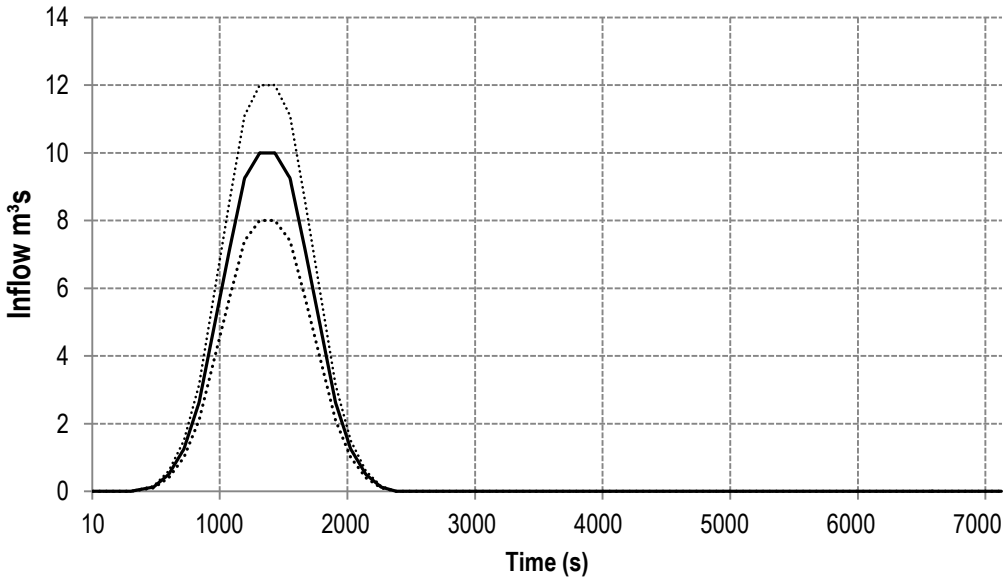

Figure 4.3: Inflow hydrograph for the Glasgow test case, with uncertainty bands of 20% (determined from Hunter al 2008).

For example points 1 and 7 are along the main flow path so allow for detailed analysis of model flow in this region, whist point 5 and 3 are storage areas, so allow for an analysis of water depths between the models.

4.1.1. Exposure Methods

In order to evaluate risk, OS data was used to determine the type of building from which a weighting value could be made and the most significant locations for regional analysis. The location of weighted cells and the values are displayed in Figure 4.4a. The values are based on a subjective range between 0 and 1, with higher weighting belonging to cells with higher vulnerability levels. The event took place during the night of the 30th July, therefore greater weighting is applied to the nearby residential buildings rather than the school as this can assumed to be empty at the time. However, given the importance of the school to overall region and the consequence of the building being damaged to the wider economy, it must still be weighted although not to the same level as the residential units. For this test case, the impact of the precise weighted value is not considered, although it could be considered in further test cases, Figure 4.4b shows the regions to be assessed in the regional F^2 calculation.

Figure 4.4: Part a, top, are the cells and weighting values for the vulnerability weighted approach (F^2VW) and part b, bottom, is the regions used in the Regional F^2

The four regions represent two areas of low quality housing (regions 2 and 3), region 1 represents higher quality housing, whilst region 4 contains the local primary school and other local amenities.

4.1.2. Sensitivity Analysis

In order to undertake the Sensitivity Analysis (SA) additional levels for some factors were developed which were not considered in the Uncertainty analysis section, for reasons of computational costs, but were included in the Morris analysis to ensure a 4 level space for each factor. In this case, the building hole (BH) method was added to building representation factor, where all elevation data is removed from the footprint of the building. The model realisations were then evaluated using Nash Sutcliffe, and the goodness of fit F^2.

Factor	Parameter Range(low to high)
Module	ATS-ACC-Rusanov-Roe
Building(type)	BH-BR-BP-BB
Friction(n)	0.015 0.075
Cell Size(m)	2m-4m-6m-8m
Hydrograph	As percent of original hydrograph -20%,-13.3%, 6.7%, 20%
DEM error	0cm- 5cm- 10cm- 15cm

Table 4.1: Factors to be used in the 4 level Morris method analysis

The BACCO sensitivity analysis provided further insight based on a smaller parameter space, and reduced. Issues concerning the use of this approach are covered in more detail in the SA analysis section of this chapter.

4.2. RESULTS

The results section is divided into 4 sub-sections. Firstly, an overview of direct model output, including analysis of depths at control points, and comparison of uncertainty extents for the 4 modules. Secondly, evaluation of the traditional model functions, F^2 and Nash Sutcliffe coefficient. Thirdly, an evaluation of the exposure based evaluation methods, before a section on the sensitivity analysis and conclusions.

Uncertainty flood extents were calculated for the 4 modules for the entire selection of model parameters (Figure 4.5).Each module shows a distinctive pattern which indicates that for this test case, the use of physical representation has an impact on the model results. Again there are similar patterns in these tests to the benchmarking paper of Hunter, in that the simpler diffusion based approach fails to overtop the blockage at the end of the main flow paths, unlike the momentum based modules. This is perhaps surprising given that some of the simulations would contain the extreme value for the hydrograph that may have been expected to influence the overall volume in the domain and consequently the extent. The relatively small blockage in this area (>10cm), may have been overcome by a simple increase in water depth. The momentum based modules manage to clear this blockage at the end of the main flow path, but the lower inundation frequency values in this region indicates the effects of other parameters in influencing the extent in this region. The uncertainty plot for the Roe solver (bottom right, Figure 4.5) shows higher inundation frequency values, and a wider extent suggesting that the increased representation in the Roe solver allows it to overcome the blockage

more consistently, and with a greater volume of water. A similar extent pattern can be observed between the Roe and Rusanov (bottom left, Figure 4.5) modules, although a higher inundation frequency value for the Rusanov solver is observed for the wider flood extent region, indicating a relative insensitivity to other parameters. Part of the reason for this wider extent relates to the use of Building representation methods, where the BR and BP method allow the code to create a wider flood extent. This is related to the simplicity of the Rusanov approach, which creates a faster moving body of water through the domain.

A key difference between the models is the flooding of the centre section of the flood domain. The ACC module (top right Figure 4.5) appears not to have flooded the centre section in any simulation, compared with the SWE modules. The SWE modules also have high inundation frequency values in this region indicating reduced uncertainty across the parameter space for this region. This difference is a product of higher velocities in the ACC module, which appears to prevent the water depth increasing in this area to the point that it can overcome the small obstacles and flow down the domain, from top to bottom. The ATS module (top left Figure 4.5) fails to overcome the urban obstacle and produces the narrowest extent. This is in keeping with the results from the Hunter et al benchmarking study. This variation provides insight into the complexity between the momentum terms and the influence of different aspects of the momentum flux in urban flooding. The high inundation frequency values for all uncertainty flood extent plots, would suggest that the other parameter factors are of low influence in comparison to the influence of the modules. However, a comparison of uncertainty plots for the ACC module taken over the 4m and 2m model domains (Figure 4.6) does suggest the importance of domain discretization in affecting model results. For the ACC module an increase in cell size appears to reduce the extent of the flooded domain to the same level as the ATS module. This is explained partly through the increased height in the blockage at this point, and highlights the potential sensitivity of reduce momentum based approaches to grid cell size and topographic representation. The implication of this is that there is a parameter interaction in the input space, where simplified physical models and the cell size of the domain is a critical aspect in the construction of an urban flood model.

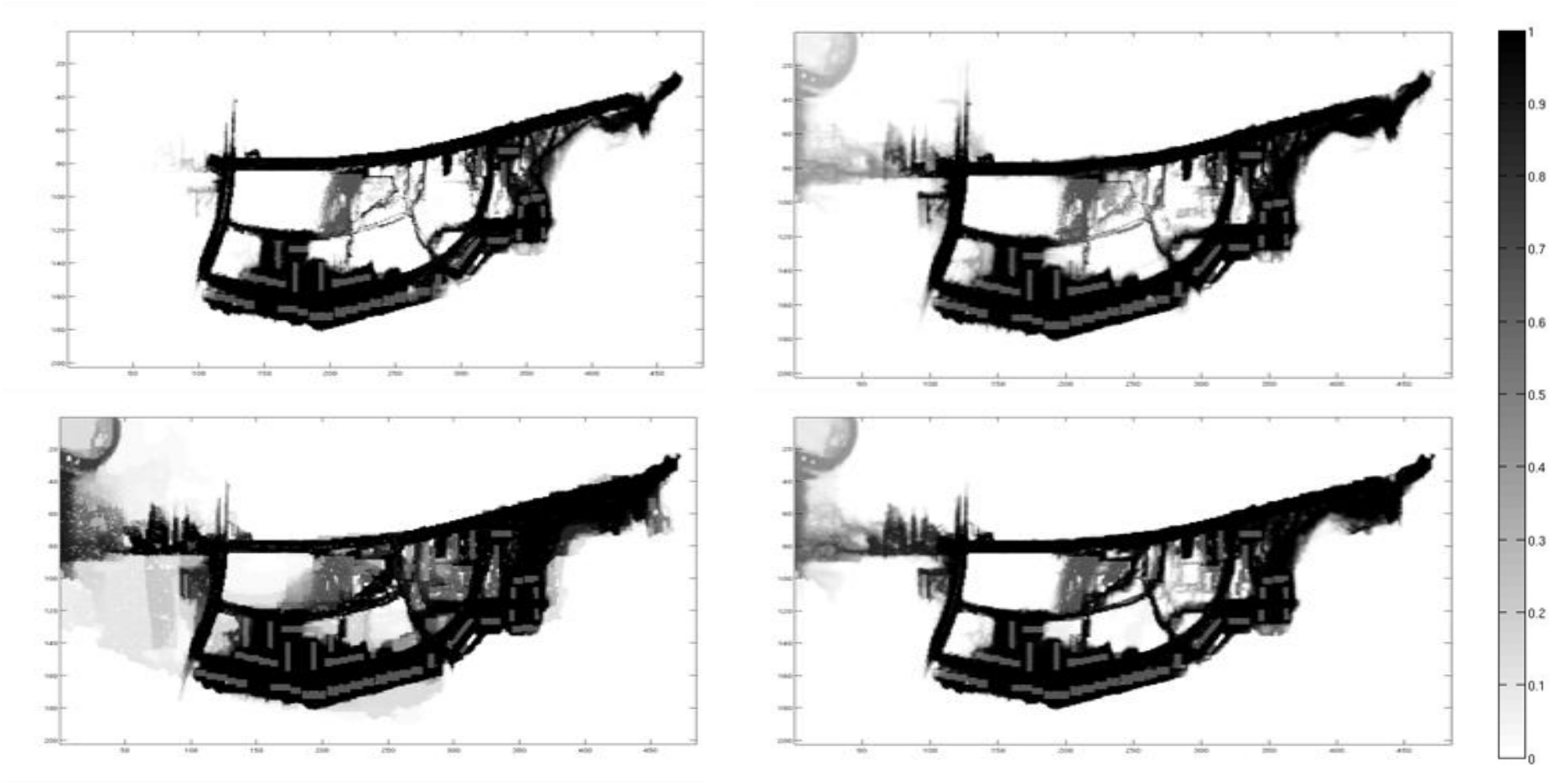

Figure 4.5: Uncertainty flood extent plots for the 4 LISFLOOD modules over the Monte Carlo results (top left ATS, top right ACC, bottom left Rusanov, bottom right Roe), where dark regions represent cells that have flooded in all model realisations.

Furthermore, the interior section of the model appears to be the most uncertain in terms of flooding for all modules. This can be attributed to the building representation techniques used. For all modules, the use of the building porosity and building resistance approaches creates a more spread out flow path, and allows it overcome some of the additional urban topographic features. This indicates that the choice of building technique may have a local impact on flow paths that may be critical in the analysis of vulnerability.

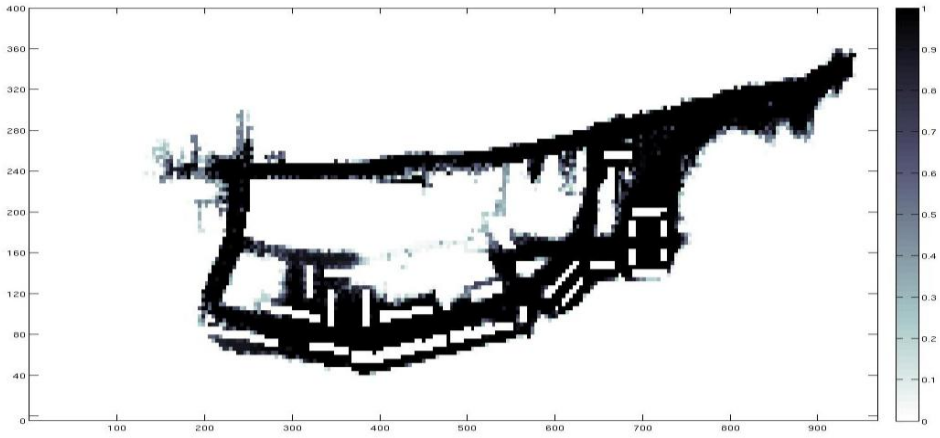

Figure 4.6: The uncertainty flood extent plot for the ACC module at 4meters

Further insight into module performance and parameter influence can be obtained by comparing the water depths between model realisations at the Control Points (Figure 4.2). Point 10 in Figure 4.2 represents the furthest extent of the main flow path, and provides insight into module performance for converging flow and transcritical flow. It also reflects another aspect critical in determining the consequence of the event, which is the time of initial flooding. In this case the depth is not of particular concern, as it is around 10cm, but the variations between model realisations and the sensitivity of the results, may prove important when taken forward into test cases with greater volumes of water. Point 10 represents the location where the flood water last reaches for all model runs, and is the furthest point from the source. A direct comparison of the water depths in all tests provides insight into variations from the modules, and also the influence of other factors.

Overall, the Rusanov module provides the earliest arrival of water at around the 900s point, and the ACC the highest peak of 0.723m. The ATS and ACC module produces a narrow range of final depths, which suggests insensitivity to other parameter factors. In comparison the Roe results displays large variations of water

depth. This highlights the point made in the comparison of flooded extent – that an increase in physical representation produces a marked change to module-parameter response. By comparison, the Rusanov module produces a very low depth for all model realisations, which relates to earlier findings concerning the Rusanov uncertainty flood extent, that the module will move water through the domain rapidly, to the point that it affects how water will pond within the domain. This highlights the potential drawbacks of the simplified numerical model, and points to the fact that where depths and inundation time are of particular importance to the final solution, simple numerical solutions to the full governing equations may prove to contain large uncertainties. The peak range values vary considerably within each module, which suggests that at this point, the choice of other factors is important. Indeed the similarity between high depth values for the Roe and ATS seem to suggest that the choice of model factor is of less concern than other parameter factors, and reflects the uncertainty in the analysis

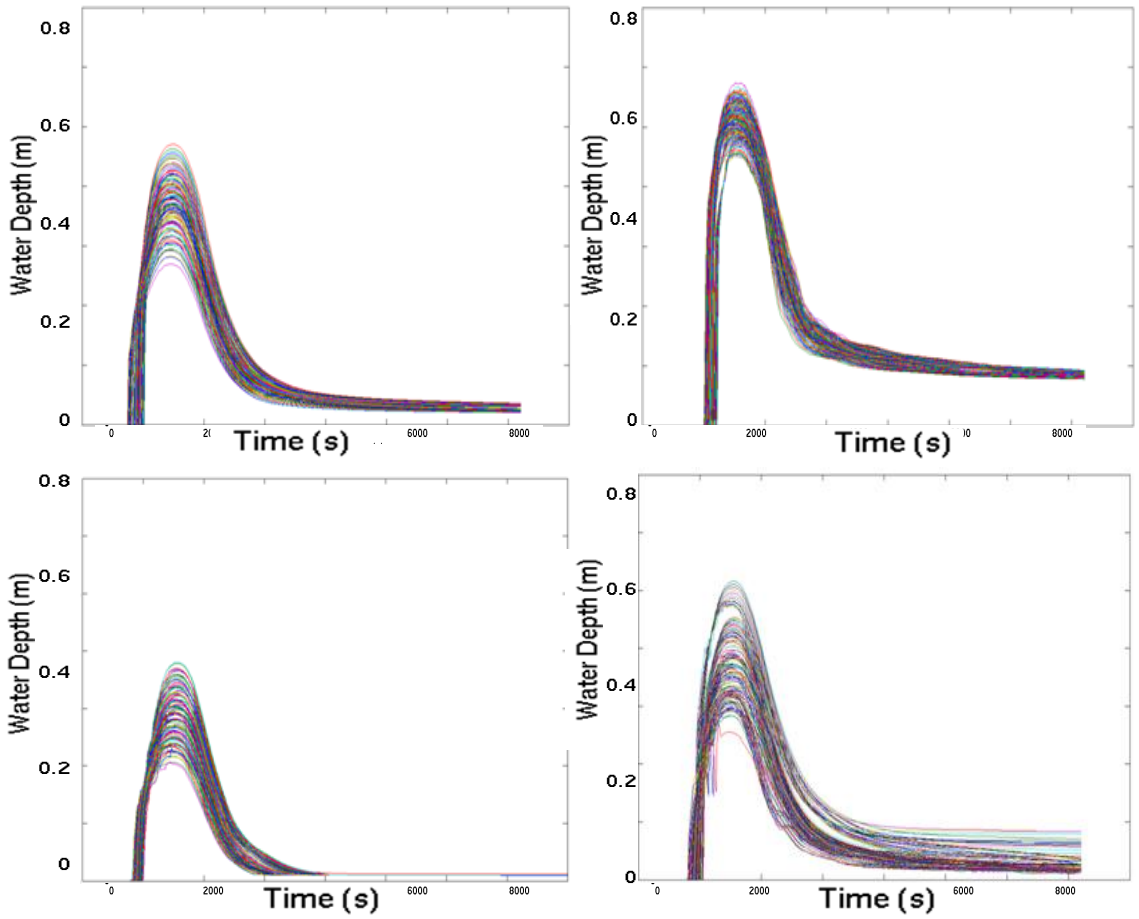

Figure 4.7: Water depths for the 4 modules over the range of tests at the control point 10(top left ATS, top right ACC, bottom left Rusanov, bottom right Roe), all plots scaled to 0 and 0.8m across the entire model run time.

The ACC module produces a later peak and a higher range of peak values, and a consistent, albeit higher, range of final water depth value. This is explained as a product of the higher water velocities down the main flow path, directing water away from this converging point. More of the water from the lower flow path finds its way to this section in the ACC module. As such there is less converging from two flow paths, and a greater quantity being provided from the lower paths. Analysis of the control points for regions that store water (CP-3), also reveal similar patterns. Arrival time is quickest for Rusanov, but a lower peak value is also recorded. For the other modules, the ATS produces the least varied range of depths (from 0.8m to 1.1m) for all depth cases, whilst the Roe solver and ACC contain greater uncertainty in the final level and time. The variation between these water levels would indicate the importance in choice of physical representation in modelling local flow characteristics, as well as final water depths, although the effect on global performance values would appear to be relatively small, based on the uncertainty plots of Figure 4.5.

4.2.1 Evaluation of Model Functions

Analysis of the measure of fit value, F^2 the Nash Sutcliffe coefficient and the vulnerability weighted measure of fit, F^2VW allow a broader comparison of the modules and the parameter space. The range of values are summarized in Table 3, and show impact of level of physical representation on results where increasing representation leads to higher mean values and ranges, for the goodness of fit functions. The highest value from any model is 0.963 from a Roe model using a 2m non degraded BB DEM, with a distributed friction value of 0.014/0.045 and hydrograph 10% below the calculated value. The Rusanov solver produces it high value (0.8577) with the same data set but a uniform distribution value of 0.1. The ACC module uses a porosity model DEM at 2m and a hydrograph 10% over the calculated value in its highest function result (0.8404). The highest ATS fit of 0.78 uses the highest distributed friction values, the highest hydrograph and a BB DEM. The indication from this is that each module has a different response to the parameter space. It also indicates that the ATS module produces this higher value as a result of extreme values that may be uncertain, and with refinement of these parameters may be incapable of producing the same peak value. By comparison, however, the simplified approach produces the highest mean Nash Sutcliffe value and the smallest range of results. This indicates that uncertainty in parameters is less critical in this measure. Furthermore, the parameters sets that create the highest values are similar, but with small variation, which will be discussed in more detail later in this section.

Model(value)	ATS	ACC	Rusanov	Roe
F^2 (Mean)	0.66448	0.71524	0.70262	0.76876
F^2 (Max)	0.78211	0.84048	0.85777	0.93648
F^2 (Min)	0.50399	0.49954	0.47314	0.61426
Nash Sutcliffe(Mean)	0.85782	0.74928	0.58309	0.79076
Nash Sutcliffe (Max)	0.99684	0.99183	0.91998	0.99887
Nash Sutcliffe (Min)	-0.69450	-0.03843	-0.16952	-0.15819
F^2VW (Mean)	0.70716	0.70716	0.71852	0.78608
F^2VW (Max)	0.863582	0.86358	0.86786	0.93607
F^2VW (Min)	0.30388	0.30388	0.42439	0.65434

Table 4.2: Summary of F^2, Nash-Sutcliffe and Vulnerability Weighted F^2 (F^2VW)

In order to explore the impact of physical representation on model results, visual inspection of the data provides insight into the range of results from each module. A visual inspection of the F^2 result is provided in Figure 4.8. and box whisker plots of Figure 4.9 which highlight the key model performance results from the modules. Figure 4.8 provides a comparison between friction values, which is typically viewed as a key uncertain parameter in modelling exercises, against module choice. In this figure, an increase in the both the separation of the lines at each Manning's friction value and differences in the gradient of the lines, which represents variation in function performance, is noticeable between each module choice. The variation in model performance relative to module choice in Figure 4.8, would suggest this is a factor of greater significance in affecting model results.

For the reduced physical models, the gradient of model performance is very low, indicating low significance for the choice of friction value. This is in keeping with other issues concerning diffusion based approaches, such as insensitivity to parameter values (Hunter et al 2005). At higher friction values, the gradient across module choice drops significantly, and convergence between models occurs..

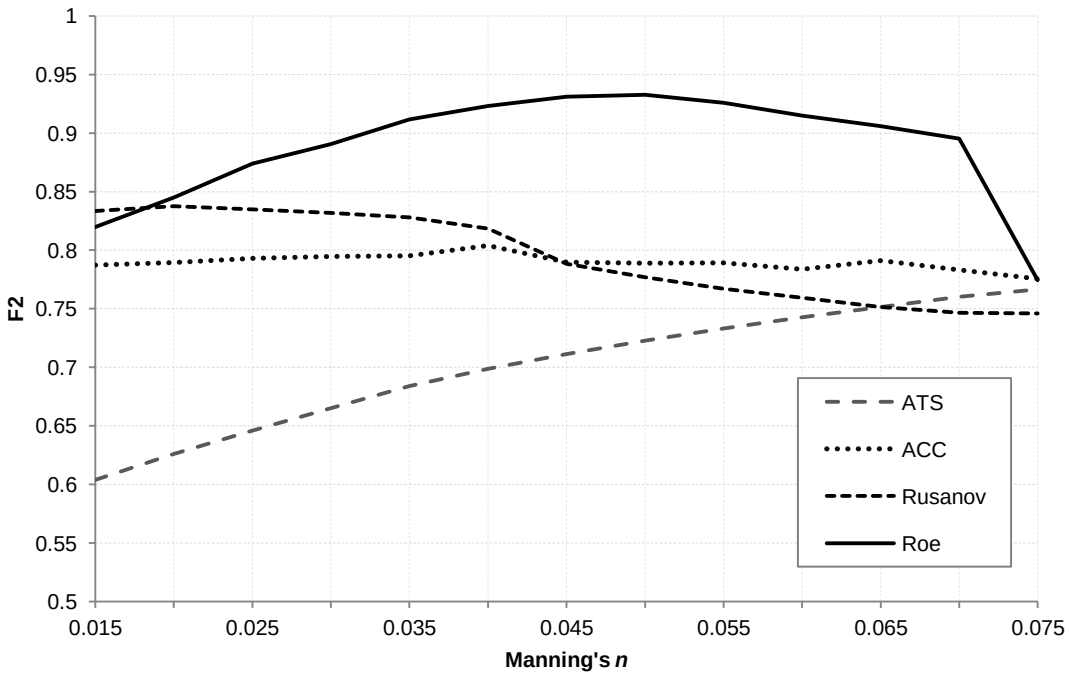

Figure 4.8: A comparison of F2 values against Manning's friction for each LISFLOOD module.

This indicates that a level of parameter interaction (where a combination of parameters influence model results, rather than the a single parameter), and that for high friction values, the significance of the level of physical representation

decreases. Overall, this figure indicates that the ACC module is insensitive to the friction value in a similar manner to the ATS module. The box and whisker plot (Figure 4.9) represents the range of function over the entire model ensemble, where the red line represents the median, the blue lines of the box, the 25 and 75th percentiles and the black lines represent the full range of the results. Single red crosses outside this range represent extreme outliers. This figure provides not only an insight into the range of results, but also indicates the distribution of results. Viewing the box and whisker figure, there is a clear pattern between increasing F^2 value and physical representation; but that the range in the level of uncertainty, represented by the box in the diagram, means that the lower level modules are capable of replicating the results of the full SWE modules. The increase between the ACC and ATS modules is small, and the variations between the two would appear to be overcome by a robust calibration process. This represents the level of potential uncertainty associated with calibration methods and with the reliance on parameter data sets to reduce uncertainty. The improvement in the value of F^2 does not necessarily mean an improvement in the model's ability to replicate reality; rather it reduces the known uncertainty down to an acceptable level. The Rusanov module produces a wider range of values in comparison with both the Roe method and the ACC module. In comparison with the ACC module as well, a higher mean value is achieved which indicates that whilst the model is capable of a higher model performance value, a higher level of uncertainty is associated with the use of this module. The analysis of the Nash Sutcliffe evaluation technique reveals a different pattern. Here, the ATS model produces a far more consistent range of results over the parameter range than either of the other codes. This is a product of its insensitivity to other parameters, which creates a consistent range of results, reflected in the uncertainty extent of Figure 4.5. Not only is the range consistent, the peak and mean value are higher than the other modules. This indicates that the ATS code has a greater ability to replicate the peak and the time of the peak, which is an influential factor in affecting the Nash Sutcliffe value (Pappenberger et al 2004, Gupta et al 2009).

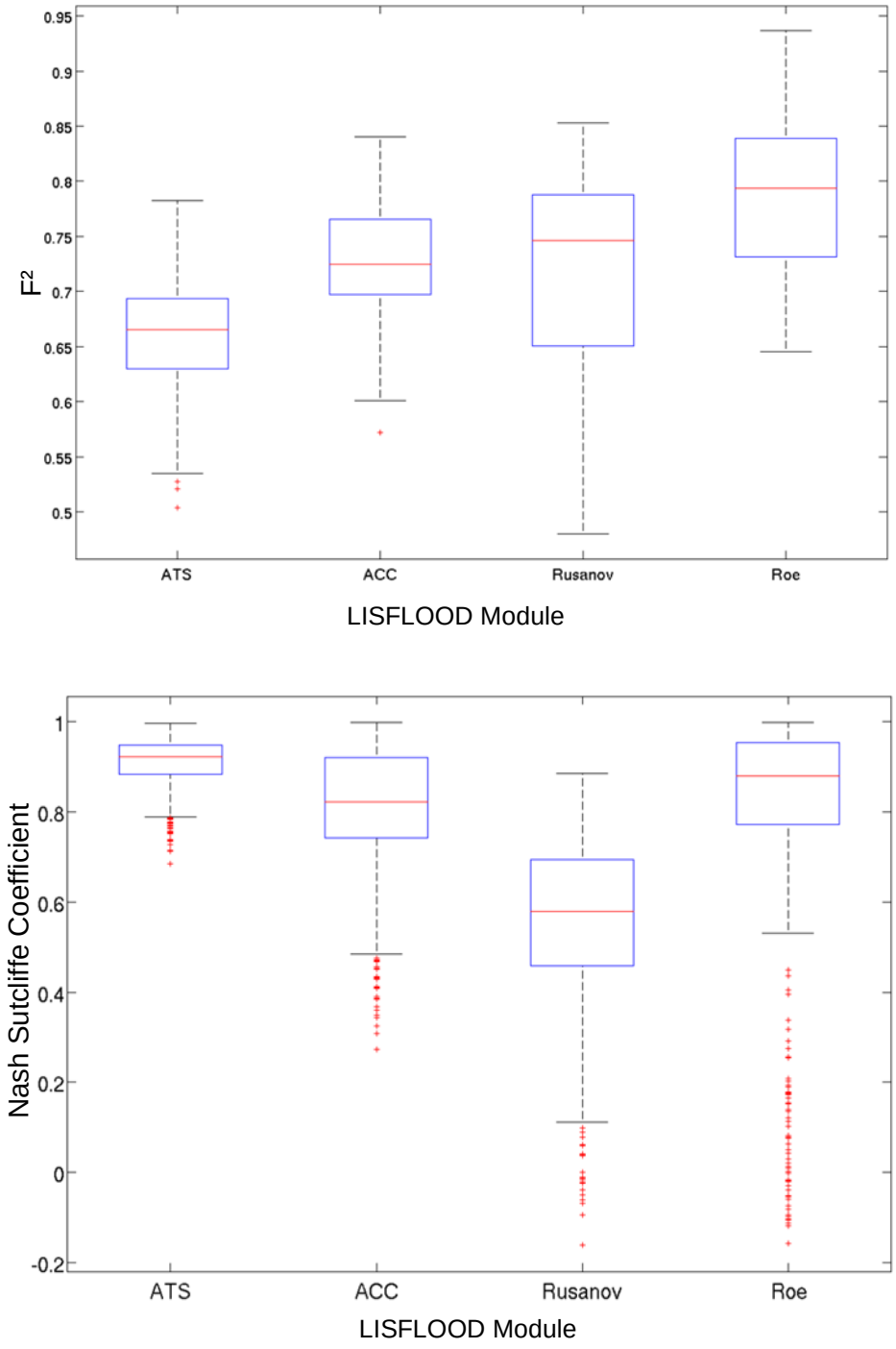

Figure 4.9: Box and Whisker plots of the F^2 value (top) and Nash Sutcliffe value (bottom) for the modules.

There is a significant variation between the full SWE modules. The Roe produces a similar range of results as the ACC module. The Rusanov module, however, produces a wide range of model results, with a low peak value in comparison to the other modules (Table 4.2). This would emphasis the point made in conjunction with the evaluation of the control points, that the Rusanov code becomes more uncertain in respect to replicating water depth. Furthermore it indicates that the simpler numerical solution may be more susceptible to uncertain parameters. The wide

number of outliers for all the modules relates to models the use the BR and BP methods of topographic representation. The affect of this has been noted with respect to the flood probabilities. This indicates that this factor may have a relatively influential impact on model results, and is therefore an area that should be carefully considered in model construction.

In using multiple criteria to evaluate models, a comparison of the objective functions is required in order to evaluate the best performing models across all model types. (Dung et al 2011). Figure 4.10 is the combined functions of F^2 and Nash Sutcliffe, where each dot represents a model run and the colour relates to the LISFLOOD module choice.

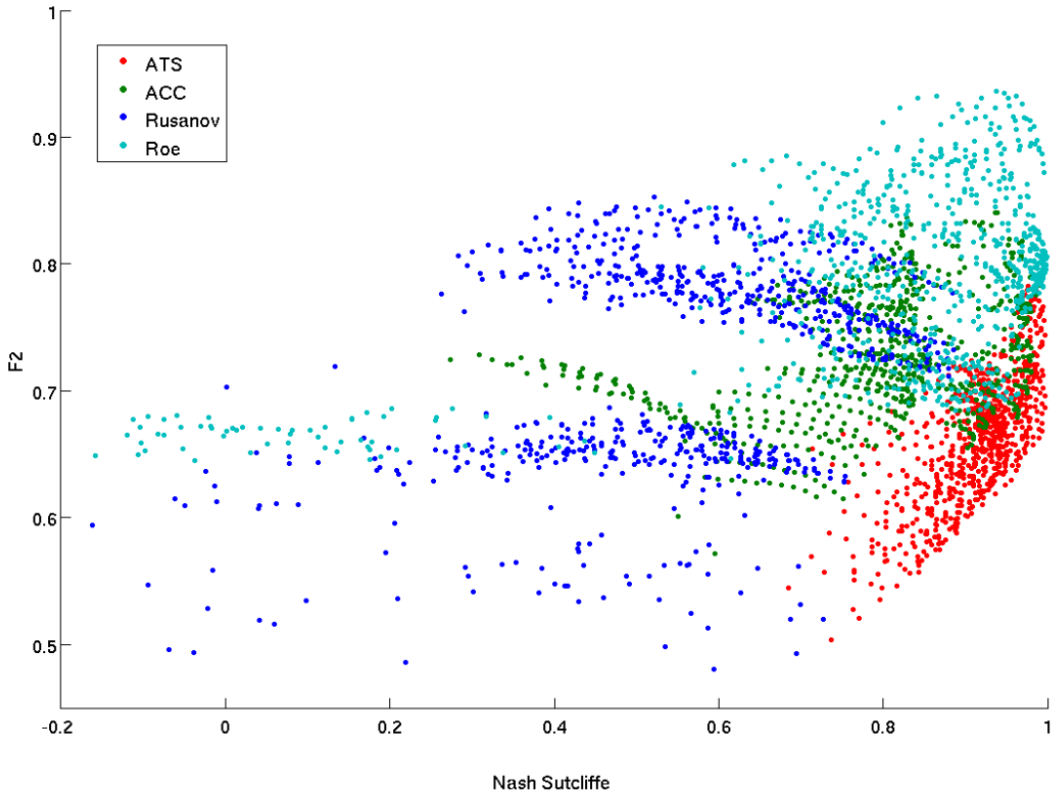

Figure 4.10: Combined results for F^2 (y-axis) and Nash-Sutcliffe (x axis)

A number of key points emerge from analysis of these results. The use of two methodologies in assessing model performance helps to refine the choice of high performance models in comparison to the use of a single function. The top right of the graph is the region in which the pareto front of high performing models for both criteria would form. In comparison to the single performance criteria approach discussed above for both comparison of extent and Nash Sutcliffe, the combined approach produces a fewer number of performing models. Analysis of these results also indicates the importance of module choice in determining results, with clear clustering of module results in certain regions of the domain. The majority of Roe

model results appear at the top right of the graph, indicating good combined performance. The parameter set for the models at the pareto front are 2m BB resolution models, with friction values around 0.045, and hydrograph values between 0 and minus 10 of the original. By contrast the Rusanov and ACC modules, although able to produce as good F^2 values fail to provide the same level of performance for either module, and produce a wider range of results, suggesting a significant amount of uncertainty in the use of these modules. This highlights the importance of multiple functions is assessing models. The position of the ATS cluster is a product of its excellent Nash Sutcliffe performance level, but a poor goodness of fit value. The shape of the cluster is smaller in extent, indicating what has been seen in previous figures, that the sensitivity of the ATS module to a range of parameters is relatively low compared to the other modules.

The focus in this analysis has been on the impact of physical representation on model results, although the influence of other factors is also noticeable on model results. Figure 4.11 is an interaction matrix plot where the columns and rows relate to the range of parameters for the labelled input. The mean value of the objective function (in this case F^2) for a value of a parameter is then calculated, and compared to the mean value of the input value for the parameter it is cross compared with. Therefore each box represents a comparison between two parameters, where the lines displayed relate to a value of parameter, noted on the right hand side of the figure. The values along the Y axis represent the mean function value of the parameter when used with the second parameter. For example, the matrix box of row 1, column 2 is a comparison of module choice (level of physical representation) and Building Representation. The lines relate to the module choice, and the position of the line relates to the mean performance level of the module, for models that used a particular building representation input (for example the middle point of the lines relates to models that used the BP method). The gradient of the line indicate the influence of the parameter in the first order (direct impact on model results), whilst crossings of the lines indicate higher order interaction between the parameters.

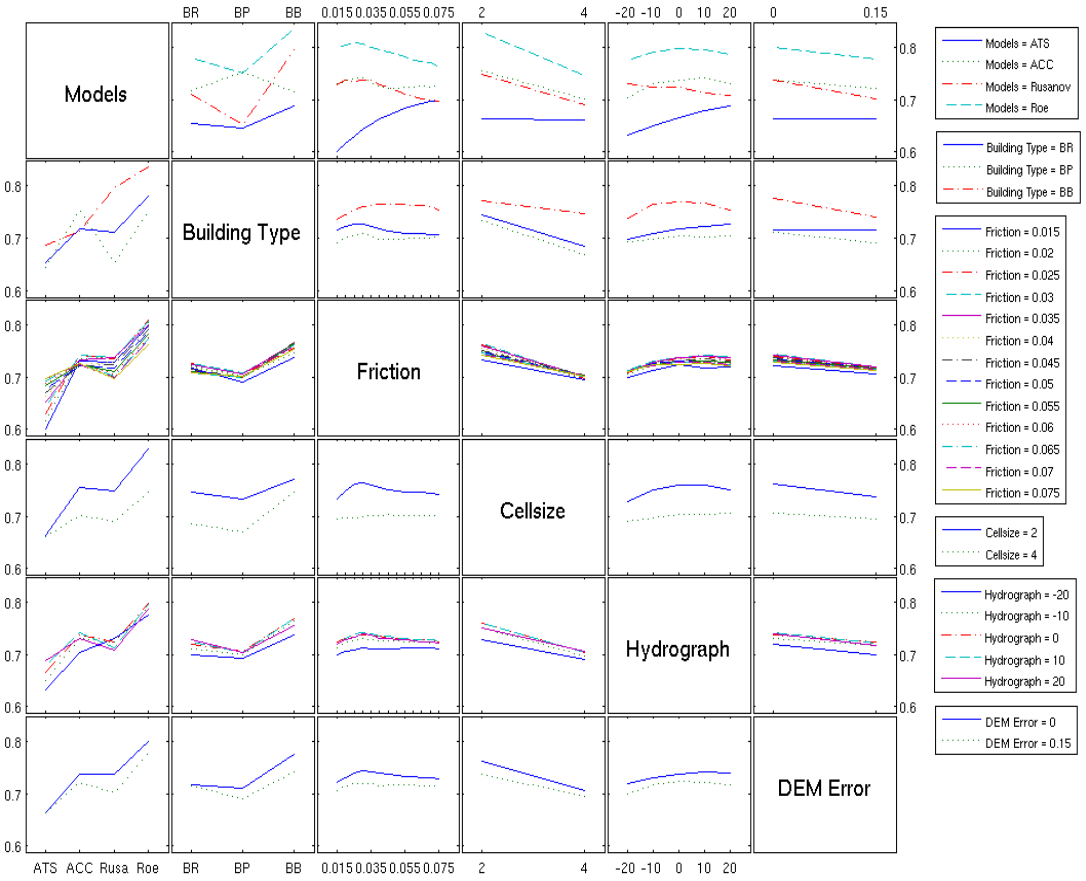

Figure 4.11: Interaction plot for each input factor comparing mean value of F^2 per level of factor other levels of factors (where F^2 is the value on the y axis of all plots and the x axis in each column is the levels for the named parameter)

This figure allows the clear dissection of factors that influence the overall results. The sharp gradients associated with the module choice (the first column) indicate the clear influence of this factor on model variation. The variations between the lines of the row relating to module choice (row 1) also indicate that each module produces significantly different results to each other. The variations between rows in the first column are also small suggesting that overall, the influence of other factors relative to this are small, although a number of additional interactions provide insight into the influence of other factors. The building representation factor also contains strong variations across the figure (second column and row). A note worth making is that the mean performance level for ACC and the porosity approach is higher than these modules mean level for either the Building block or Building resistance method (top row, second column). This can be accounted for by the fact that the highest F^2 value produced by the ACC model contained the porosity model as the building representation factor. By comparison the full SWE modules and BP method produces a lower mean value. The influence of cell size as a key factor can also be observed in this figure, and its influence as an interactive factor relative to the module choice can also be seen at key areas within

the graph. For module choice, as cell size increases, the mean level begin to converge at a lower point. This indicates that with increasing cell size, the influence of the acceleration terms within the SWE equation begin to lose significance, although this leads to a drop in overall model performance. This is in keeping with other research that indicates at larger gird resolutions, diffusion wave properties begin to dominant (Hunter et al 2007). A similar pattern can also be noted for with the gradients of the friction parameter. Here, as with cell size, the increase in friction value leads to a convergence of mean levels, which has been previously noted with the comparison of F^2 values between friction value and module choice (Figure 4.8) By contrast, the influence of the DEM error and the value of the hydrograph all appear to be of lower significance than these effects, which can be seen by the low gradients of the lines in their related comparison plots.

4.2.2. Distributed Friction Values vs. Single Friction Values

A key aspect in the construction of the model is the choice of the friction value, as well as the distribution of the friction value across the spatial domain of the model. For these cases the choice between using uniform values or spatially distributed values based on the underlying surface has been explored. Using the rows that relate to the value of Manning's *n* from two separate interaction matrix plots, one that used distributed values and one that used single values, a comparison of the mean level of F^2 performance across the two approaches can be made (Figure 4.12). In comparison, the variation between the approaches appears to be small and significantly less than the influence of both level of physical representation and the representation of the topographic surface. However, small variations between the two methods can be observed which influence the uncertainty associated with this factor and provide insight into the significance of this factor.

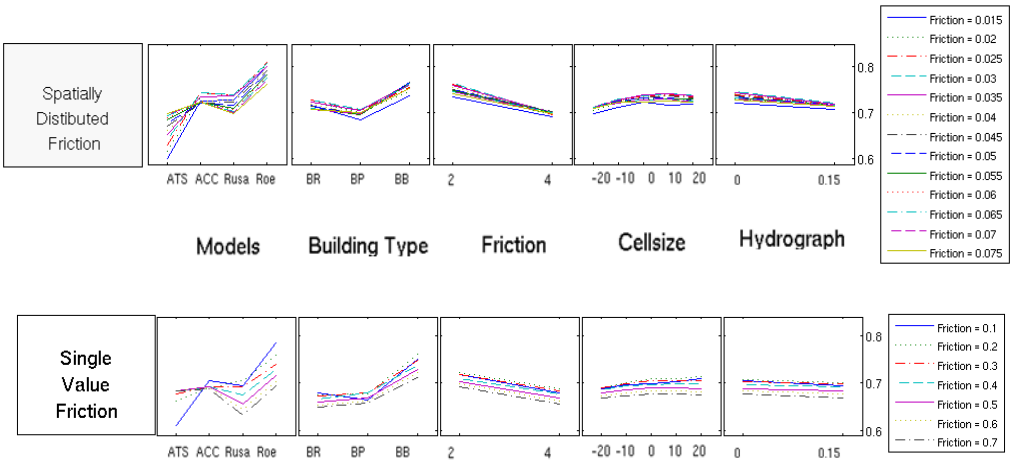

Figure 4.12: Comparison of mean level per parameter for spatially distributed friction and single value friction. The comparison parameter from left to right is model type, building type, cell size, hydrograph, and DEM error.

Whilst the shape of the interaction lines remains approximately the same across each of the factors, the range between each line is larger for the single value than the distributed value, indicating an increased uncertainty associated with choosing a single friction value. Furthermore, the value of F^2 reduces between friction representation approaches, where the highest values of 0.8 can be observed for the comparison with model type. Partly this can be accounted for by the use of a spatially distributed friction value in the creation of the benchmark model, but it also indicates that the use of spatially distributed friction values can reduce uncertainty in the model performance.

Further analysis of friction distribution techniques is taken from a comparison of water levels over the duration of a model run for control points along the main flow pathways. The change from a friction value of around 0.01 to 0.07, causes a dramatic increase in the way water moves through this flow path, creating wider storage regions, as water is slowed through the model domain. These affects though appear to be localised, and do not affect global model performance values in the same way.

This indicates that where low value friction surfaces are represented explicitly, the precise friction value becomes less critical to the overall model uncertainty. However, where broader generalisation of friction surfaces is used, the choice becomes more influential on results, and consequently become critical to model output uncertainty. Given the use of a benchmark model in this test case, that used spatially variable results, this may be less of an issue in the following test cases, and further analysis will provide insight into how important this is to overall modelling practices.

4.3. Exposure Based Analysis

For each module the uncertainty of a certain hazard level being reached (being defined as maximum water levels and velocities over 0.7 qumecs) was evaluated across the range of model results, and is displayed in Figure 4.13. This uncertainty hazard plot approach, similar in principle to the uncertainty flood extent plots of Figure 4.5, has first been used by Aronica et al (2012). This approach highlights not only the regions where the hazard is likely to occur but also identifies the certainty of a region flooding based on the model ensemble. The variation between SWE models and simplified approaches is clear, with the full dynamic models producing a wider region of hazard. The use of velocity in the definition of hazard has a clear impact in the appropriateness of the simplified approach in defining flood risk. In this respect, the choice of model would be critical in determining the spatial aspect of hazards associated with flooding.

A comparison of the Rusanov and the Roe flux appears to show that the Rusanov flux defines a larger region as being high in hazard and inundation frequency. This is in keeping with the analysis in the first part of this section where higher velocities were a product of the simplified approach. An interesting point of comparison is that whilst the ATS and ACC regions suggest smaller regions of high certainty hazard occurrence, they are in the same spatial region, near to the inflow point. The extension of the hazard region by the ACC model is highly uncertain, indicating that both ATS and ACC produce similar and reliable regions of hazard across the test cases. Further insight is also gained into the overall modelling uncertainty. The majority of the SWE's models hazard frequency is around or below 0.6 indicating a high level of uncertainty for defining this hazard region. Furthermore, the regions are not well defined, particularly for the Rusanov image and between regions of high hazard frequency, a number of individual cells appear across the model domain. These areas could be removed through a resampling process, but here they provide information about how the models behave. This speckled affect appears to be the product of instabilities within the model that only occurs for a small number of test cases, where low friction values are used. Another insight that this approach provides is insight into main flowpaths across the floodplain, as the hazard rating allows the identification of where the depth of water and velocity is significant. Not only can this help to improve the understanding of the model output and the hydraulics of the event, over simple extent comparison methods, it also improves the evaluation of models by determining key regions in the model domain that could be determined as hazardous. In this test case, the two main flow paths

can be determined from the full SWE model results, and the ACC model to a lesser extent. The lack of definition of flow paths in the two simplified approaches indicate that a potential underestimation of the flow paths and therefore the risk of the event have been made. Moreover, it indicates that the simplified approaches provide a different hydraulic output to the full SWE that cannot be determined from extent and depth comparison points. How this relates to deterministic approaches to evaluating exposure and overall risk can be determined from the other exposure methods.

4.3.1. Cost of Damage Method

The results from the Cost of Damage method are summarised in Table 4.3, for each module. The model output is a total cost of damage, based on water depths in cells surrounding the buildings which are then compared to a depth damage curve for the property building type. The wide range of results can be attributed to a step change in value output between the two cell size values of 2m and 4m. The variation in results is not only attributed to variation in the results of the two cell sizes, where water depths at key locations can vary between 0.7m and 0.25m between the same parameter set models for 2m and 4m

Model	Mean Value	Minimum	Maximum	Standard Deviation
ATS	£4,614,684	£1,206,528	£9,357,948	£2,297,434
ACC	£6,507,373	£3,298,732	£11,211,234	£2,304,670.8
Rusanov	£6,147,870.3	£3,281,268	£9,255,382	£1,971,267.9
Roe	£6,022,814.2	£3,017,600	£9,328,534	£1,961,454.2

Table 4.3: Results from the Cost of Damage method.

The results are also potentially affected by the approach used. This represents a key component in the use of monetary methods for evaluating results, in that consistency between a range of approaches for cell sizes, building representation techniques and spatial discretization techniques is critical in drawing significant conclusions. A comparison of this method compared to the traditional methods identifies a number of similar characteristics between the two approaches as well as the similarities between the influences of model parameters on output, as well as highlight other parameter interactions.

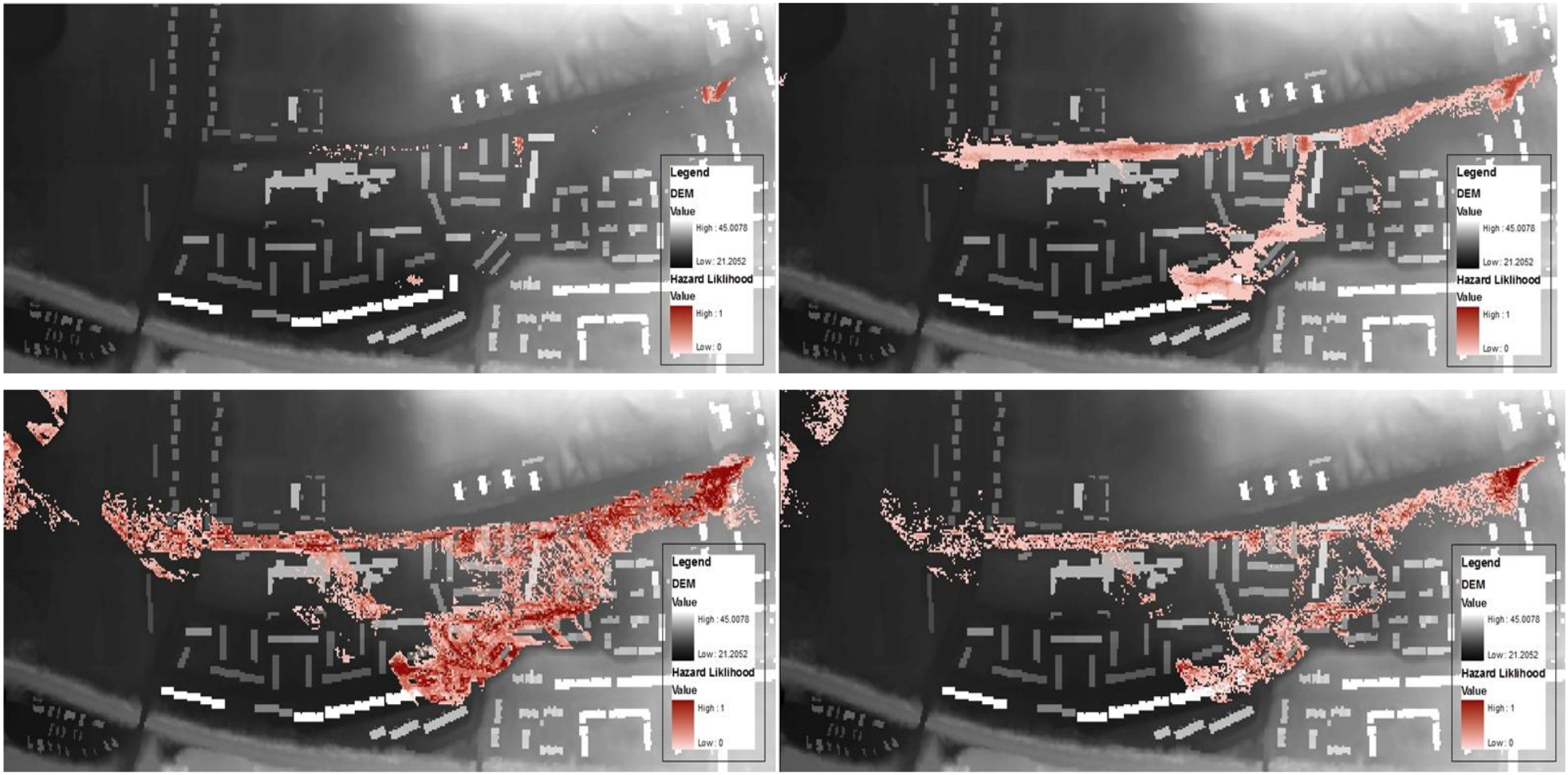

Figure 4.13: Uncertainty Hazards plots, where red regions represent high hazard (top left ATS, top right ACC, bottom left Rusanov, bottom right Roe). Darker red represents regions of high hazard frequency and areas defined as hazardous in all model runs

Figure 4.14 is a comparison between F² (blue line, left hand y axis) and Cost of Damage method (green line, right hand y axis) for the entire range of Roe model results. This allows a broad overview of the change in model results over the range of parameters. Each grid line on the *x* axis is a split between cell size and building type to allow for easier comparison. Within each section, the hydrograph and friction values increase from left to right. The midpoint of the *x* axis represents the change from 2m grid models, to 4m grid models. At this point, a significant drop in calculated cost can be observed, which is also replicated by the F² value, but to a smaller extent. This drop is observed for all model types, indicating that the choice of cell size is a highly significant parameter in determining cost.

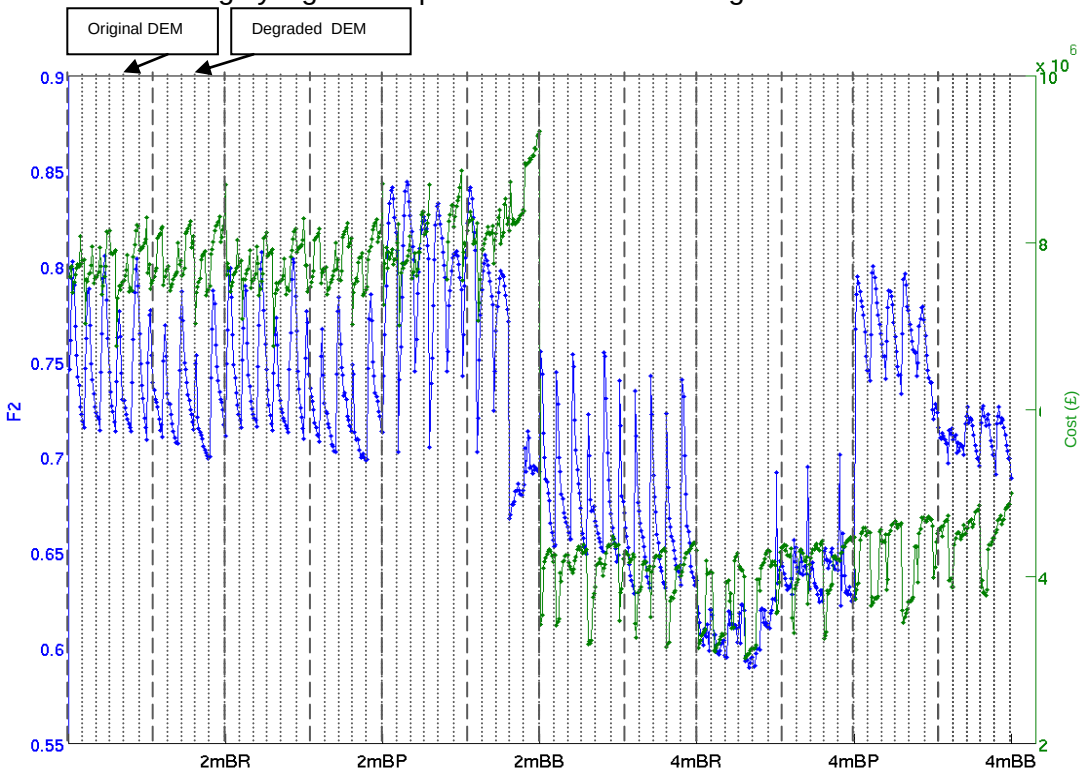

Figure 4.14: Comparison of the Cost of Damage (green line right side axis) and the F ² value (blue line left axis) for the Roe model results. Each dot is a model run, with the lines to provide an aid to the eye. The x axis represents groups of models with the same cell size and topographic representation, which are labelled to the left of the marked grid lines. The dashed line in each section represents the split of models with non degraded DEM's and degraded DEM. The minor grey lines represent the models with different hydrographic values – the first subdivision is -20% moving up to +20%. In each minor section, each model represents increased friction value from low surface value to high surface value, although this rise is not marked as a gridline on the figure.

Within each cell section, the variation between the individual parameters appears to be similar to the F² value in that the value of friction is significant. As with the $F^{\langle 2 \rangle}$ and Nash Sutcliffe, the value of the hydrograph appears to be relatively insignificant to creating a variation in model results, in comparison to other parameters, which is perhaps unusual considering that the depth of water in the cell is critical to the

calculation of the financial cost. This approach allows further insight into the impact of the level of physical representation and topographic representation technique. The highest value across the total range of model belongs to the ACC model, and specifically to 2m resolution model, using the building resistance methodology and single value friction across the rest of the domain. These models produce a spike above the typical range of cost seen across the rest of the model results (Figure 4.15). This also relates to a variation in the goodness of fit function, which relates to the use of the randomly degraded DEM.

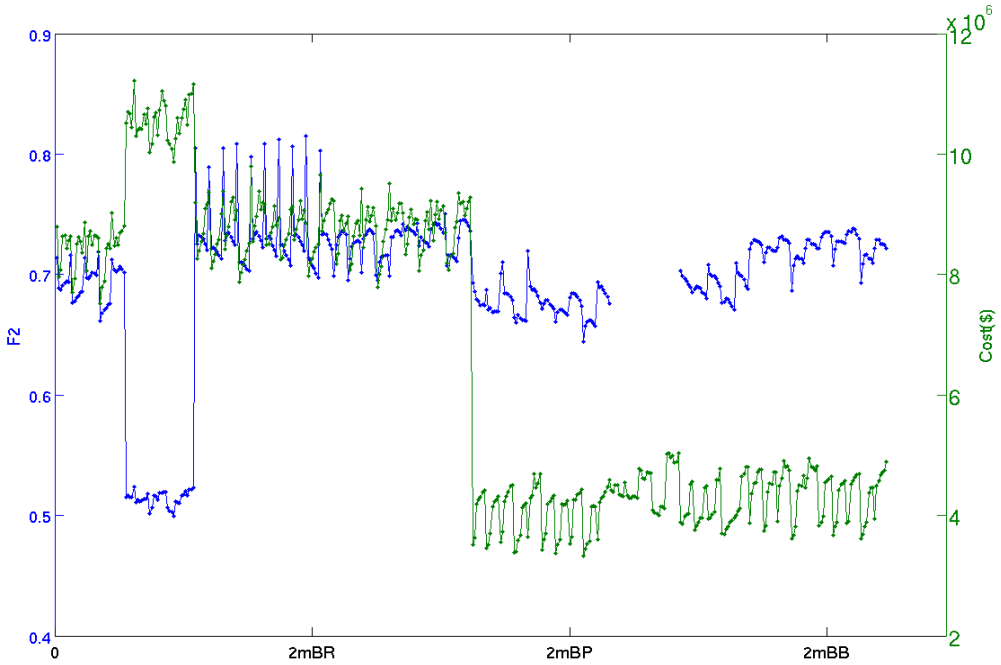

Figure 4.15: F^2 (blue line, right axis) vs. Cost of Damage (green line left axis) for a range of ACC module results based on a 2m grid resolution.

For this set of models, the random error increases the resistance to the flow path, and reduces the velocities along critical flow paths and therefore the maximum flooded extent. This may be a product of the approach used here (randomly degrading the DEM), which leads to an uneven surface. This sensitivity to the DEM error in association with the building resistance method provides further insight into model performance, whereby simplified approaches to both the level of physical representation and topographic representation increase the sensitivity of the results to the underlying topography. This approach also confirms the importance of using multiple methods for evaluating model performance, and some of the advantages of using depth/damage curves, whereby modelled depths can be incorporated into a single function with which to describe a model realisation. This allows local variations in models to be more directly applied to the model output and provides a global model output where depths at critical locations can be directly interpreted from model results.

4.3.3. Regional F²

A comparison of regional extent fit values compared to the global performance value provides an insight into how regional variations contribute to the overall global performance value. Using the comparison of different performance measure approach used by Pappenberger et al (2006), the performance in each sub-domain can be compared to each other and the global performance level. The regions, displayed in Figure 4.4 were determined from regions of similar vulnerability based on the underlying building type. Analysis of the global method versus the regional value (Figure 4.16) provides key insight into how the performance in each region compare to the global value as well as how regions perform relative to the other region. Each region appears to have similar levels of model performance. The strong positive correlation between each sub domain also indicates that in these regions model performance similar to other regions. Therefore, in these regions, identified as being of critical importance to the vulnerability of the model, model performance is similar across the range of parameters, suggesting that the level of uncertainty in the global model performance value approach is less critical in determining the associated risk.

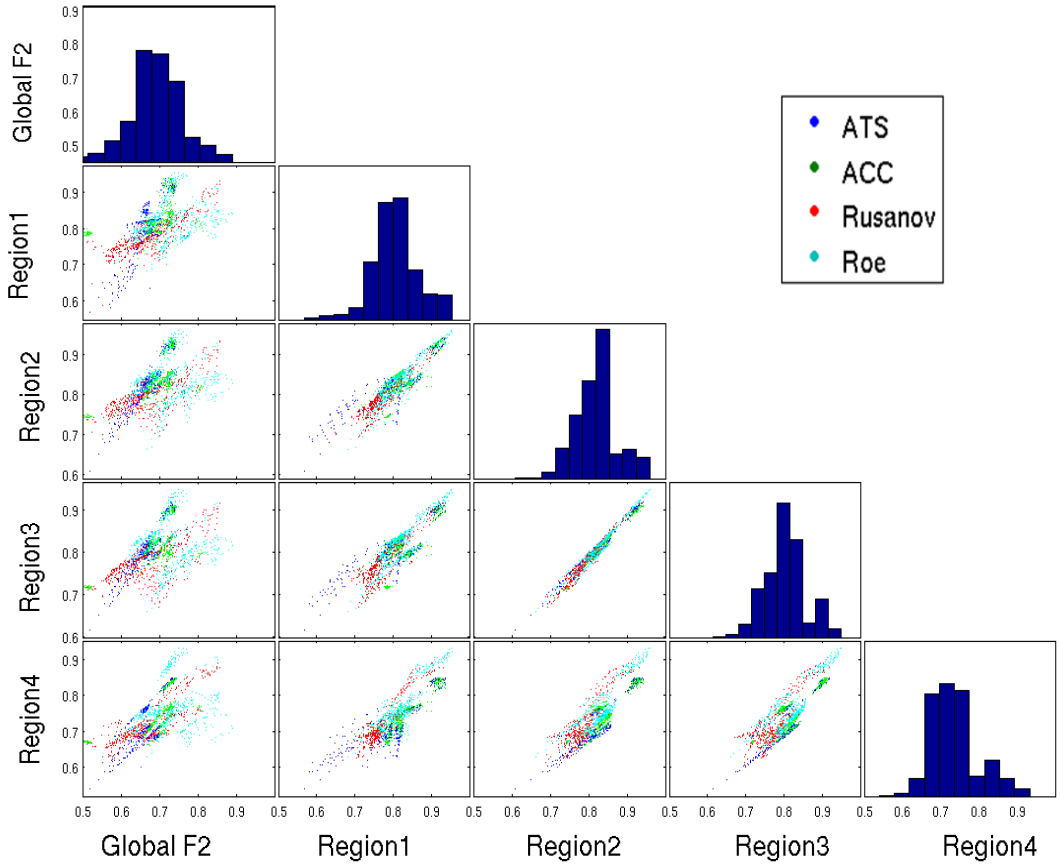

Figure 4.16: Comparison of regional values of F² versus Global F² value (first column). Each dot represents a model run.

Figure 4.16 also highlights that all modules are capable of replicating the benchmark extent to a high level in these sub domain regions, with F^2 values being over 0.64. These regions also display low levels of variations of model outputs in comparison to the global model performance, indicating not only adequate model performance for all modules in the regions of significant hazard and vulnerability, but also reduced impact of uncertain parameters for this test-case. This is further confirmed by examination of the first column which displays the global performance against the regional models. Each row appears to be similar to the other rows, indicating that each region has a similar impact and model performance level to the other regions. This indicates that these regions have a similar contribution, and that wider model variation, seen in the original evaluation of F^2 is from regions which are less critical to determining risk (i.e to the west of the model domain).

This approach can also be used to determine regional modular variations, by using the model colour to determine patterns between models. The spread of Rusanov models in each region (red dots) confirm the findings of the earlier analysis concerning the sensitivity of the Rusanov solver to uncertain parameters and the wider range of results seen relative to the other parameters. The ATS module also contains a similar spread across the regions, indicating that this module may have more local variations than previously identified with the global model. By comparison both ACC and Roe produce narrower extents across the regions indicating a similar regional performance level and sensitivity to model inputs. Determining the regional variations in a more detailed way and examining the parameter space and function response provide a useful additional area of work to provide further insight into local model variation, as well as enhance understanding of exposure as a critical model evaluation technique.

4.3.4. Weighted Vulnerability Approach

The weighted vulnerability approach produces a range of results similar to the traditional evaluation F^2 techniques, although it has increased the peak value of and F^2 decreased the minimum value. This is an important component of model evaluation techniques, in that the function should determine a clear set of ‘best performing’ parameters to avoid the issues of equifinality (Beven 2002). Although the peak value is increased and refined to a smaller set of parameters, the general trend is a reduction in the variation between the function of different models, creating a larger number of similar performing models. An example of this is in Figure 4.17, which compares the F^2 weighted function for each module over friction value, similar to Figure 4.8

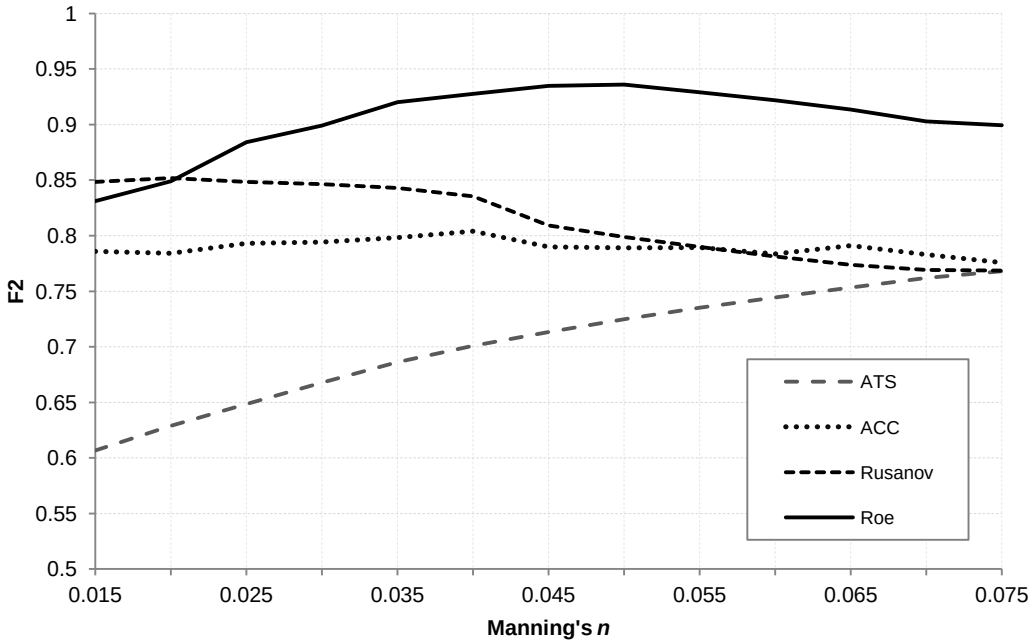

Figure 4.17: The F^2VW value plotted over the Manning's friction parameter space for each LISFLOOD module

A clear reduction in the clarity between the model types, compared to Figure 4.8 can be seen, particularly with the ACC and Rusanov models, indicating that the use of weighted cells will in some cases improve the level of model performance by reducing the impact of poorly performing models. Part of this can be accounted for by the subjective approach used to define the cells that are weighted, in that the weighted cells are located near the input source and along the main flow paths, and are likely to be flooded in most test cases. The majority of cells are also weighted with the same value, to reflect the fact that the level of housing is similar across the model domain. In more varied economic regions this may prove to be less of a defining factor in the application of this approach, and allow for greater variation between model output.

In comparison to the Cost of Damage method, a weighted cell method lacks the precision to be able to differentiate at a greater resolution between the parameters used here. However, this method does still provide the means to explore model results and to see how localised model variations can impact a global model output. As with the Regional Vulnerability approach, further test cases will provide more information about the approach and the usefulness of it as a appropriate methodology for calibration.

4.4. Sensitivity Analysis

In order to quantify the results seen, GSA methods are used. The Morris method is used to provide a overview of each model function, whilst the BACCO method uses only F^2 to analyse results. The Exposure and Vulnerability methods have been excluded to allow for more direct analysis between parameters and output. For the Morris analysis, a wider range of cell sizes were used to further explore this parameter, with the parameter range extended to include an 8m and 6m grid domain. These values were used to represent the extreme limit of the range of cell sizes that could be used. Here an 8m grid cell would be near half the width of the road network in this model region, and provides an indication of the sensitivity of this factor in model results. This analysis can only be used to rank the input factors in terms of significance, and not quantify the impact. The figures below are used to provide a visual analysis to how the factors rank and compare

The results from the Morris method confirm some of the findings identified in the previous analysis. For the F^2 and Nash Sutcliffe objective functions, the choice of physical model appears as the most influential factor, with a high mean value, which indicates significant first order impact and high standard deviation which indicates higher order interaction. The other parameters vary in order of significance for each of the objective functions, but a significant aspect of the analysis is the influence of higher order effects, which is represented by the standard deviation of the mean elementary effect. Where the F^2 function has been used to evaluate the model, the parameters form an approximate linear relationship with the exception of the topographic representation factor and the grid cell size, which have low first order effect, but a higher interactive effect on model results. This is in part due to the response of the ACC model, which for the BP method has a higher mean compared to the BR and BB methods. This effect is less pronounced across the other module choices, hence the higher interactive influence. The effect of the cell size has been less pronounced in the earlier analysis. For this analysis the cell size was increased to 8m, and this appears to create a higher order effect. This follows from previous results in this section, where increasing cell size decreases the performance level of the SWE modules. The use of F^2VW as an additional means of evaluating the model results produced similar results to Figure 4.18, indicating that the additional weighting scheme did not provide further insight than the traditional function methodology in sensitivity analysis

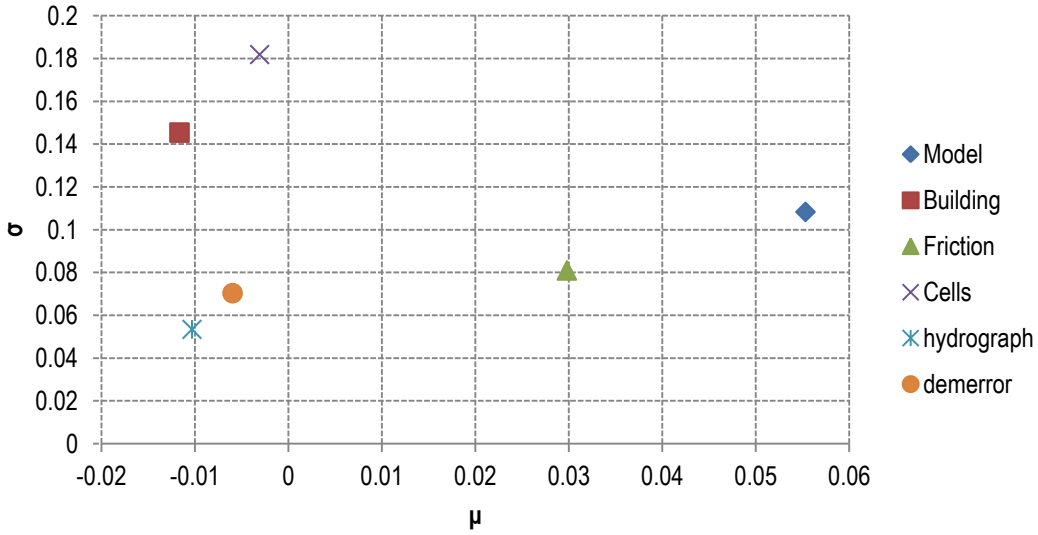

Figure 4.18: Morris Results for F^2, where the x axis represents the mean elementary effect. The values themselves only provide a relative ranking position and do not quantify the level of interaction.

Analysis using the Nash Sutcliffe coefficient reveals a more linear relationship between first order and higher order effects. The influence of topographic representation is as an interactive factor with other parameters. A variation between the two functions is the relative position of cell value relative to the other factors, which would be ranked second for the two functions. This is a product of the impact of cell size on water depths in cells, which also affected the Cost of Damage evaluation technique. The hydrograph has also improved in terms of relevance, indicating first that where water depths are critical to model performance, the significance of this increases, and also highlights the need to move away from simple extent fitting exercises, which can often ignore information critical in affecting hazard results.

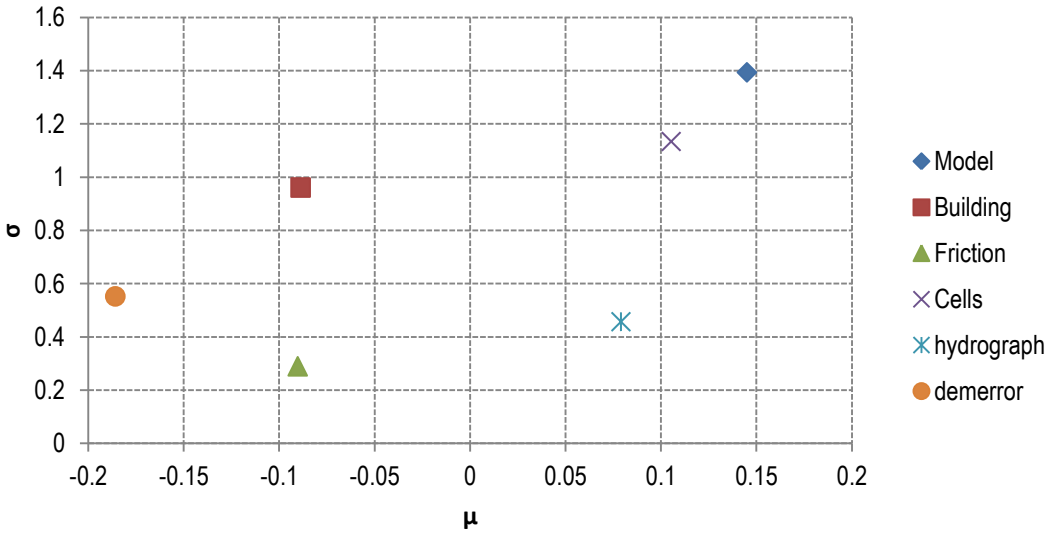

Figure 4.19: Morris Results for the Nash Sutcliffe function.

The results for the BACCO GSA are summarised in table 4.4. A much reduced parameter space was used in this test. The level of physical representation and the building representation method were both excluded as input to the analysis. Instead a comparison of a simplified approach (ACC) and a full SWE model (Roe) was undertaken in order to explore how parameter significance changes across the levels of physical representation. Consequently the analysis of these results is related to how a parameter varies in its level of significance between the two modules. This is partly the product of complications relating to the functions used to test and build the emulator. An essential requirement of the use of BACCO methods is an output function that varies smoothly with model input. For the ATS model, whose function remains relatively in sensitive to model parameters, creating an emulator proved to be problematic, and a poor emulator variance was returned. A typical way to improve the emulator performance would be to increase the number of training points and reducing the spacing between these points. Given the additional computational costs associated with it and the potential for not improving the results sufficiently, a smaller parameter space was explored, which included Cell size, friction and the hydrograph value and used 2 of the LISFLOOD modules to represent the simplified approach (ACC) and the full SWE approach (Roe). The parameters were sampled between a range of 2m and 4m for cell, 0.015 and 0.075 for friction and +-20% of the original hydrograph using a LP-Tau sampling method, ensuring a good coverage of the sample space. 100 data points were used to train the emulator, which gave an adequate fit between the emulator and observed data. The results are summarised in Table 4.4;

Parameter	Total (ACC)	Total (Roe)
Cell	4.42	78.53
Friction	52.23	2.13
Hydrograph	30.07	13.02
Cell Friction	>1	0.69
Cell Hydro	>1	3.86
Friction Hydrograph	13.29	1.20

Table 4.4: Contribution of parameters to overall variance as a % of total variance

The results therefore represent a smaller part of the analysis, but still provide some means of quantifying earlier patterns. The influence of the hydrograph appears to be significantly higher for the simplified approach compared to the full SWE model. The significance of the cell size is also significantly smaller for the ACC module than the Roe solver. This counters what had previously been seen in the rest of the analysis. This is partly down to the occurrence of ACC flowing over the blockage at the end of the main flow path, when the Building Block method is used. As suggested by Figure 4.5, the likelihood of ACC flowing over the blockage is relatively small compared to the rest of the region, and is smaller still when the building block method is used, as the flow path is well constrained, leading to a greater volume of water pooling at the south end of the site. In this scenario the F^2 value is likely to be unaffected by cell size as it is for the other parameters. This indicates again that parameter interaction is critical in affecting not only results, but larger scale trends between parameters and model evaluation techniques. A further variation between the approaches is between the Roe solver and the ACC solver relating to the significance of the friction value. For the ACC solver this is much higher than the Roe solver Overall, the importance of the module choice appears to be critical, this can be infered by the variaiton between the results. This test case however contains a number of features that favour the level of physical reprsentation, such as the occurence of transcritical flow, and the highly constrained flow paths that prevent larger flood extents to occur in the middle of the model domain. This however maybe inhernt to all urban test flow cases. Further test cases will develop this point.

4.4. Overview

The Glasgow test case has shown evidence of the importance of the level of physical representation on results. Each module creates different extents that relate to variations in the value of the model evaluation function. The overall model results indicate the significance of the acceleration terms in replicating hydraulic conditions of the urban environment, and in overcoming obstacles. This confirms previous findings with this test case in research by Hunter et al (2008). What the research here has furthered is that the impact is often local in both space and time, but this is often not reflected to the same degree in the model evaluation function. It has also indicated that the choice of other input factors are less critical to model results, and has further suggested that the uncertainty related to physical representation must be considered first. The use of exposure based methods has highlighted this, and can potentially provide further insight into the importance of localised model results over the entire model domain. The conclusions must be considered in the context of the test case. The use of the a high resolution model to evaluate model runs will impact results, and the conclusions that can be drawn, but does provide an indication of the significance of the model input to model results. The use of this benchmark model approach has advantages including reduced uncertainty related to assessing model output based on uncertain observed data sets, but overall the use of actual data with which to evaluate models helps to validate the conclusions drawn.

5. TEST CASE COVENTRY

The Coventry event is based on a canal embankment failure that occurred during the night of 15th December, 1978, which was the result of ill considered building development by a company adjacent to bank of the Coventry Grand Canal. Removal of the towpath bank, by as much as one third of the available material, lead to a sudden rotation-slip failure and an outburst event that flooded nearby engineering facilities and workshops, local housing and the Coventry General Hospital, an area of over 1km². Figure 5.1 is the surveyed outline of the flood, over a modern day OS map of the region. The event is unique in terms of the hydraulic conditions of the event, the probability/hazards of the event and the data captured for model evaluation – there are few examples of observed canal outbursts in urban areas.

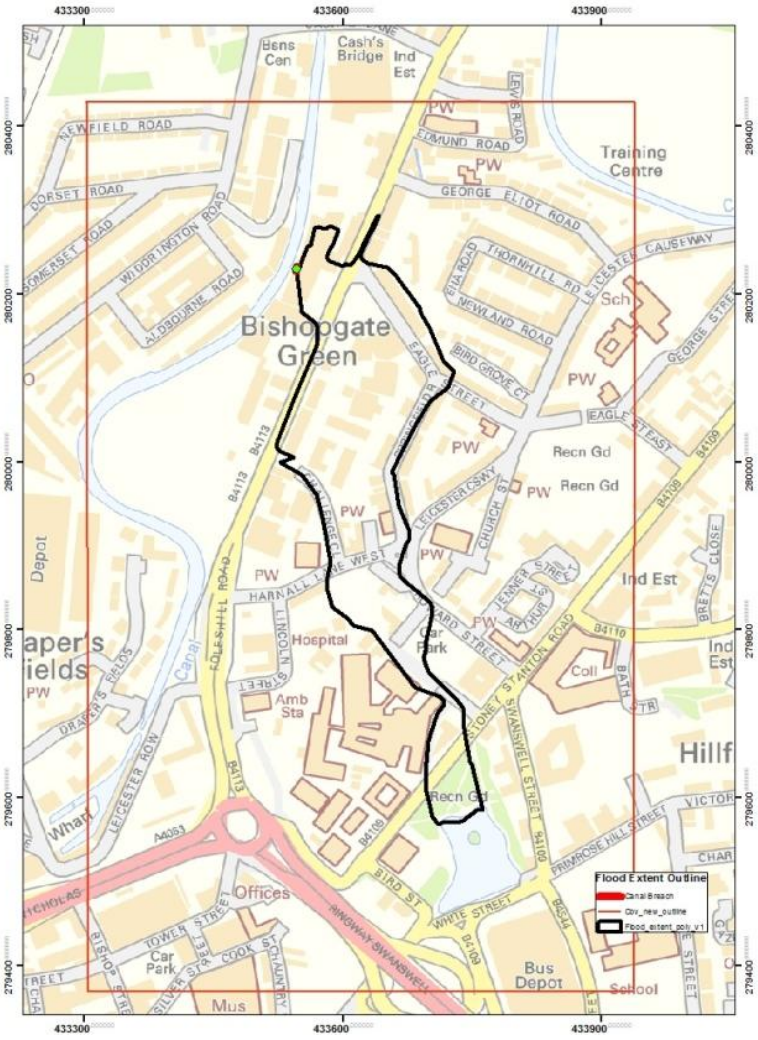

Figure 5.1: Surveyed extent of the 1978 Coventry Canal failure. Outburst location is the green dot, located centre of the image, adjacent to the Canal.

5.1. Overview

A detailed report was made of the event by Halcrow, which identified the key causes and highlighted some of the impacts from the outflow. The basic outline of the event is as follows; An extension of a factory involved removing material from the embankment at the back of a warehouse, and setting down a concrete base at the location. This was around a meter below the base of the canal and around 2.5 meters from the top of the bank. A slip failure then occurred allowing a breach in the bank to drain out the canal. Due to the sudden nature of the event the lack of knowledge of the workings and the availability of suitable blocking material, to prevent water from the entire pound exiting the breach, the event had a larger impact then might normally be associated with outburst events. In combination, these effects created an event which would last for several hours, caused extreme damage to nearby houses and lead to the evacuation of a nearby hospital. The nature of the outflow from this event is unknown, but would have been sudden following the slip failure. The report of the event in the local newspaper cites eyewitness accounts of a wall of water moving down Eagle Street, which was powerful enough to move cars. The engineering report notes that the main flow paths were in a southern direction, down Springfield road and Eagle Street and towards the hospital before discharging into natural systems via the a small pond, located at the south of the domain. A smaller amount of water also moved laterally down the Foeshill road, inundating several shops, a petrol station and some houses, although this flow was smaller in quantity and velocity than the main flow path. This test case presents a similar set of hydraulic conditions to a dam failure, although the precise nature of the failure creates a slower and lower energy event than a typical dam burst event, due to the lower amounts of water, the smaller elevation variation between the lowest point of the breach and the ground, and specific hydraulic conditions of canal outburst, which are explained in greater detail below. (Dun personal communication). The relative slope of the domain is low, at around 0.6%. However, a large depression can be observed in both the DEM (Figure 5.2) and is noted in the report that this topographical feature provides a major control for directing outflow. In comparison to the Glasgow test case, the slope of the underlying topography is shallower. Despite the level of flow from the initial outburst, and the extent of regions of low friction, transcritical flow is much less apparent in this test case. This was confirmed by evaluating the Froude values of initial model runs, and indicates that the ability to simulate transcritical flow is much less important.

The region is a combination of industrial units and warehouses, and residential buildings with associated amenities including shops, pubs and recreational areas. It

is noted in the report that many of the people who lived in the terrace houses along the main flow path of Eagle Street were both poor and elderly, and therefore had a high level of vulnerability to the hazard. The Coventry Hospital is located to the south of the domain, which includes an ambulance facility, to the west of the primacies, although located beyond the surveyed flood extent. According to the report, the hospital was flooded, although specific depths are not mentioned. Disruption was minimal, as only the basement, which contained a store of out patients records were kept, was flooded. The lack of impact is noted in the local newspaper which notes that less than 10 hours after the event, the hospital is back up to full function.

5.1.1. Inputs

As both Coventry and Glasgow are urban events, the parameter space was defined and discretized in a similar way and are summarised in table 3.2. The friction surfaces were again defined by OS data although data from 1978 has been used. As with the Glasgow test case, this was divided into road/path and non road surfaces, with appropriate levels of friction assigned. As with the Glasgow test case, a spatially distributed and spatially uniform friction value was compared. Other factors required greater consideration in construction.

5.1.1.1. Hydrograph

A key component with this model is the creation of the outflow hydrograph of the canal that represents the inflow boundary condition. Hydrographs were calculated based on the ISIS/British Waterways hydrograph outflow calculator, and the inflow hydrograph used for the test case is displayed in Figure 5.2.

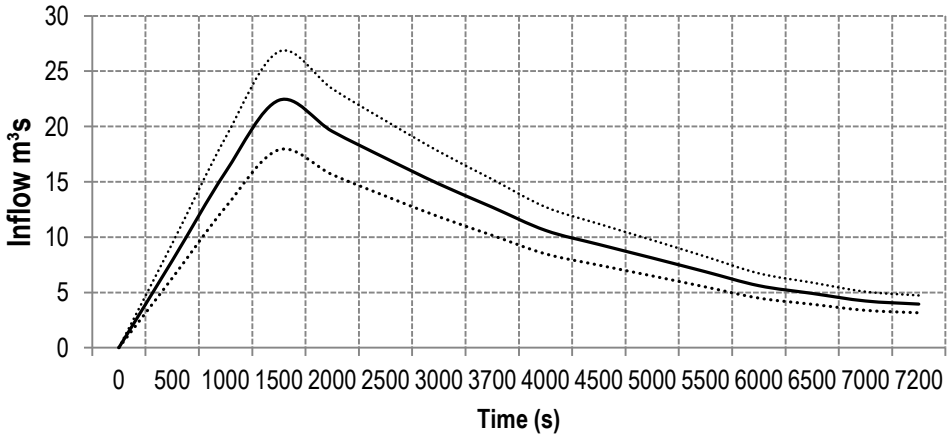

Figure 5.2. The Inflow hydrograph for the Coventry test case, with upper and lower uncertainty bands

This was developed as part of a nationwide report into the risk associated with events of this nature (Dun and Wicks 2013). A detailed study of outflow

characteristics was undertaken, using ISIS 1D models to simulate an embankment failure. Parameters which are particularly key to developing the volume of the hydrograph includes the width of the break, depth of water, length of pound (the length between two break points, such as locks, along the canal), and elevation difference between top of surface water and elevation below the canal. These values can be determined from the detailed survey plans of the event. These parameters were determined from the technical report, and as such it can be assumed that these parameters can be treated with a high confidence level and in a deterministic fashion and are described in Table 5.1. The most critical parameters in determining the shape and peak of the outflow are the pound length, which determines the volume of water available and the evolution time, which defines the phase from initial outflow to peak value).

The value chosen for evolution time to peak outflow represent the minimum value of this parameter allowed in the calculation, and represents the sudden nature of the failure. The initial value for the length of pound represents the distance between the break and the end of the Coventry canal to the south of the embankment failure point. Further values such as roughness are based on the material of the embankment and were determined from the Engineers reports, which include borehole data for the region. The report also provided the input for channel width, pound depth and constriction width. Variations in some of the parameters were considered, but affected the hydrograph less than the predefined uncertainty bounds in the hydrographic calculator. Therefore the parameters below were used, with the knowledge that uncertainty in the parameters would be with the range of the uncertainty boundary.

Pound Length(km)	12km
Roughness (Manning's n)	0.02
Evolution Time to peak (mins)	25mins(minimum value)
Channel Width(m)	7m
Initial Pound depth	0.6
Breach location	0.1
Constriction distance	0
Constriction Width	13m
Embankment Height	1.2
Embankment Material	Sand/Clay

Table 5.1: Input Parameters for the BWI Hydrograph calculator, for the 1978 Coventry event

The calculated input included a peak outflow value of 21 m^3/s, with an evolution time of over 6 hours (the time from initial outflow to a near zero levels of outflow). This is broadly in keeping with the description of the event from local reports.

5.1.1.2. Digital Elevation Model

In order to replicate the 1978 event, an elevation model of the 1978 topography and urban layout is required. A comparison of 1970's OS data and modern day reveals a considerable amount of building work and removal has taken place including the road network which has also been considerably reworked around the hospital and centre area of the domain. In order to create an accurate replication of these features, the available LiDAR data, which was captured in 2008, would need to be reworked to the 1978 topography. Assuming that the underlying surface topography is similar or variation is negligible, the present day LiDAR was sampled at points that did not correspond to present day building layout. 450 points were sampled, at around 2m resolution, which were then used to create a new digital surface model, by using a linear interpolation method between the sampled points.

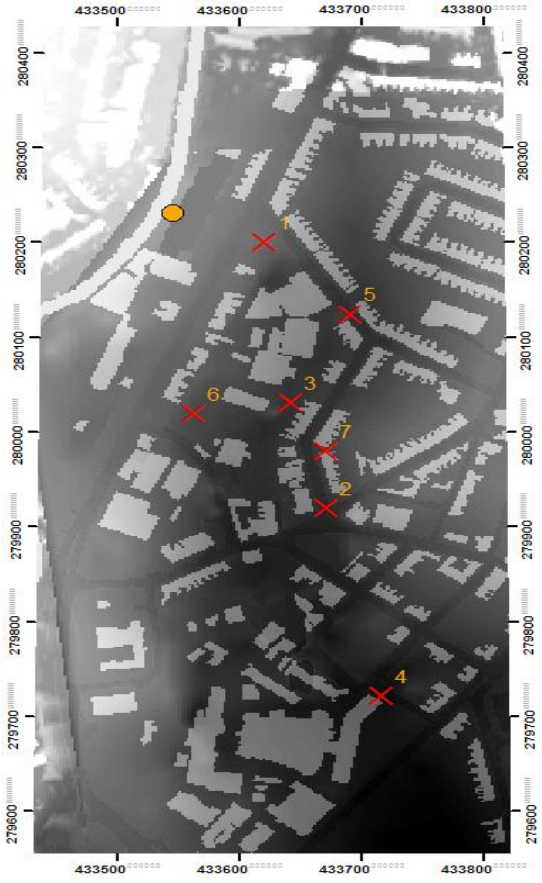

Figure 5.3: DEM of the test case, with location of control points, used in the analysis section of this chapter

Using the 1976 OS map, the road network was digitised and elevation within the footprint was dropped by 10cm. Finally, the 1976 building network was also added to model the building block method, using the location of building footprints to increasing elevation by 6m. This process is treated as a discrete process, and is not

included in the uncertainty analysis. Further analysis could be made of the point sampling and the creation of the digital surface model, to determine the uncertainty associated with this process, but given that this is a unique test case that requires an approach not typically undertaken in most modelling exercises, analysis of this pre-processing step may not have much bearing on future modelling processes.

5.1.1.3. Exclusion of factors

In order to reduce the computational cost of the uncertainty analysis, a number of initial models were realised with variations of physical boundary conditions. The representation of the breaching point was tested as multiple points, a 13m gap, and a single point input through which the estimated hydrograph was tested. The results indicated that this made little variation to either extent or to water depths, so was discounted from further analysis. The southern boundary was also investigated with variable conditions. A free surface slope boundary and fixed water depth boundary were investigated. These were found to influence the depths near the hospital, but not the wider flood extent, or global model results. The fixed water depths were also found to cause a small backwater effect at these locations, which would be counter to the engineering report that determined that once the water had reached the nearby pond it quickly moved through natural water courses. Given the subjective nature of the water depth boundary, and the potential computational issues with using a fixed water depth boundary, this was also excluded from the wider uncertainty analysis, and set to a free surface boundary condition. The impact of boundary conditions could be included in further work. An additional factor that must be considered in this case is the use of a drainage network. A traditional assumption in the use of 2D modelling of urban domains, is that the drains are surcharged due to the quantity of water, and therefore do not require representation in the model. This assumption does not strictly hold for this case. However, examination of the report reveals that one of the initial jobs of the cleanup operation was the removal of clay material that had covered up the drains. As a result of this, the assumption that the drains do not need to be represented can be carried over to this test case.

5.1.1.4. Output

Model assessment was made by comparing the surveyed flood extent with modelled flood extent. The surveyed extent is taken from a map based on a 1:1250 Land Ranger map, from 1975. The image was digitised, georeferenced and the flood extent extracted. The surveyed data was extended to include the pond that the flood waters reached, and then dispersed through. The model boundary was located at the southern edge of this pond, and this total area was used to compare model extents and evaluate the goodness of fit function, F^2. The outline of the event

is displayed in Figure 5.1, as the black outline. There is little detail in the report about the nature of the survey, either the techniques or the observed media (e.g. wrack marks or tidal water marks), so a certain amount of uncertainty is to be expected. Unlike the techniques used by Fewtrell et al (2009) for the Carlisle dataset, where accurate approximations to the uncertainty could be included in the model evaluation, this data set lacks the same detail with which to make an approximate approach to quantifying it. As with the Glasgow dataset, the need here is not to validate the model, but to use a global evaluation technique with which to describe model output for further analysis. A number of approximate depths are recorded in the engineers report. The depth of water at the junction of Eagle Street and Hanwell Road is noted as being 4ft, whilst houses near Hanwell square are noted as being around 3 feet deep, which provide an additional, if uncertain means of evaluating models. For this test case the depth of water at Hanwell Road is assumed to be accurate. A comparison between this and model output in this region is then used to provide another model evaluation technique. Whilst this could not be used to validate models, it does provide another means to gain insight into model performance and parameter interactions. Consequently, conclusions relating to this must be considered in this context.

5.1.1.5. Exposure Based Methods

The number of at risk properties and regions is considerable given the density of residential buildings, industrial units and the Coventry hospital complex. The key areas defined in this region for the local F^2 calculation are displayed in Figure 5.4b

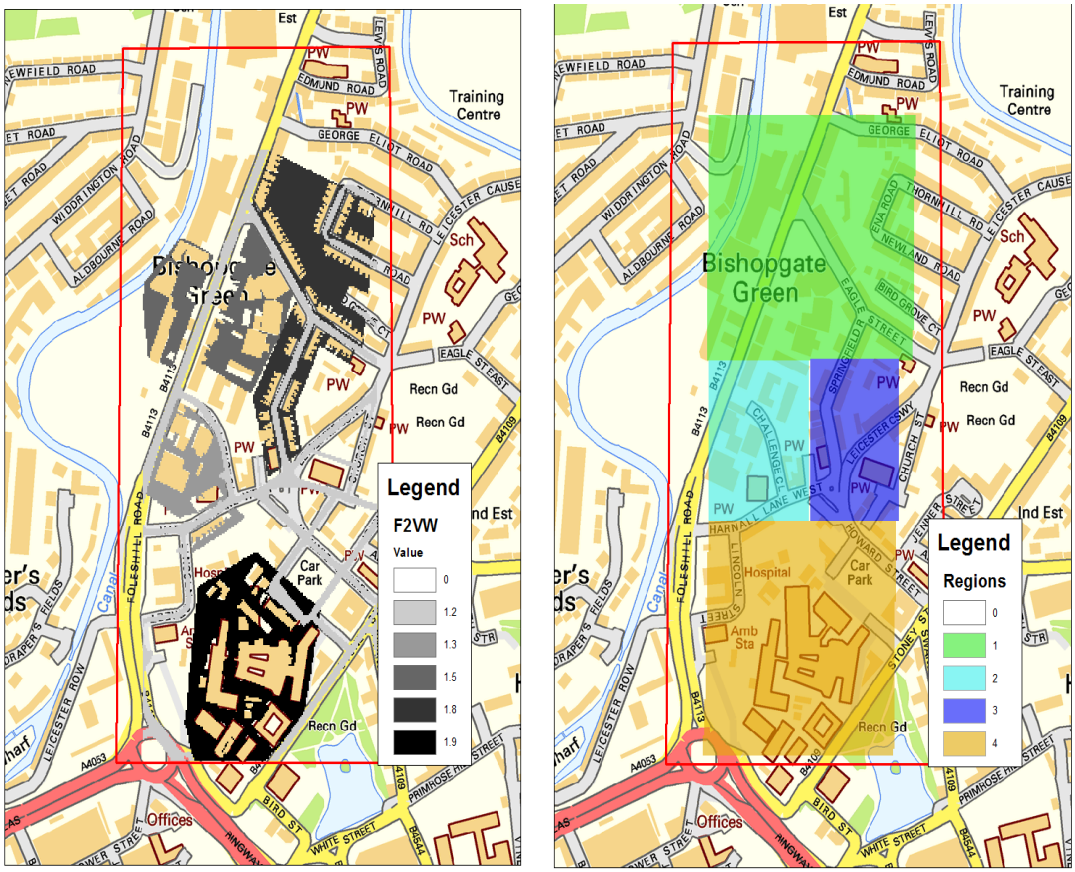

Figure 5.4: Part a, left, is the cells and weighting values for the vulnerability weighted approach (F^2 VW) and part b, right, is the regions used in the Regional F^2 approach

Using the engineering report, and noting that the event took place at 9P.M, the weighting cell approach has been weighted in favour of residential buildings, and the hospital buildings. The weighted values assigned cells are noted and displayed in Figure 5.4a, which also displays the distribution of these cells. Greater weighting has been applied to the hospital region. The large numbers of residential buildings are also applied with a high weighting value, not only due to the potential vulnerability of the residents, but also as a result of the proximity to the outflow point. This highlights the advantage of using such methods, where the nature of the event and region can be included into the evaluation process to improve how decision makers interact with the process. Further areas where weighting has been applied include warehouses and workshops. As with the school in the previous test case a lower weighting is applied to recognise that immediate hazard is low but that the long term impact of the damage to these units and to the region may be significant.

5.2. Results

This section begins with an overview of the hydraulic output of the model and a comparison of flooded extent and detailed analysis of water depths at 2 of the control points. This is then followed with analysis of the model function results, exposure based functions and sensitivity analysis. Each section provides an overview of how each module compares and varies and how the precise hydraulic output of the models contribute to the model results and uncertainty.

The model outputs indicate that the key aspects of the event, as reported in the engineers report, are simulated here. First, from the initial outburst a main flow path is created in a south easterly direction, down a nearby street. This is where the majority of damage was reported, and, at the south end of the street, where the water was at its deepest. From here the water move towards the hospital and out through the southern boundary of the model. Secondary flow paths emerge from the breach point, inundating the Foeshill road (the main road, running north south, Figure 5.1) and nearby residential and industrial units.

In comparison to the Glasgow case, this event created smaller flood inundation extent variations between the 4 module types, despite the presence of transcritical flows and complex topography associated with urban environments, as can be seen in Figure 5.5. For each module, a similar extent outline appears, with a maximum extent ranging from the junction of Foeshill road and Challenge close, around 200m away from the breach point and 75m in the opposite direction. The flow is well constrained through the centre section of the model domain. Here, a natural bowl formation severely restricts the width of extent for each module, and the results are in keeping with the events as reported in the engineer’s report. This region also contains the greatest water depths, which then move toward the southern boundary which represents the point at which the outburst waters returned to natural water catchment systems. Each module approximately captures this series of events, as demonstrated in Figure 5.5, and although the variation between each module is smaller than Glasgow, and the regions of uncertainty are less defined for each module, a number of key point still emerge between each module type.

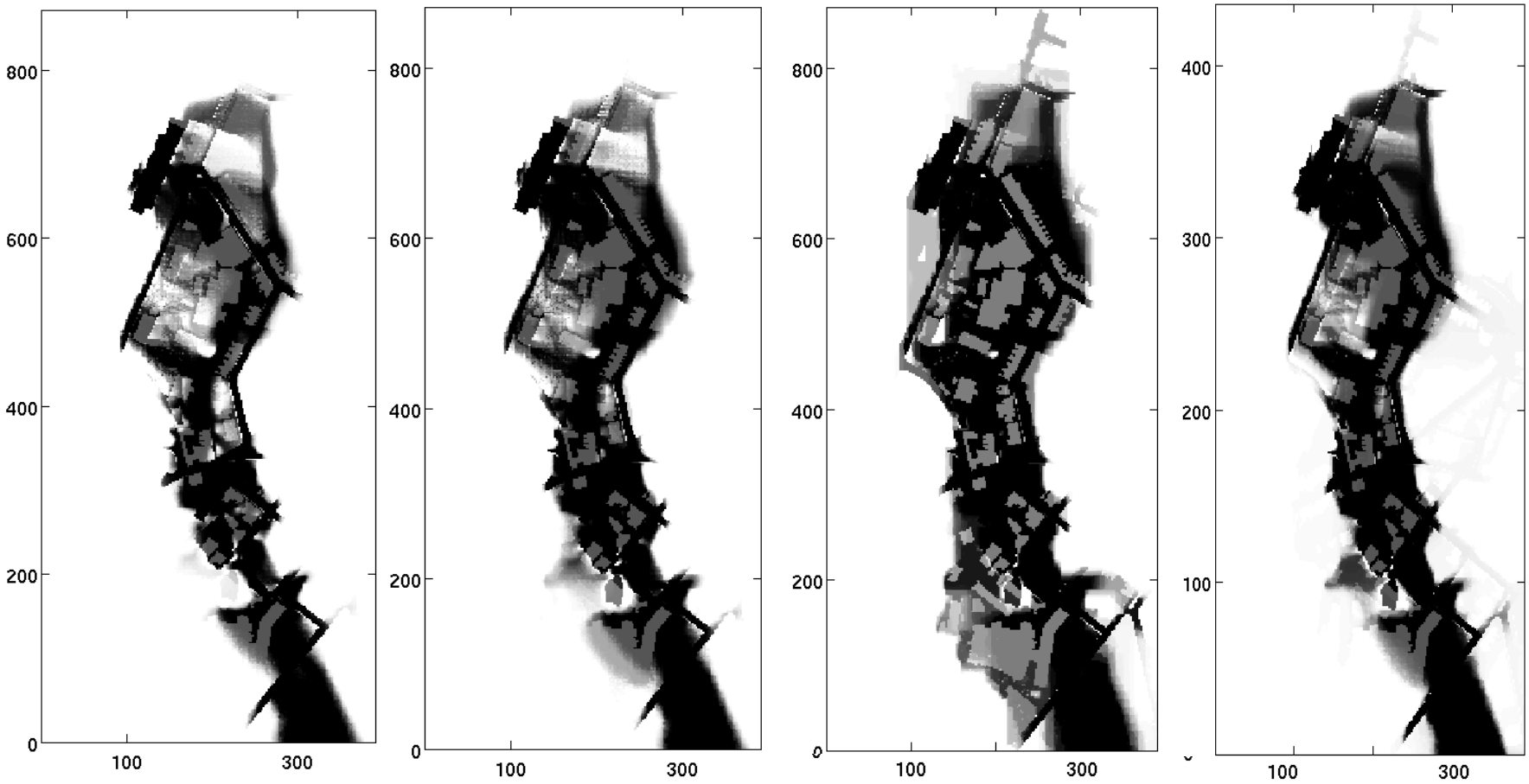

Figure 5.5: Uncertainty flood extent plots for the 4 LISFLOOD modules over the Monte Carlo results (from left to right ATS, ACC, Rusanov, Roe), where dark regions represent cells that have flooded in all model realisations.

A main variation between the full SWE models and the simplified approach is the northerly extent of the flow. For both the ACC module and the ATS module a smaller extent with low inundation frequency values are modelled. In comparison the full SWE modules create a wider flooded extent, with higher inundation frequency value. This region, lies outside of the surveyed flood extent, which suggests that the full SWE may over represent the processes of this type of event. Investigation into the causes of this indicates that the water flow in this region is low in depth compared to the main flow path (>0.6m). It emerges into this region through the Foeshill road (at 125m on the lower axis, 600m y axis Figure 5.5), and from a gap between buildings along the main flowpath. After the initial inundation phase, water returns through this same point and by the end of the model run has either returned to dry conditions or shallow water pools. This indicates that additional urban representation is needed such as drain systems, or that the surveyed extent may have been taken sometime after the initial flood when this region may have appeared dry. Further insight into this can be achieved by a comparison of F^2 values, which will identify how critical this is to overall model performance and which is undertaken in Section 5.3.

The centre section of the model domain (around 100 to 200m and 500-600m on Figure 5.5) contains a number of industrial and residential units which are critical in determining exposure to risk for this event. The engineers report indicates that the majority of units were affected by the flood waters. Here, the flooded extent is most uncertain for the simplified modules, in particular for the ATS module, where a number of cells value equal to zero, indicating that no flood waters were encounted in this area, even at high values of the inflow boundary conditions. This provides insight into the relative performance of both the module choice and the hydrograph, where the reduced physics approach of the ATS module will potentially undervalue the hazard associated with the event as a result of its inability to overcome urban obstacles.

A number of points emerge from the comparison of extents; first that overall module sensitivity is low, indicating that in determining the extent, the choice of module is relatively insignificant. Secondly, the smaller regions of uncertainty also indicate that other module choices are, as with Glasgow, relatively unimportant in determining the flood extent. This indicates not only the complexity of modelling outburst events, but the significance of extent as a means of evaluating risk and therefore models. It also suggests that the underlying topography is a key control on model results and the sensitivity of other factors.

5.2.1. Analysis of Water Depths

Comparison of the outputs of the control points provides a greater insight into module performance and for some of the other input parameters. Evaluation of the Roe solver results indicates that supercritical flow occurs along the main flowpath for the initial phase of inundation. After this, lower Froude value conditions (>0.3 <0.9) prevail for the majority of the control points through the model runtime. A variation to this is for the control point located at the southern end of the main flowpath (CP-2, Figure 5.3). The control point is located central to the main flow, and remains in a state of supercritical flow throughout the model domain with Froude values of 1.3 for the majority of the model runtime. The spatial and temporal variation of the critical flow level indicates that the correct representation of this process may not be essential for an accurate simulation of extent. This leads more into how critical representing transcritical flow is to the efficient modelling of urban environments. Further analysis of water depths will provide more detailed answers.

5.2.1.1. Control Point 4

Figure 5.6 is a comparison of the water depths for the modules at CP-4, across the whole parameter range. At this point the water flows past the hospital into a confined part of the basin (around 300m and 200m point on Figure 5.3). A number of points from the analysis of this control point provide insight into the significance of the module choice on outputs. First, the time from initial inundation to peak value is relatively short, around 2 minutes across the model results for each parameter space, but varies considerably between each module. For the simplified approaches, the time to peak value from arrival is very short at around 100-200 seconds, where as the full dynamic wave models are around 200-300 seconds.

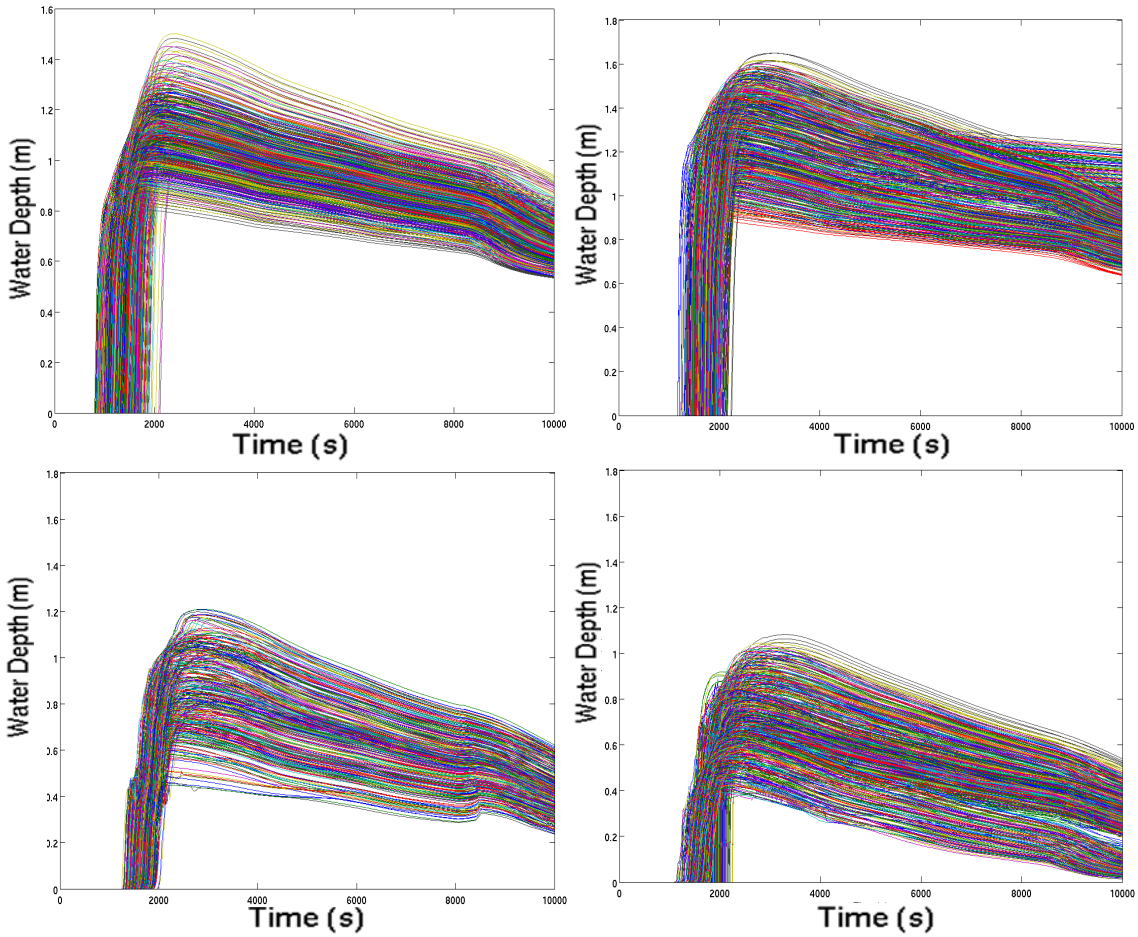

Figure 5.6: Water depths for the 4 modules over the range of tests at the control point 4, all plots scaled to 0 and 1.8m across the entire model run time. (from top left to bottom right ATS,ACC,Rusanov,Roe)

This is related to the ability of the full modules to overcome critical obstacles and to inundate a larger area of the domain, which creates a longer inundating phase for the full SWE modules. This is similar to the uncertainty flood extent (Figure 5.5), and is related to the formation of secondary flow paths in the centre area of the model domain, which creates a longer lag time, as water rejoins the main flowpath. The value of peak also varies considerably between the modules and this can be attributed to the ability to represent transcritical flow in the full SWE modules, whereby water is moved through at a greater rate due to the higher velocities associated with supercritical flow. Peak depth values for the Rusanov and Roe modules are significantly lower than the simplified models, around 0.8m lower, which is a product of both the transcritical flow conditions, where high velocities for water depths are encountered, and the ability to flood a wider extent, providing less water to the main flowpath.

The timing of initial inundation for this control point also varies considerably between each module. This is a critical aspect of determining the risk associated with this event, as the moment of initial inundation is essential. In the historic

event the impact of the event was high due to the sudden nature of the failure, and it is a critical output of the model. The ATS module produces the earliest arrival time, and the widest range of arrival times, in comparison to the other modules, which produce similar values. The ATS module mean value of 22 minutes compares to 30 minutes for the ACC module and 28 minutes for the full SWE modules. The indication here is that the diffusion wave approach provides a well defined main flow path that channels the water effectively through the model domain. The inclusion of additional terms from the governing equations reduces this defined flowpath by creating additional flood paths for the water to move across the domain leading to a long lag time. The increased level of physical representation also appears to create more robust solutions where uncertain parameters are used, hence the small range of arrival times. Further insight can be gained by a comparison of the 2 levels of physical representation (between full SWE modules and the simplified approaches). The variation between the two full SWE is minimal. Both modules produce similar range and mean value of arrival time, the first example across the range of tests cases that indicate that the choice of numerical approach is less critical than the inclusion of the full momentum terms. The variation between the two simplified approaches is more critical however, with the ACC module producing higher mean, minimum and maximum values. This is in part, a product of the increased levels of physical representation allowing it to overcome some of the obstacles in the model domain. It fails to maintain the same level of momentum as the full SWE, causing a slight increase in the timing of the flood wave. The increased level of representation relative to the diffusion wave ATS approach, also appears to increase the drying phase of the model, evidenced by the lower rate of water draining in Figure 5.6, which is also critical to the timing and rate of water flow across the domain.

Each module however produces a wide range of values for the arrival time, with minimum arrival times of 13 minutes for the ATS module, 19.5 for ACC, 18 for Rusanov and 17 minutes for Roe compared to maximum values of 35 for ATS, 41 for ACC and 38 for Roe and Rusanov, indicating the choice of other factors and parameters has a significant influence on the timing of flood waves.

The Spread of arrival times and the final depth levels provide information on the significance of the other factors within the model sample space. The wide range of arrival times for the ATS module indicates that this is far more sensitive to other parameters. This has not been seen in relation to the flood extent, which is generally insensitive to the range of parameters tested in this study. By comparison, the full SWE modules produce a more consistent range. This indicates that where full physics modules are used, the relative uncertainty of

other factors become less critical in affecting the level of risk associated with the timing of inundation, and are therefore more robust to the influence of uncertain parameters. The arrival time is only an indicator of the hazard associated with the test case. Further impact of other factors can be seen, however in the final water depth level for the full SWE modules, and the peak value. A wide variety of peak values are modelled from 0.4m to 1.2m for the Rusanov solver, whilst the Roe solver produces a lower maximum value of 1m. Water is then slowly released across the duration of the model run. The spread of results from peak to final level is similar for both models, although the Rusanov solver creates a narrower final range in comparison to the Roe solver. It is also clear that the Roe solver moves to a distinct drying phase with very low levels of water value present at the control point by the end of the model runtime. By comparison both the simplified approaches demonstrate no clear drying phase, and water remains pooled in this section for the duration. This indicates that conserving momentum in this test makes a critical effect on water flow, even if this not observable in the flooded extent values.

5.2.1.2. Control Point 2

Control Point 2 also highlights significant variations between the LISFLOOD modules. CP-2 is located at the bottom of the main flow path, through Eagle Street (Figure 5.3). Here the water is flowing from the main flowpath towards the narrow section of the natural bowl that occurs here and on towards CP-4. A number of aspects concerning the performance of each module are apparent here. The water depths for all the model results are displayed in Figure 5.7.

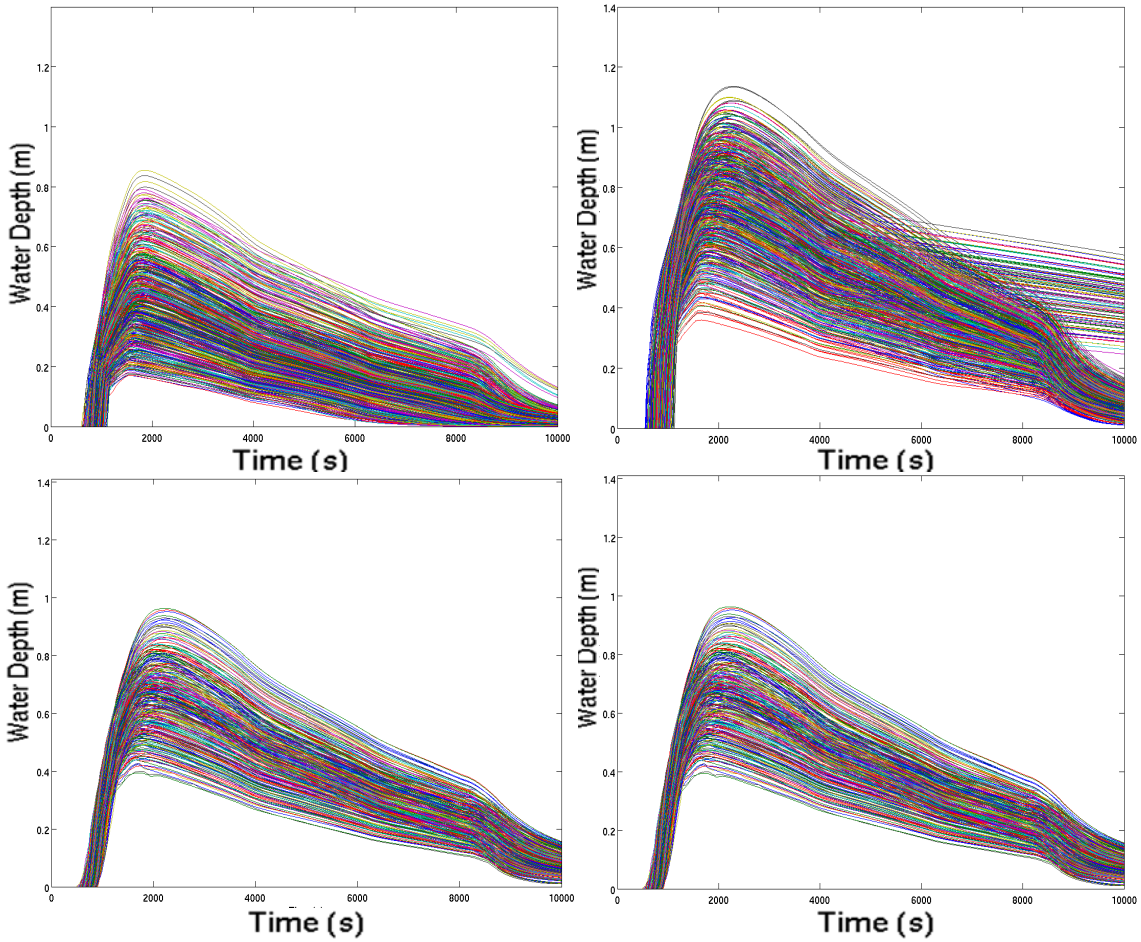

Figure 5.7: Water depths for the 4 modules over the range of tests at the control point 2, all plots scaled to 0 and 1.8m across the entire model run time. (from top left to bottom right ATS,ACC,Rusanov,Roe)

Again the similarity of performance between the two momentum based modules and the variation between the Roe and Rusanov solver and the two simplified approaches can be observed. The water arrives at around the 600 second mark for both Roe and Rusnaov modules, compared to 600-1000 for the ATS approach. The peak values are higher for the Roe and Rusanov solvers, between 0.35 and 0.9 meters, compared with 0.17 and 0.82 for the ATS module. In contrast the ACC module, whilst creating a similar arrival time variation to the ATS module, has a higher peak value and a longer arrival to peak time than the other three modules. The time to peak value is also significantly different across the three levels of physical representation. The ATS module produces the shortest arrival to peak time. This combined with the lower peak values, indicate that water is being moved quickly through this section, in comparison to the other modules. At this point the slope associated with the underlying topography is relatively high, which creates a greater flow rate in the diffusion wave approach and enhances the defined flow path created with the ATS module. This provides some further insight into the early arrival time of the diffusion wave module at CP-4. The higher level of

physics of the ACC module creates a longer pooling period, indicating that the inability to replicate supercritical flows whilst still maintaining the local acceleration term can restrict the flow rate value, which has additional impact on flood levels and arrival times at further points.

Other control points also reveal more about the impact of module choice. CP-3 and CP-7 have lower depth levels for ATS and ACC compared to the full SWE modules. For CP-7 the ATS module only produces a limited number of model runs capable of inundating this control point. These two control points relate to water depth in areas away from the main flow path, and indicate the importance of detailed local analysis of model results in determining model performance, and in evaluating risk levels for regions and sub regions. The combined analysis of control points indicate a strong spatial variation associated with modular choice. It also raises significant questions about the use of extent as a means of evaluating model performance, which are discussed in greater detail in the following sections.

5.3. Model Evaluation

Each model result was compared to the surveyed flood extent to produce a goodness of fit value. A summary of the main results for this and for other model evaluation techniques are summarised in Table 5.2. In comparison to the Glasgow test cases, the variation between the modules is significantly smaller, as demonstrated by the box and whisker plot of Figure 5.8, which is plotted to the same scale as the equivalent figure from the Glasgow chapter. As with the Glasgow plot, the box represents the 75 and 25 percentiles, the red line is the mean value, and the black lines the total range of data in the normal distribution. Outliers are represented as red crosses beyond the range of results.

Model(function)	ATS	ACC	Rusanov	Roe
F^2 (Mean)	0.635	0.618	0.550	0.601
F^2 (Max)	0.710	0.687	0.673	0.687
F^2 (Min)	0.511	0.464	0.402	0.332
Depth(Mean)	-0.004	-0.326	0.084	-0.108
Depth(Max)	0.303	0.002	0.956	0.652
Depth (min)	-0.352	-0.696	-0.463	-4.148
F^2 VW(Mean)	0.619	0.590	0.522	0.575
F^2 VW(Max)	0.706	0.646	0.644	0.322
F^2 VW(Min)	0.503	0.514	0.390	0.675

Table 5.2: Summary of Model Evaluation Results; F^2 , NS and Vulnerability Weighted F^2 (F^2 VW)

In the previous test cases, an increase in model performance relative to the module choice could be determined from the box plot figure. For this test case, there is less of a discernible pattern between the levels of physical representation. In fact, the opposite effect appears to be occurring. The ATS module has the highest mean F^2 value, and both the simplified physics models produce a higher mean value than the full SWE modules. This is perhaps not surprising in the case of the Rusanov solver, which produces a wider range of values, and therefore a lower mean value, but is interesting in the case of Roe solver, which tends to be a robust to a variety of flow types and uncertain parameters. The range of F^2 values for the ACC module is also the smallest, which confirms previous findings in this research about the general sensitivity of the ACC formulation to input parameters.

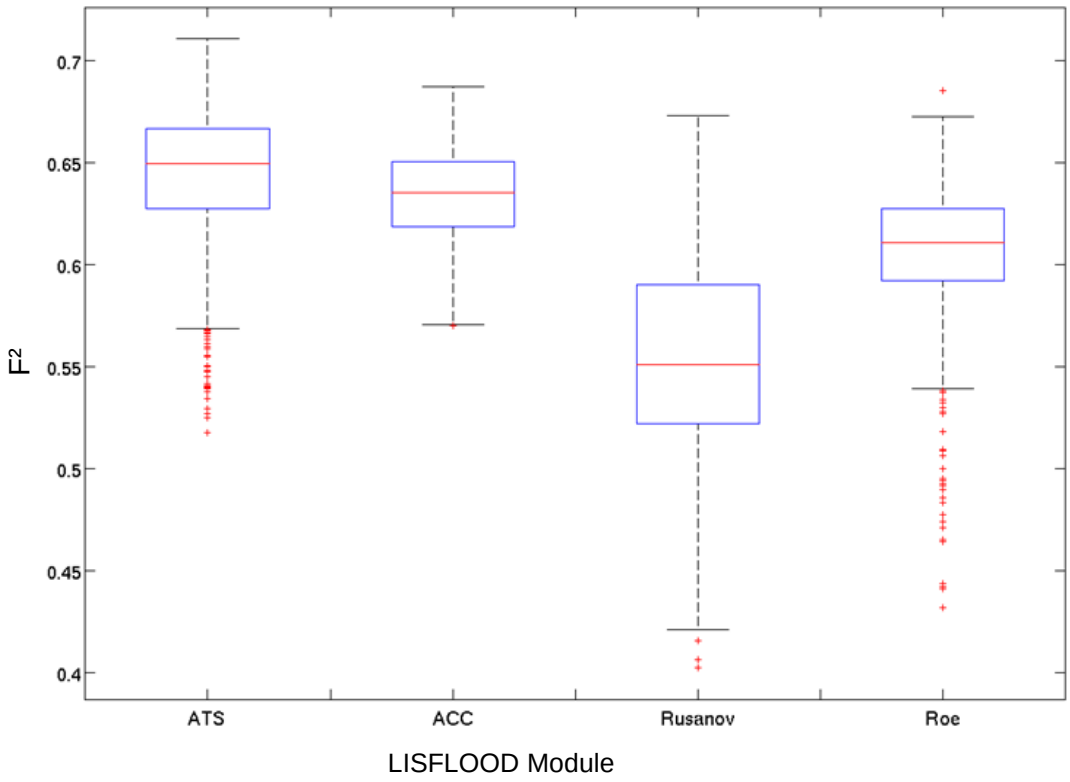

Figure 5.8: Box and whisker plot of the F^2 value against module choice

As with the Glasgow test case, the Rusanov solver produces the widest range of results, highlighting the potential drawbacks of using a simplified numerical approach to solving the full SWE, which is an increased sensitivity to uncertain parameters, leading to a larger model evaluation uncertainty boundary. This is despite the ability of the module to produce similar higher peak values to the other codes. When this is compared to the Roe solver, which produces a narrower range of results with similar maximum value, indicates that the Roe solver is robust to a variety of flow types and uncertain parameters. The approximate downward trend in F^2 value to the module choice is related to the wider flood extent modelled by the full SWE. This increase in extent leads to the number of incorrect model/observed values increasing, thereby

lowering the overall F^2 value. For all the modules a similar level of correctly predicted inundated cells is returned across the factor, indicating that the reason for the poor level of F^2 is caused by over prediction of flood extent by the full SWE modules in the upper region of the model domain. The ACC modules also floods this region, but not to the same extent as the full SWE modules. This indicates that the model may require additional representation of urban sewerage networks, in order to account for additional low level flow paths that emerge from the main flow path. The engineer's report states that the initial outburst contained a large amount of debris from the failed bank, which covered drains and blocked sewerage networks, but this might not be the case away from the main flowpath. Furthermore, the surveyed flood extent may represent only the main flooded region and exclude smaller amounts of water. The precise details of the survey were not recorded in the engineer's report and this information can be critical in making further assumptions about the precise nature of the flood, which can then be incorporated into the assessment process (Neal et al 2009).

The engineer's report detailed depths of water at locations in the affected region, which is reported as an approximate value. A single depth value is used to evaluate the models. Given the uncertainty associated with both the location of the measurement and the accuracy of the measurement, this value must be considered as an uncertain and approximate measure of model performance, due to the uncertain nature of the measurement. Here it is used to gain further understanding of model performance, and the overall results are summarised in Table 5.2 and variations between modules highlighted in the box and whisker plot in Figure 5.9

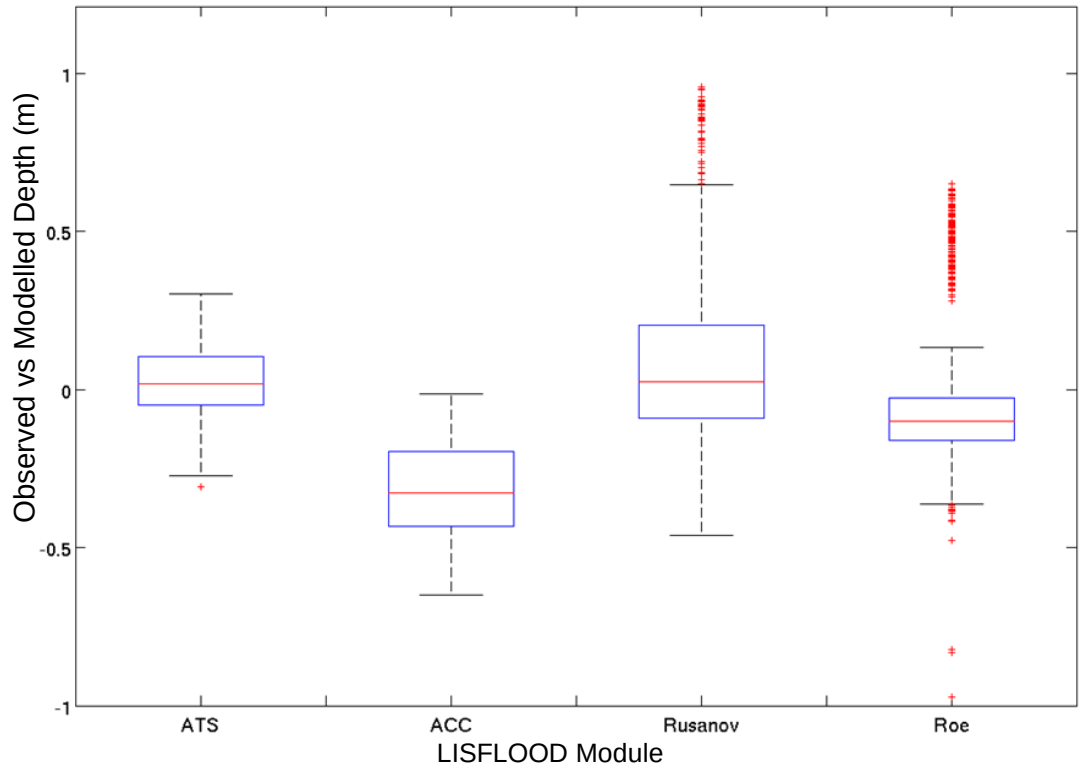

Figure 5.9: Box and whisker plots of the modular choice against the variation of depth

The value of *y* axis represents the difference of modelled and recorded depth for each module, where 0 is the same value as the observed depth, located in the centre of the *y* axis. The ATS module produces the lowest mean value and a narrow range of results, with the majority of its results falling within 0.31 meters of the recorded value. Both the full SWE modules also produce mean values close to zero, albeit with larger ranges. The large range of depth values from the Rusanov solver is consistent with previous evidence relating to the sensitivity of it to uncertain parameters. The contrast with the Roe solver is also consistent with previous results, in that the Roe solver provides a robust solution that is not adversely affected by the parameter range. A small variation between the mean values of Roe and the Rusanov flux, relates to the higher velocities of the Rusanov flux that tends to reduce the peak depth value. The ACC module consistently over predicts the depth, with a mean value of 0.32 above the observed data, a maximum difference of 0.69m and a minimum variation of 0.002m. This is in keeping with what had previously been observed at CP-2 where, the ACC module produces a higher peak value at a later time period than the other modules.

Using a similar approach as in the Glasgow test of comparing the two objective functions and creating a pareto front provides further information about how the different LISFLOOD modules behave over the different functions. The two functions are compared in Figure 5.10

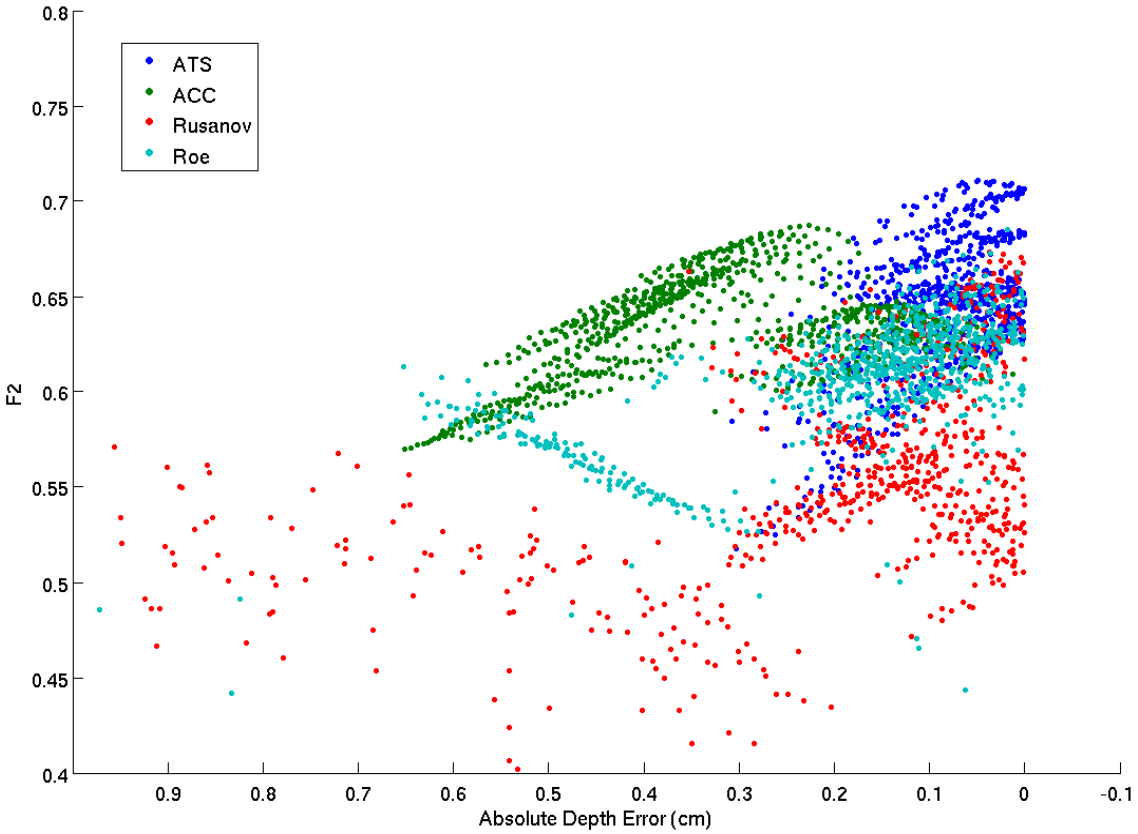

Figure 5.10: Combined results for F^2 (y-axis) and Depth Error (x axis)

The pareto front of best performing models across the two parameters is again located in the top right of the figure. The depth error is converted to an absolute depth error for clear analysis of this space. The patero front is dominated by the ATS model results, whilst the ACC produces similar F^2 values, but due to its over-prediction of depth, occupies a region behind the pareto front. Both the Roe and Rusanov models produce a greater spread of model results across the F^2 and depth functions. The implications here are that increased levels of physical representation allow the development of additional flow paths that create a greater variation of model results. These models therefore require a robust and detailed calibration process in order to reduce uncertainty. Within this aspect, the Roe solver provides a solution less susceptible to a wide range of inputs. The conclusions that can be drawn from these combined results is limited by the uncertainty associated with using a single approximate depth value, but it does emphasis that multiple model evaluation techniques enhances the calibration process and reduces the number of models that can be termed as adequate, if a GLUE methodology was applied (Beven et al 2006).

5.3.1. Impact of Inputs on Model Results

In order to evaluate further the impact of the parameter set on model results a comparison is required between the parameters to determine parameters with significant contribution to the model function variance.

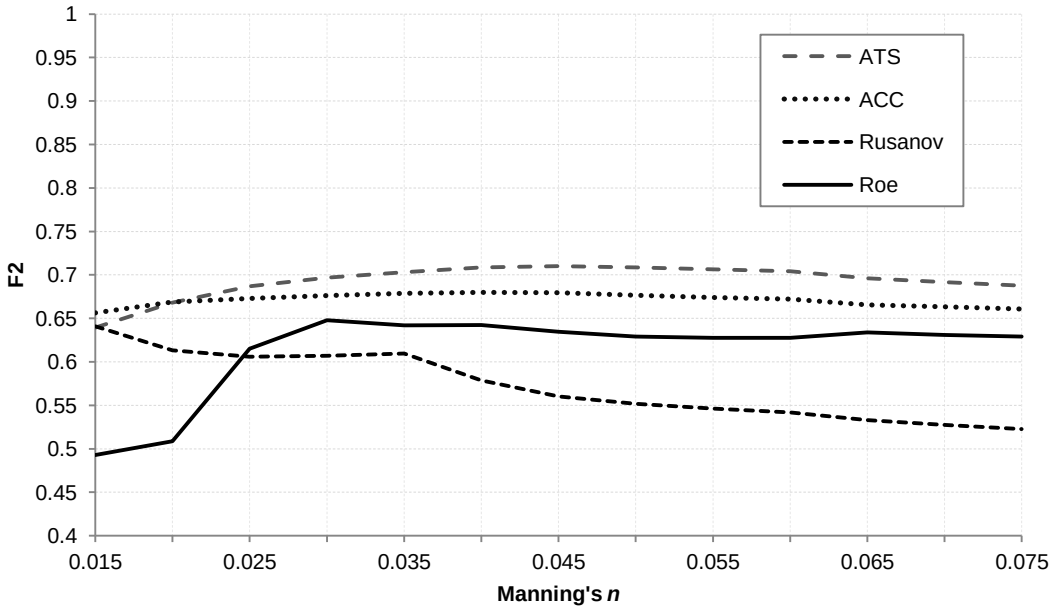

Figure 5.11: A comparison of F^2 for each module against each friction value.

Mapping the F^2 function over the parameter space of Manning's friction value and each module, provides insight into the relative influence of the parameters. The resultant graph is markedly different from the previous graph in the Glasgow test case (Figure 4.8). Here the order of the line is reversed, with the highest peak of the value belonging to the ATS module, whilst the lowest relates to the Rusanov solver. A variation from the Glasgow case is that, with the exception of the ACC code, all modules here show some variation relative to the friction parameter. The sensitivity of the Rusanov solver to friction values is clear from the variation in gradient across the X axis, with a low value of 0.49 occurring with a friction value of 0.07. There is a general trend for the F^2 value to decrease with increasing friction value when using the Rusanov solver, as increased resistance reduces the water flow rate through the main flow path which appears to increases the simulated extent. Two key points can be determined. Firstly, it should be noted that the variation is much smaller than with Glasgow, with the range being within 0.2 from maximum to minimum F^2 value, in comparison with the Glasgow range of 0.38. The implication is that the topography and the hydraulic conditions of the event provide the greatest controls on model results, and that the smaller variation between module choice and friction choice relative to Glasgow indicates the relative low importance of the model factors and inputs. Secondly, the level of physical representation appears to have a non linear

impact on model results. The Glasgow test case demonstrated an monotonic relationship between the output and this factor, whereas here the direct relationship between module choice and output is less clear. This can also be determined from Figure 5.10 This implies that the choice of physical representation is complex and not as clear to determine

The F^2 interaction matrix plot (Figure 5.12) further confirms this. In this figure, the effect of a parameter on overall model results can be visually assessed, and is the equivalent of Figure 4.10 from the Glasgow test case chapter. This figure has been scaled to the same level as the Glasgow test case to highlight the relative lack of variation between parameters.

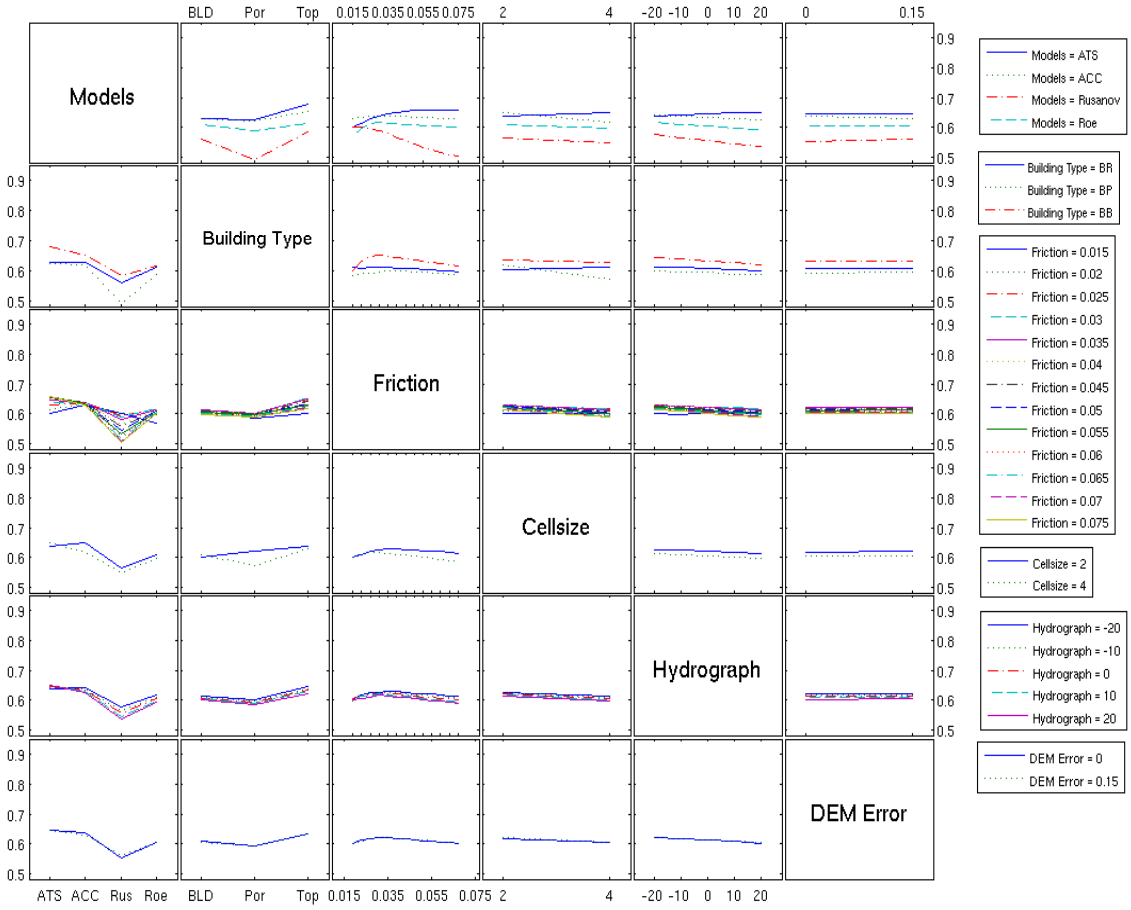

Figure 5.12: Interaction plot for each input factor comparing mean value of F^2 per level of factor other levels of factors.

As with the Glasgow test case, the choice of module appears to be the most influential factor based on the gradient of the line, with a clear downward trending gradient of mean F^2 value towards the full SWE (as the line moves from left to right across the rows of column 1). The gradient of the lines in the first row and first column compared to the other rows and columns highlight this, and again confirm the significance of this choice to model results over other factors. The choice of building

representation and the choice of friction value also have an influence on model results. In the previous test case, the influence of friction proved to be lower in relation to other factors, although the overall influence on the results was greater, implying low interaction but more critical first order influence on results. The influence of both these factors on the Rusanov solver can also be seen. The use of a porosity model produces the lowest level mean level for the Rusanov solver (column 2, row 1), which is not the case for the other solvers, which produce a similar mean level for the BR method. The influence of increased friction values and reduced F^2 level can be noticed, in the column relating to this factor and it appears that this accounts for the majority of the influence of the friction value.

A level of interaction can also be observed between the three key parameters. With the use of the Building block method the influence of the friction value increases, with a distinct peak at the 0.02 level, (column 3) which produced the highest F^2 level for all samples. This can also be seen between the building representation and the module choice, which produces a wider variation between the module types. The implication is that factors that affect the flowpath either through increasing resistance or defining the point appear to have the greatest influence on model results.

An interesting aspect is the relatively low influence of the hydrograph and cell size. Both factors display little variation in gradient in either the related row or column in comparison to the other factors. The significance of grid resolution to urban inundation modelling has been shown to be critical to model results (Fewtrell et al 2008, Sampson et al 2012). The levels of grid resolution for this test case provides a useful range by which the criticality of this factor can be tested. The fact that the influence appears to be low relative to other factors suggests that this measure is less critical when transcritical flow and large volumes of water are present, such as in these events. A variation that was noted in the previous test case, the influence of cell size on ACC model, is also noted here, although not to the same level as had previously been noted. The range of hydrograph levels also prove to be less influential in affecting the modelled extent. This point to other factors providing a greater control on the model results, implying not only that the uncertainty associated with the hydrograph is less critical in functions that evaluate extent, but that the use of extent as a means of evaluating models may critical underestimate variations in depth. Both the low sensitivity of the F^2 measure to hydrograph, and the higher level of sensitivity demonstrated to the cell size point to the influence of underlying topography as a controlling influence in model performance, particularly where it provides an important factor in controlling the formation of flow paths. This is emphasised by the significance of the choice of building representation, which also influence the effect and definition of the flow path.

5.3.2. Distributed Friction Values vs. Single Friction Values

As with the Glasgow test case, the method of how to model Manning's friction has been split between a distributed friction value and a single value friction for the model domain, and the results from this test case indicate similar findings to the previous case. Figure 5.13 is an interaction plot for distributed friction (top) and single value friction.

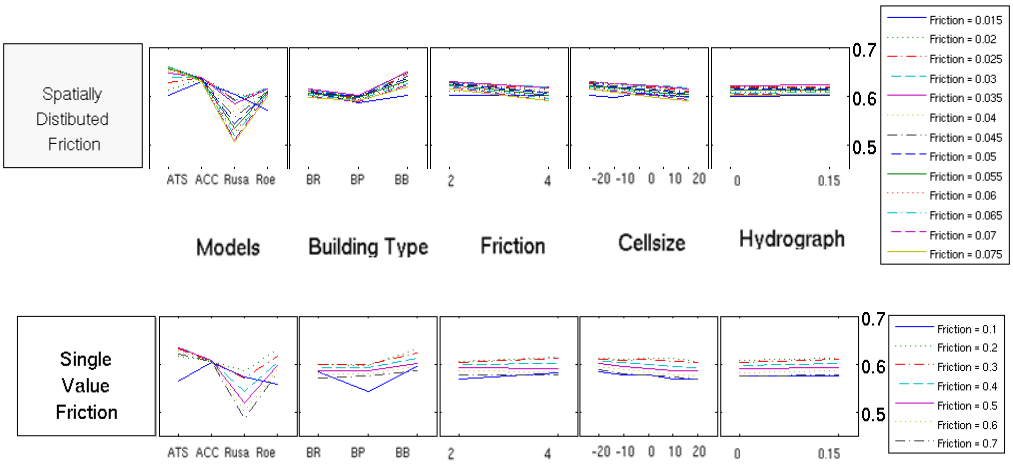

Figure 5.13: Comparison of mean level per parameter for spatially distributed friction (top) and single value friction (bottom). The comparison parameter from left to right is model type, building type, cell size, hydrograph, and DEM error

As with the Glasgow test case, the spread of results is greater for the single friction value than the distributed friction value, with the mean level of model performance also higher for the distributed value. Both plots display the same level of interaction with other factors, although a variation can be noted with the choice of building representation and friction choice (second box from the left). The small variation is caused by the decrease in mean level performance for the porosity model in the lowest value of single friction (0.01), which counters the general increase for the other levels of friction. In comparison to the distributed value, the porosity method has a slightly lower mean value than the Building Resistance method. This implies that the choice of friction distribution method and value is far more critical when using a porosity method, and that low friction surfaces which are typical in urban areas must be explicitly represented when using this building representation approach. The overall trend observed for friction values in this case and the Glasgow case are similar, in that the use of distributed friction appears not only to reduce the range of model results, and therefore reduce uncertainty, but also increase the overall mean result. For the Glasgow test case this may have been causal as a result of using a benchmark model which used distributed friction values. With the same results displayed here, the indications have greater emphasis.

5.3.3. Evaluation of Timing of Initial Inundation

The use of extent as a means of determining risk and differentiating between models may underestimate the consequence of the event, and ignore other critical aspects of the flood. In this test case, the timing of the flood wave moving through the domain is a significant model output. Evaluation of this provides details of the significance of parameters on model result. The influence of the parameter set on this output can be clearly seen in the interaction plots relating to the time of initial inundation which has been used to assess the model choice. Figure 5.14 is an interaction plot for the timing to reach the out reaches of the hospital complex (CP-4, Figure 5.3).

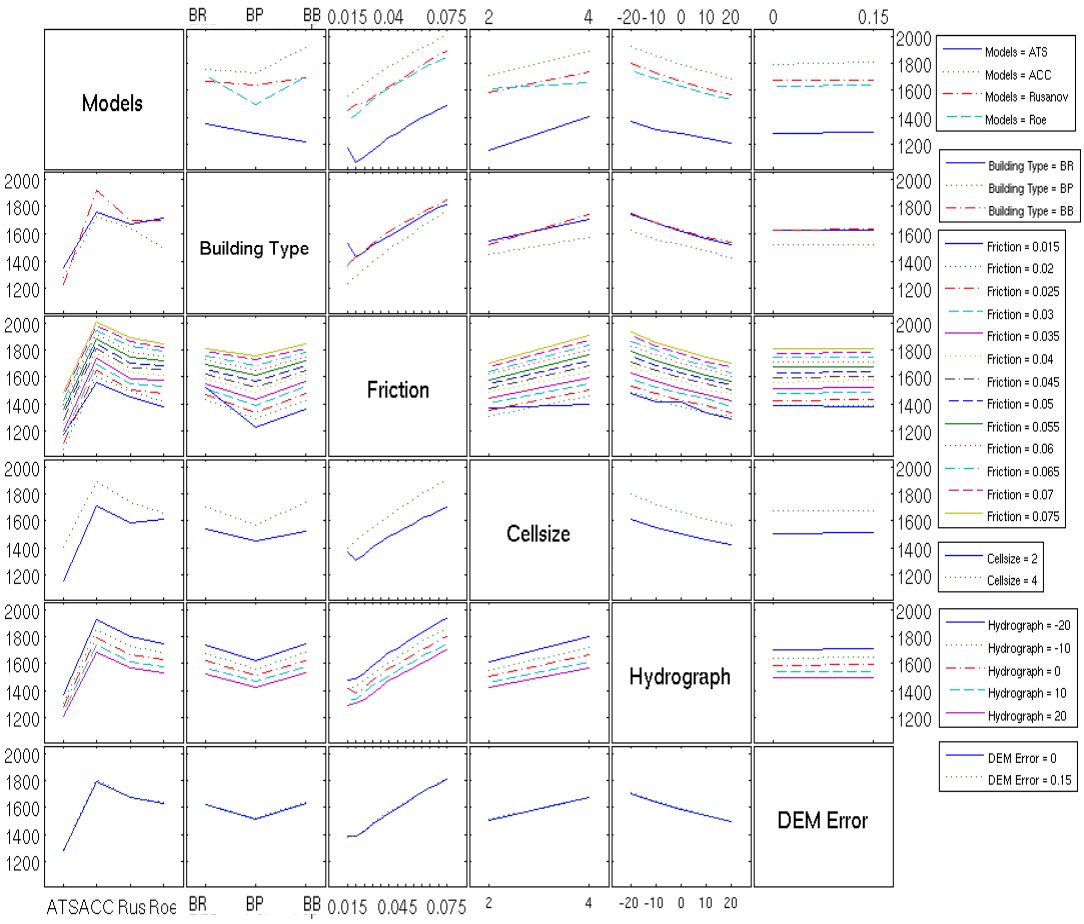

Figure 5.14: Interaction plot for each input factor comparing mean value of time of arrival for each level of factor against other levels of factors.

Here a much greater variation can be seen between all inputs and between all levels of factors, compared to the measures of extent, However there is little interaction between the inputs. Friction appears to be the most critical factor in determining arrival time, with a significant increase in arrival times across the range of times, and a steeper gradient across the friction levels (column 3). The choice of module also appears to be a significant choice and the variation between ACC and ATS, and ACC

and Rusanov and Roe modules can be seen. The first row also highlights the large variation between the ATS module and the other codes, and also suggests that module choice is most significant for describing variation in inundation timing. In comparison to the F^2 interaction plot, both cell size and hydrograph have a clear influence of the arrival time of the water. As the grid resolution is refined the arrival time increases, and as the hydrograph level increases, the arrival time decreases. The use of a porosity model also appears to influence the model results, creating a shorter time period to inundation than topography and building resistance method. The level of building representation and module choice represents the only evidence of parameter interaction across this measure, (column 2, row 1). The Roe model is particularly affected by the porosity model, producing a far quicker return time than the other modules which remain at similar time levels to the building resistance method. Considering the general robustness of the Roe solver to a variety of uncertain parameters in comparison to the Rusanov solver, as has been seen in previous analysis of control points and extents, this represents a significant sensitivity to building representation. Whilst this is less critical than the choice of friction it emphasis the need to consider building representation techniques cautiously.

The variation between the interactions plots for time and F^2 emphasis the restrictive qualities of the F^2 method for calibrating the model. The variations in times between the parameters sets are significantly larger than for goodness of fit. This may be critical depending on the future use of the model and the need to which decision makers may apply the simulation results. In this case an ensamble approach may be more appropriate to account for the large uncertainty associated with the timing of inundation, but emphasis the need for more detailed model analysis. As has previously been seen the non linear impact of level of physical representation on model results can be determined from the gradient of the line in column 1. This implication again is that determining the level of physical representation is not a clear decision and requires consideration of all model outcomes before determining an appropriate approach to represent the governing equations in the model code.

5.4. Exposure Based Analysis

With the relative insensitivity of the F^2 value to alterations in parameters, this provides an excellent case study with which to explore whether methods enhanced by incorporating vulnerability and exposure can improve the ability to discriminate between model and parameter sets. The first method used in the vulnerability weighted F^2 method, summarised in Table 5.2, and is compared to the standard approach using the same objective function modelled over a parameter space as Figure 5.11.

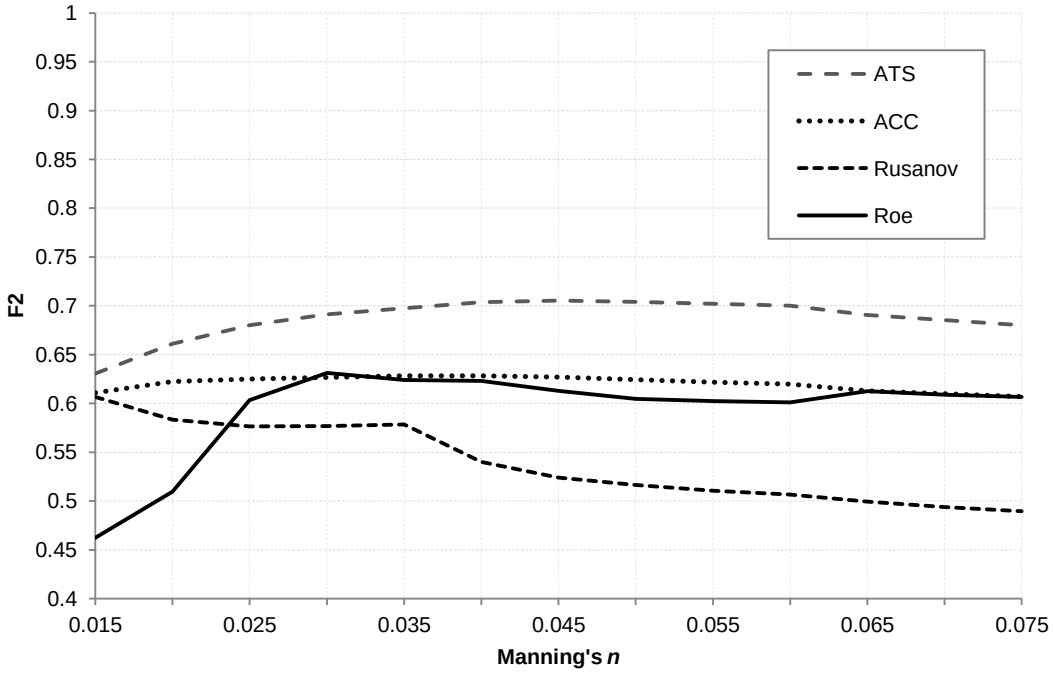

Figure 5.15: The F² VW value plotted against Manning's n value for each module. The pattern is similar to Figure 5.8.

A clear variation from Figure 5.8 is the smoothing of the parameter space by the F²VW method. This however reduces the applicability of the method which seeks to increase the spacing between models. Conversely though it also indicates that the choice of model is less significant when considering risk in the evaluation process. Over all the test cases, the general trend is a reduction in the peak value of the function, and a decrease in the lower values that enhances the lower value of the function from the mean value. This is an important aspect of using weighted methods, to enhance the peak values and produce a clearer analysis of higher performing models. The approach adopted here failed to refine the function over the parameter space. This is partly due to the high concentration of cells to represent vulnerable areas, which is a product of the size of the model domain and the nature of the domain. A more defined weighting system, using higher value produced a similar range of results, indicating that the precise values are less significant than the distribution of weighted cells in the model domain.

The regional goodness of fit values are summarised in Figure 5.16., where the far left column represents the global F² function, and each dot represents a model simulation. The colour of the model simulation relates to the module used in the simulation. The regions have been divided into sections of vulnerability, where the level of vulnerability can be approximately grouped together and can be seen in Figure 5.4b.

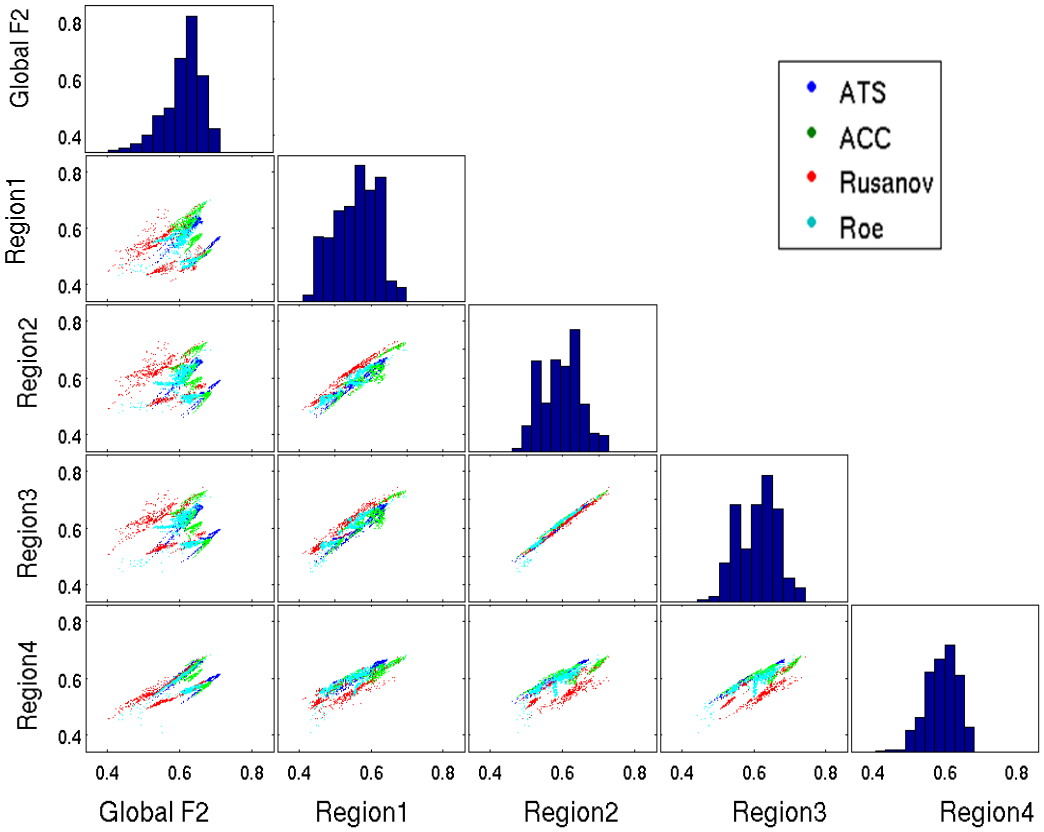

Figure 5.16: Comparison of regional values of F^2 versus Global F^2 value (first column). Each dot represents a model run, the colour relating to the LISFLOOD-module

The ATS code has returned the highest F^2 value, and this regional analysis allows detailed dissection of the precise regions where the F^2 value is greatly affected, as well as determining how critical module choice is in vulnerable regions. The overall trend is for a higher level of model performance and for a much closer range of model results for all the modules in the vulnerable regions. The second key trend is that each module, identified by the colour range, produces a narrower range of results for each subsection. The full SWE modules also produce higher values of regional F^2 than the ATS module. This indicates that the global F^2 value achieved by ATS is related to the fact that it produces a narrower extent with fewer incorrect cells, rather than an increase in the number of correctly matched cells. The wider range of values for Rusanov can also be identified from this figure. Each region produces a correlation between the other sub regions indicating that model performance is roughly comparable throughout the critical areas of the model domain. The mean goodness of fit value does improve in the lower sub regions (2-4), located near the boundary. The top region (column 2) has a lower mean and a wider spread of results. This region includes part of the domain where the full SWE modules created a wider flood extent, and the impact of this on the overall results can be seen. By comparison the ATS module does not contain the highest level of F^2 for this region, although it

was the only module not to produce the wider flood extent previously seen with the full SWE modules. The lack of correlation between the main global F^2 values and the sub regions indicate that for each sub region, there is a change in the general performance of each model, typically an improvement. This highlights that each module is capable of replicating the observed extent in key regions of the modelled domain, and that the global value is affected by regions which are less critical to the overall risk analysis of the model domain. It can also be determined from this figure that it is the regions closest to the inflow boundary that contributes more to output variation. This can be determined from column 1 where a comparison of region 1 and the global value show a wider variation that a comparison of region 4. This indicates that once flow paths have formed variations between modules and parameters reduce, and indicate that a critical region of the model is that closest to the inflow. Therefore a significant amount of uncertainty in the model is associated locally and with close proximity to this point in the model. In order to reduce uncertainty efforts should focus on this region of the model domain.

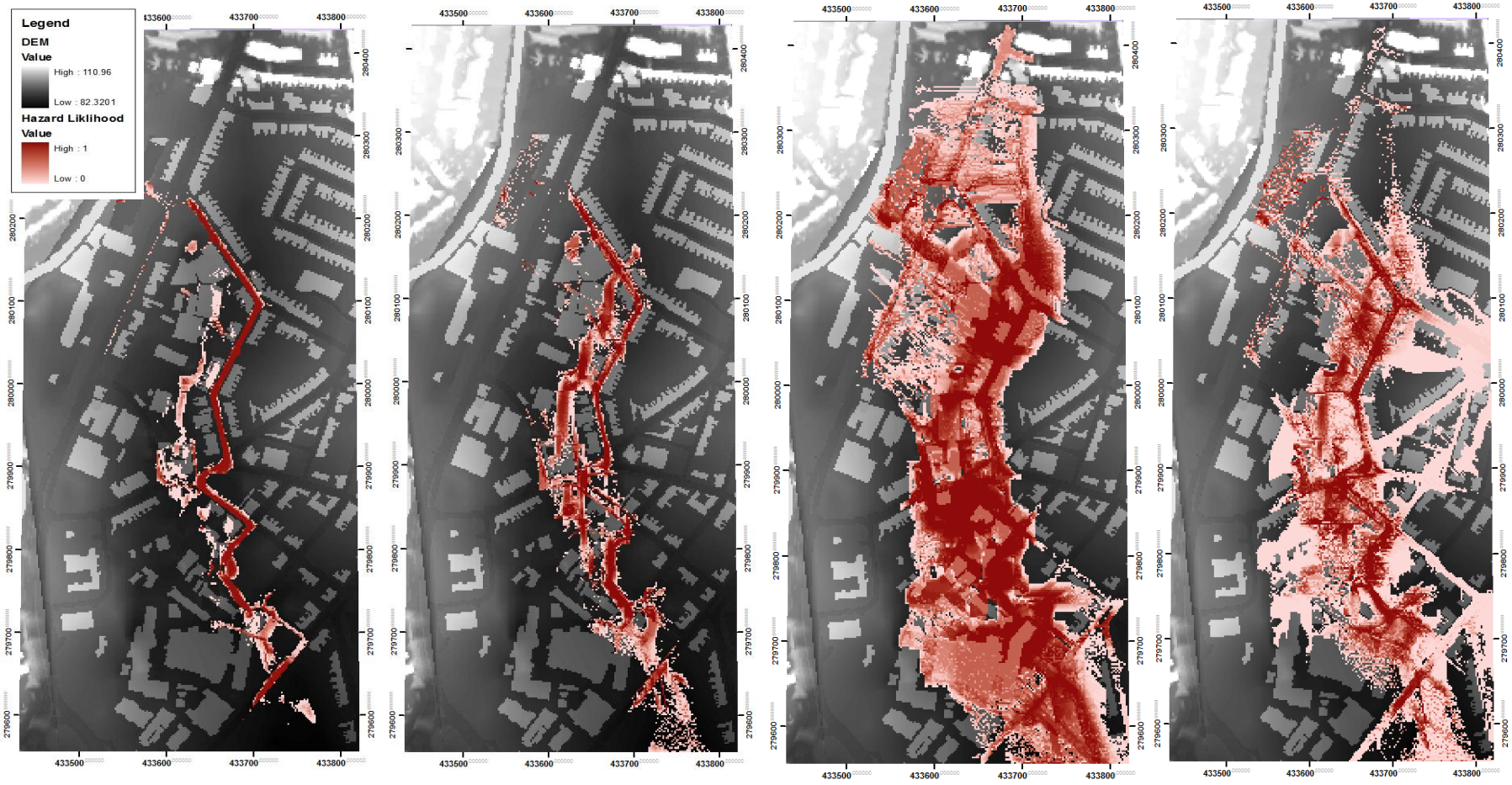

Figure 5.17: Uncertainty Hazard Plots, where red regions represent the high hazard areas, from left to right ATS, ACC, Rusanov, Roe, where dark regions represent cells that have flooded in all model realisations.

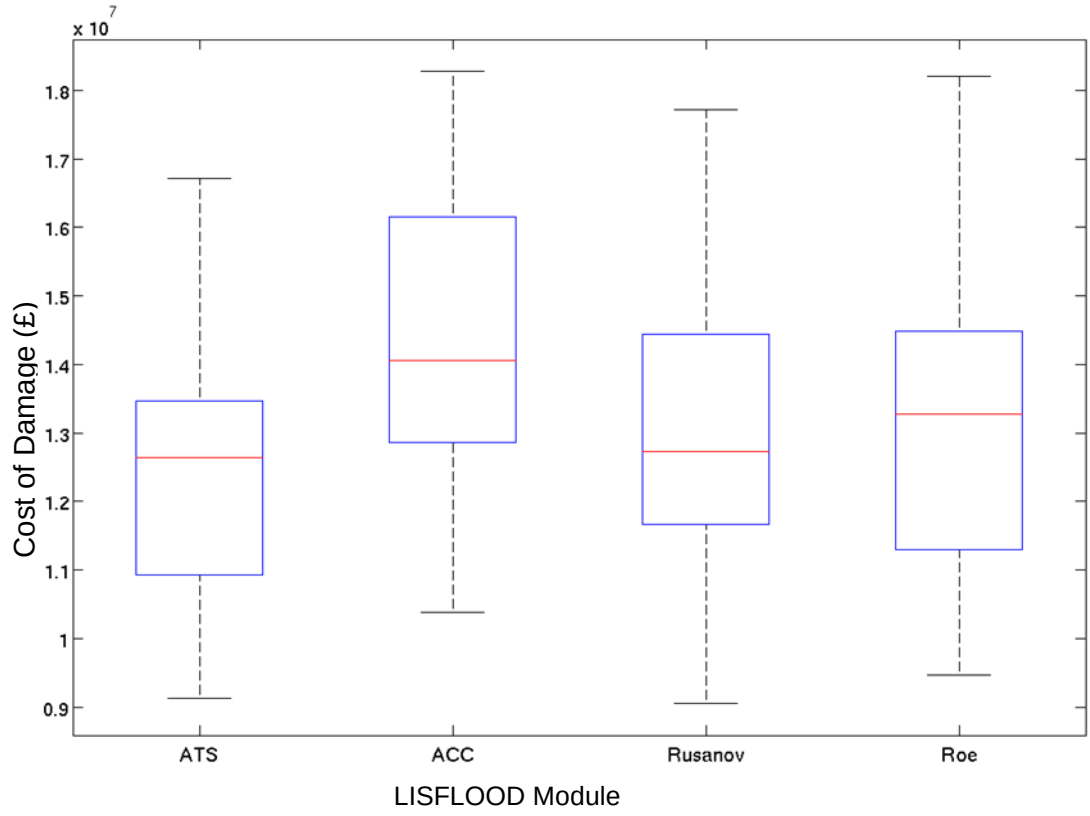

Figure 5.18: Box and Whisker plot of the Cost of Damage of the event for each of the LISFLOOD modules.

The use of Cost of Damage provides a useful insight into the modular performance and provides a method with greater discrepancies between model realisations. A visual summary of results shows a similar pattern between the modules that had been highlighted in the analysis of the control points. The greater depths simulated by the ACC module (Figure 5.6) here translate to a higher total cost, with a maximum total cost of £18.28million, with the parameters set of BB method, a hydrograph with 20% increase in volume and a distributed friction value of 0.065 for vegetated areas and 0.018 for low friction surfaces. The range of cost for the two simplified approaches is also similar, as is the range of costs for the two full SWE modules. The wide range of both Roe and Rusanov incorporate the entire range of the two simplified approaches. This is a noticeable effect of the range of depths simulated by these two modules, due to the wider ranges of extents, and indicates greater sensitivity to the other factors, a key component in a model evaluation technique. The previous case study saw a large variation between the grid resolution levels, which has been reduced by using a mask based on the number of cells around the footprint of the building, rather than a demarked area. There is still a noticeable step, but this is now a product of small variations between the grid

resolutions, rather than an issue with the method by which the cost of the event is calculated. A key feature for this method is to capture some of the factor influence and parameter interaction that has been captured in the standard model evaluation techniques.

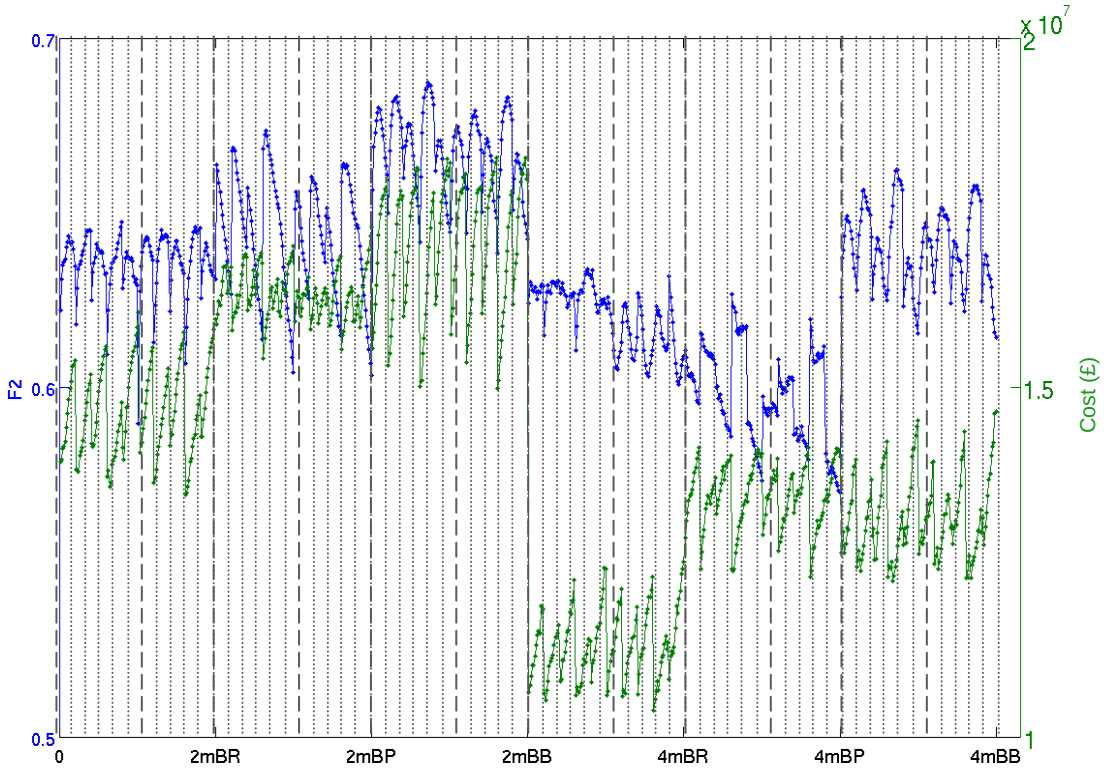

Figure 5.19: F^2 value (blue line) against Modelled Cost value (green line, right axis) for ACC module result. Each vertical line represents a different building representation method and cell size, going from (left hand side) Building resistance, Poroisty model and Building block, with the first 3 sections being 2m and the final 3 sections 4m grid resolution. The division of results is described in detail in the caption for Figure 4.14

Figure 5.19. is a comparison of the F^2 value vs. the cost of damage. Similar model response to parameters that are noticeable with the F^2 also appear here. Each function creates a series of small increases that appear as a range of spikes across this figure space relate to the increase in friction size. A clear variation can also be noted between building representation techniques. Each section marked out by the vertical lines represent a different cell size and building technique. A clear variation can be seen in both functions, but the variation of the cost of damage is significantly wider than F^2. This is confirmed when the cost of damage values are normalised, reflected the fact this method can provide a much clearer definition of model types and model performance. This combined with the ability to reflect general model/parameter response highlight the potential of this method to evaluate models, and the sensitivity of the approach to model uncertainty.

The hazard maps of Figure 5.17 demonstrates the variation of the regions of significant hazard between each module type. The ATS has the smallest area defined as being of high risk compared to the other modules. The main flow path is clearly highlighted in the ATS figure, and the hazard frequency value is high. This has been a particular feature of the ATS code, in replicating the high flow rate of this flow path with consistency over the parameter set. The ACC by comparison produces a much wider uncertainty extent, with evidence of secondary pathways forming adjacent to the main flowpath. This in part describes the variations seen between the ACC and ATS and can be attributed to the increase in level in physical representation. The wider hazard extent of the ACC solver also has lower hazard frequency values and is therefore of increased uncertainty. The main flowpaths and secondary flowpaths have high frequency values indicating that these are a consistent feature of the simulations involving the ACC. This is further evidence that increasing levels of physical representations impact results in key ways, and in identifying hazards and regions of high velocity. Both the Roe and Rusanov solver produce a much wider hazard extent region. The Rusanov solver also contains a high level of hazard frequency, with the majority of the flooded extent being labelled as hazardous. This represents an over estimation of the hazard however, as regions that had not been affected such as the northern section of the domain are calculated by the code as being part of the hazard extent with high frequency levels. This is a potential issue with the Rusanov solver and other simplified methods, which can prove to be unstable in transcritical conditions where maximum wave speeds will increase, and potentially cause the wave front to produce higher velocities, which are then used in the calculation of hazard. Typically, the use of maximum global wave speed values increases the diffusion of the numerical solution, but for this test case it appears to increase the wave front to potentially high levels, that increase the calculated hazard level. The hazardous regions are more defined with the Roe solver. The main flow path contains high frequency values, as does the secondary flow paths indicating that at this level of numerical solution and physical representation the main regions of risk are consistently replicated across the parameter space, increasing the confidence with which conclusions about areas of high risk can be made from this module. The issue of the wider flood extent in the middle section of the domain is also not highlighted as being critical to overall hazard frequency values. The larger extent for the Roe solver in the lower regions of the model domain are caused by low friction value models and high hydrograph values (0.015 and +20% of the calculated hydrograph). This has not been previously highlighted as the frequency value is very low and relates to only 2 model runs within the parameter space. This highlights a potential advantage of using the uncertainty hazard approach, as this

method helps to refine and identify non performing models. Furthermore, the region which had been identified as being inundated for both the full SWE modules and part of the ACC module, can be seen in the Roe and ACC figure to be contain either low frequency values or to be outside the threshold of the hazard. This again highlights the advantage of including risk as a means of evaluating model performance, by explaining some of the discrepancies between the modelled output and the observed data set. In this test case a region that impacted the F^2 value for these modules can be identified as being of low risk in terms of the impact of the flood water to life and in causing structural damage. It also highlights the potential drawbacks of simply using F^2 to evaluate models. The ATS provides better goodness of fit, but poorer estimations of the hazard, in terms of cost and in this analysis, which may have detrimental effects on decisions made related to a similar modelling exercise.

5.4. Sensitivity Analysis

The results for the Morris method are presented in Figure 5.20. 210 model realisations were used to complete a 4 level 6 factor parameters space, based on the same parameters as the Glasgow data set. The function used in this analysis is the F^2 value, chosen to highlight the relative influence of factors for a typical model evaluation technique. The other methods evaluated were excluded either due to the high uncertainty associated with the function (such as depth comparison) or the subjectivity of the method (Exposure based analysis).

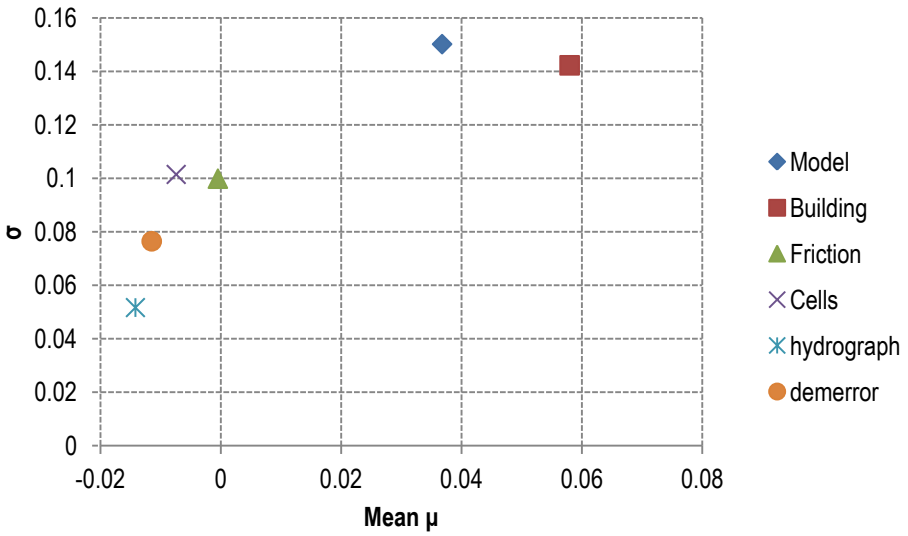

Figure 5.20: Morris Results for F^2, where the x axis represents the mean elementary effect. The values themselves only provide a relative ranking position and do not quantify the level of interaction.

From Figure 5.20, the relative importance of the building representation technique can be identified, as well as the relative importance of the level of physical representation, represented by the Model choice. Both factors can be ranked as being the most influential inputs to the model. The ranking choice of model factor again indicates that this factor is highly significant in affecting model results, and should be well considered in any modelling exercise. The third highest ranked factor is the value of Manning's friction, which confirms findings from the systematic analysis, that the three most influential parameters are model choice, building representation method and friction choice. These factors affect the flow rate and the ability to define the main flow path which has proved to be critical in the modelling of this event. The hydrograph and the cell size are both low ranked factors, which confirm findings from the systematic analysis and the visual analysis of model interaction in Figure 5.12. This highlights particular aspects of this case study, and the appropriateness of using a goodness of measure fit to evaluate models, where variations in hydrograph level and cell size appeared not to affect the extent considerably, despite the variations in water depth. The error associated with the LiDAR DEM is also a low ranked value, which confirms what has been identified in the analysis of this test case and the previous test case. It identifies that this factor is less influential in affecting model results, although the method by which it is examined (a random valued DEM, based on the limits of the RMSE) may also impact on the results.

The level of interaction between parameters is relatively similar to the factors first order influence. This is a variation from the quantitative analysis in the Glasgow test case, where interaction between parameters accounted for a large amount of the total variation of the overall model results. The implication here, as was determined from the systematic approach, is that the nature of the event reduces the significance of non essential parameters, due to the constricting nature of the topography.

The implications from these results must be considered in a wider aspect. The Morris method can be a useful method for identifying the most influential factors, but where factors are of similar rank can often fail to differentiate between factors. The uncertainty of using this method is reduced by the analysis of the Monte Carlo study, and both approaches highlight the significance of the level of physical representation and building representation, which increases the confidence in the conclusions of the approach.

The Bayesian emulator approach used a 100-point LP-Tau input space based on the ranges of parameters used in the Monte Carlo approach. As with the Glasgow test case, the approach has been reduced to the ACC and Roe solvers and the two

are compared directly, rather than include them in the factor space. The building representation has also been fixed to the Building Block method to reduce the parameter space and reduce computational cost. This approach still highlights the importance of an input in each approach and reveals details about factor interaction over different physical representation levels. The effect of each factor is described as a percentage of the total variance.

Parameter	Total (ACC)	Total (Roe)
Cell	83.9	27.21
LiDAR	0.81	0.8
Hydrograph	1.86	0.46
Friction	12.38	69.96
Cell and Friction	7.13	>1

Table 5.4: Contribution of parameters to overall variance as a % of total variance

For each approach the significance of the factors appears to evolve considerably in comparison to both the other module and the overall factors. The influence of the cell size appears to be considerably greater than has previously been identified. This can be explained through the absence of the building representation method in the analysis. The factors in combination control the numerical description of the main flow path, which is the most influential in determining model output. This factor by itself provides the description for this feature and explains the relative high significance in the GEM analysis. The factor of grid resolution accounts for the majority of the variation seen in the ACC code. This pattern has been noted in this test and more significantly in the Glasgow test, where increasing cell size reduced the extent of the ACC model output. The level of friction value accounts for most of the variation with the Roe solver, and confirms what has been identified in the interaction plot. As has been highlighted in the Morris method, the interaction between parameters is low, with only one combined factor creating a significant variation (cell size and friction for ACC). This combined influence is unsurprising considering the high value of the two factors in the first order.

5.5. Conclusions

The Coventry test case is a unique type of event, which can be classified as a low occurrence probability, high impact event. The restrictive topography, relatively small model domain and the unique hydraulic properties provide a useful examination of the LISFLOOD-FP code and modules. The highest goodness of fit value was returned by the ATS module, while increasing levels of physical representation appear to reduce this function. Further examination by the use of

risk and vulnerability measures have highlighted not only how this value can potentially be misleading, it can also lead to an under prediction of the hazard and associated cost. It also indicates that a possible requirement of urban modelling is he use of a connected sewage networks when modelling outburst events. This can be assumed based on the development of a flood pathway to the north of the outburst point seen in the full SWE models. A significant assumption made in the use of flood inundation models in urban environments is the surcharging of drains. This test indicates that for certain scenarios this may not be valid. Overall the level of physical representation appears, as with Glasgow to be the most influential in determining model results, but that this influence is not monotonically increasing with complexity of the physics model in comparison to the previous test case. In this test case, the overall trend is for the ATS and ACC results to be markedly different than in comparison to each other and the full SWE modules. This indicates that assumptions about appropriate levels of physical representation may be difficult to quantify, but the impact maybe critical overall. As has previously been seen the impact of models with different levels of physical representation tends to be local. In this test case, the outputs of each model had a wider range near to the inflow boundary, which reduced in distance from this point. This indicates that evaluation techniques and model checks should include further detailed checks at this point, or use methods such as the regional analysis and depth of cost method which is weighted towards this area. In that way these local variations are brought to the attention of the modeller in an immediate way.

6. TEST CASE MEXBOROUGH

The final test case is a river overtopping event in Mexborough, South Yorkshire. The test case is based on an event from the summer 2007 flood events, and data includes RTK GPS surveyed high water marks. The model domain is predominantly rural, giving a useful counter point to the two urban test cases, in that theoretically the hydraulic conditions could be represented with a diffusion wave approach, and the inclusion of additional terms from the governing equation is less critical. The model will also require a representation of the river and represents a problem typically undertaken with inundation models. The conclusions based on this result and the other test cases will reduce model bias from the research conclusions.

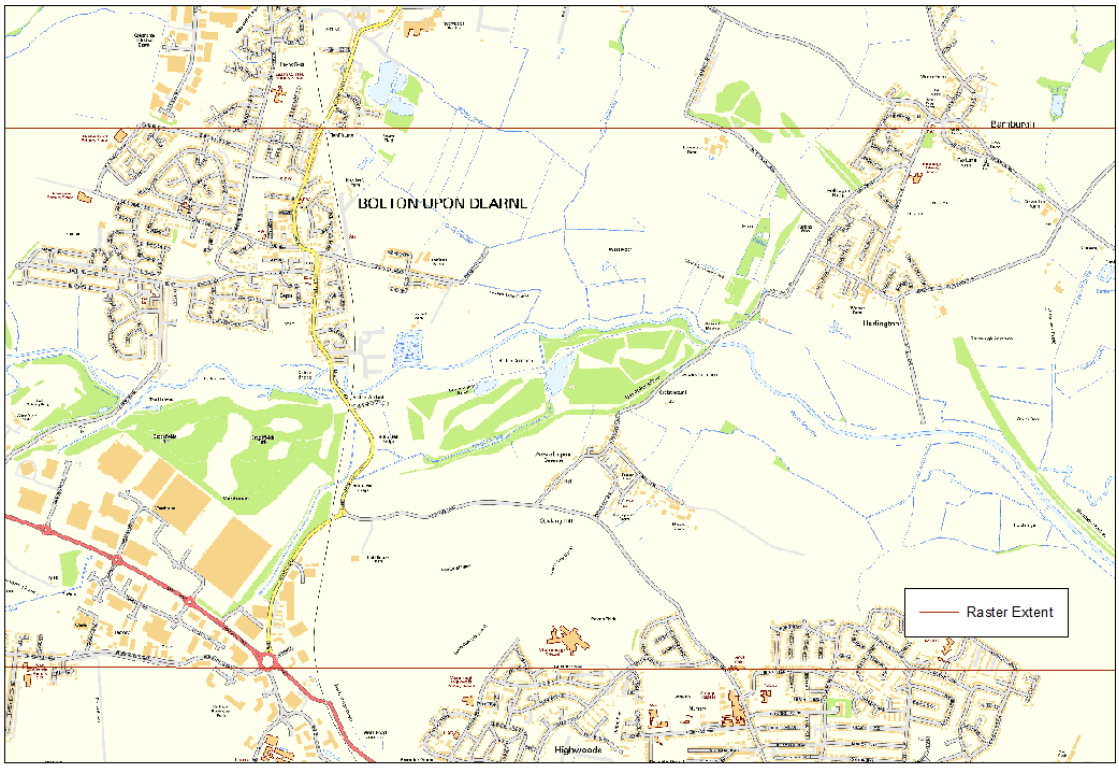

Figure 6.1: Map of Mexborough, the model extent and location of buildings

The Mexborough event was part of the wider UK 2007 flood events, and occurred on the 25th June as a result of prolonged rainfall. At the time of the flood water levels at the gauge station of Dearne upon Adwick was already recording discharge values above the mean value of 3.5m³/s at around 25m³/s. A period of prolonged rainfall in the preceding days increased the peak flow to an estimated 280m³/s. The test case is based on a 20km² region, south of the village of Bolton upon Dearne and north of the town of Mexborough. A series of surveyed wrack marks measured by the EA identify the maximum flood extent of the event, and allow some initial

identification of key risk regions. Key features within the flood plain include a number of business premises and a waste water treatment facility. The village of Bolton and Mexborough both fall outside of the flood extent at higher ground than the regional floodplain, and consequently the probability of fluvial inundation are low. The road system is also less developed than urban test cases with only a few A and B grade runs crossing the floodplain. A major north south railway runs through the model domain crossing the river south of the village of Bolton, although at an elevated height from the flood plain (10meters above the floodplain). The event has been modelled in previous research. Beven et al (2009) explored the methods of communication of uncertainty in flood extent maps and diagrams using this data set. Focus was on a key water treatment facility which fell within the uncertainty range of the flooded extent, but which had fallen outside of the extent during the event itself. Leedal et al (2010) furthered this work, and further detailed modelling work was conducted by JBA consultancy in part of their report to the local council.

6.1. Test Case Design

In order to model this test case, a representation of the river was required as a boundary condition. A 1D channel linked to a 2D floodplain was found to produce model instabilities in both the Roe and Rusanov solver that caused the simulation to fail. Therefore an explicit 1D channel river model, which is typical to most modelling exercises that require a river to be represented, was excluded. The river was represented as a series of wetted cells that correspond to bank full level, with the hydrograph inflow represented as a point source at the western boundary, with the QMED value removed. The DEM is based on a 5m DSM that was resampled to 20m and 40m scale. Initial model runs revealed that the computational cost was significant for LISFLOOD-ATS at 5m and 10m resolution. The grid resolution was therefore decreased to 20m and 40m. This approach can be further justified by an analysis of the building size within the flood inundation extent. A number of large scale buildings with footprints greater than 100m² occupy an area to the south of the village of Bolton, which would be sufficiently well resolved at a scale of 20m and 40m. The majority of smaller, residential buildings are located outside the flood extent, and at significantly higher elevation than the floodplain, which would require higher levels of representation that can be afforded from the 20m cell size.

6.1.1. Test Case Factor Inputs

The Bolton at Dearne gauge, located with the model domain provides input data for this test case scenario.

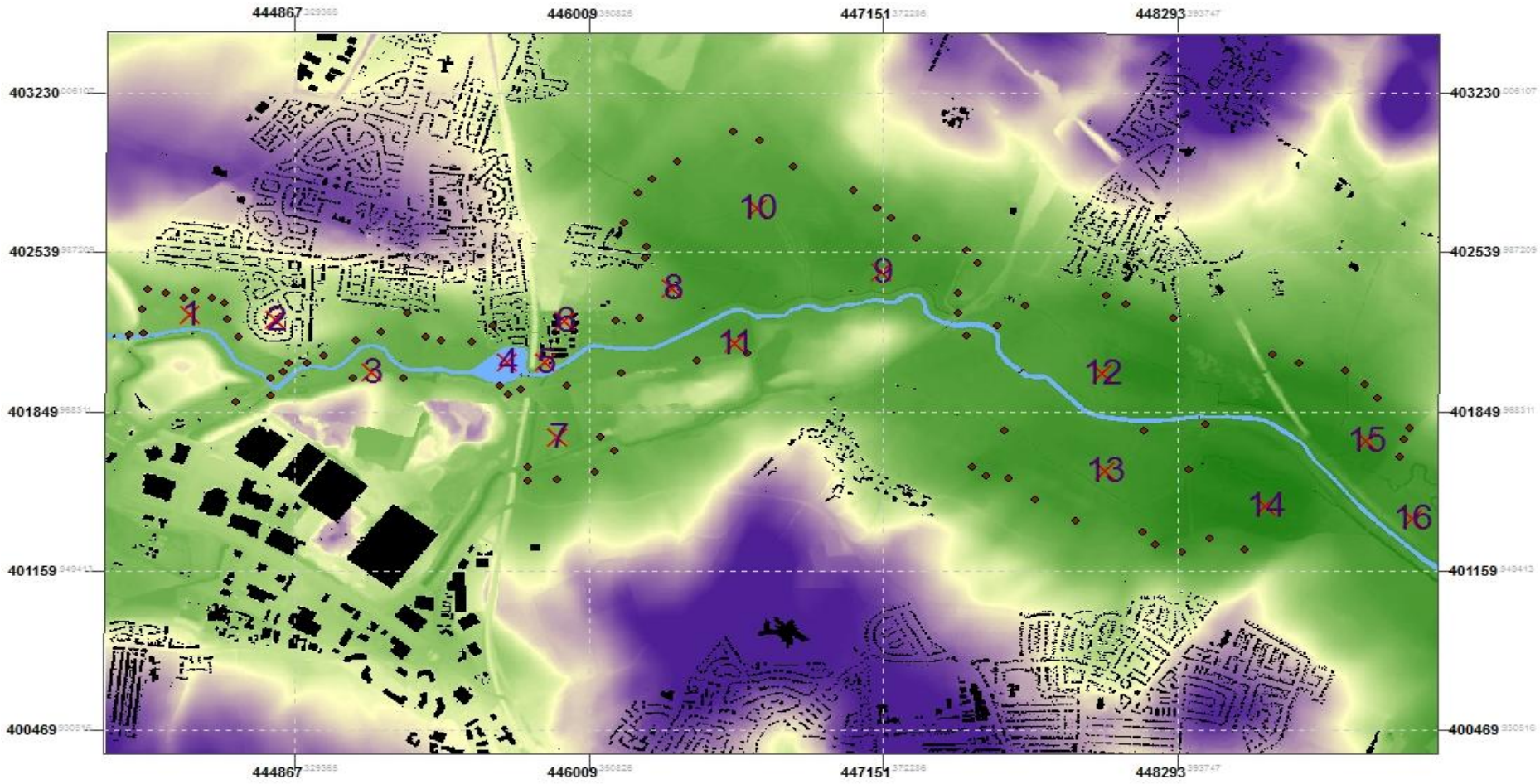

Figure 6.2: Digital Surface model of the test case, with location of buildings, Control Points, and outline of River Deane

The magnitude of the event was greater than the range used to calibrate the gauge, which consequently impacted on the predicted discharge value. Initial model runs that used the gauge recorded values produced a flood extent far in excess of the observed data set. Previous work by Beven and Leedal (2010) had identified that a 76m³/s discharge value represents a return period of 1 in 100 year event (0.1 AER), whilst a 96m³/s discharge represents a 1 in 200 year event (0.05 AER). Given that this event was identified as a 150 year event (0.075), it is likely that the event will be between these values with a lower peak value of 86m³/s. The inflow is displayed in Figure 6.3.

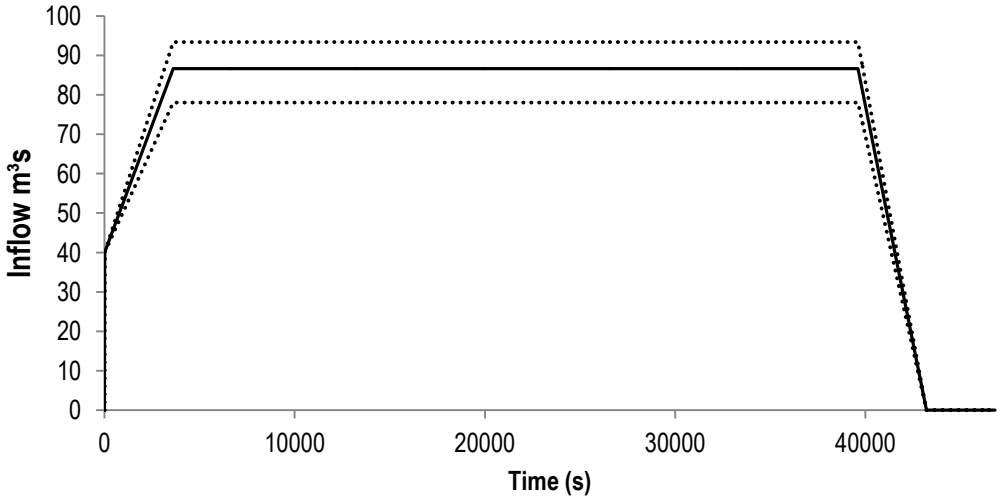

Figure.6.3: Inflow hydrograph for the Mexborough Test Case, with uncertainty bounds, based on Return period flow added. The simulation is effectively steady state.

The gauged data indicated that the duration of the event (the time over the mean discharge value) was over 5 days, which provided additional computational costs. Further model runs were required to reduce this factor, and a final inflow hydrograph based on a peak flow of 15 hours was created, with a model runtime of 17 hours. This provided enough time for the model codes to allow water to flow across the model domain to the eastern boundary, with an open boundary condition applied. Further evaluation of downstream boundary conditions was excluded due to computational cost.

As with the previous test cases, a compromise between accuracy of simulation number of input factors, and computational time was required. The final set of factors evaluated in this test case is outlined in Table 6.1.

Factor	Parameter Range(low to high)
Module	ATS-ACC-Rusanov-Roe
Building(type)	BH -BP-BB
Friction distributed(n)	0.015 to 0.075
Friction Single (n)	0.01-0.07
Cell Size(m)	20m and 40m
Hydrograph	Minus 20 to +20 of 150 year event
DEM error	0cm -15cm

Table 6.1: Parameter space for the systematic analysis section

This input space created a total number of 2020 simulations, 505 for each module. The sensitivity analysis section was also based on this with the exclusion of building type as an input to both the BACCO GSA method and the Morris method. This was due to instabilities with the Rusanov solver when the building porosity method was used.

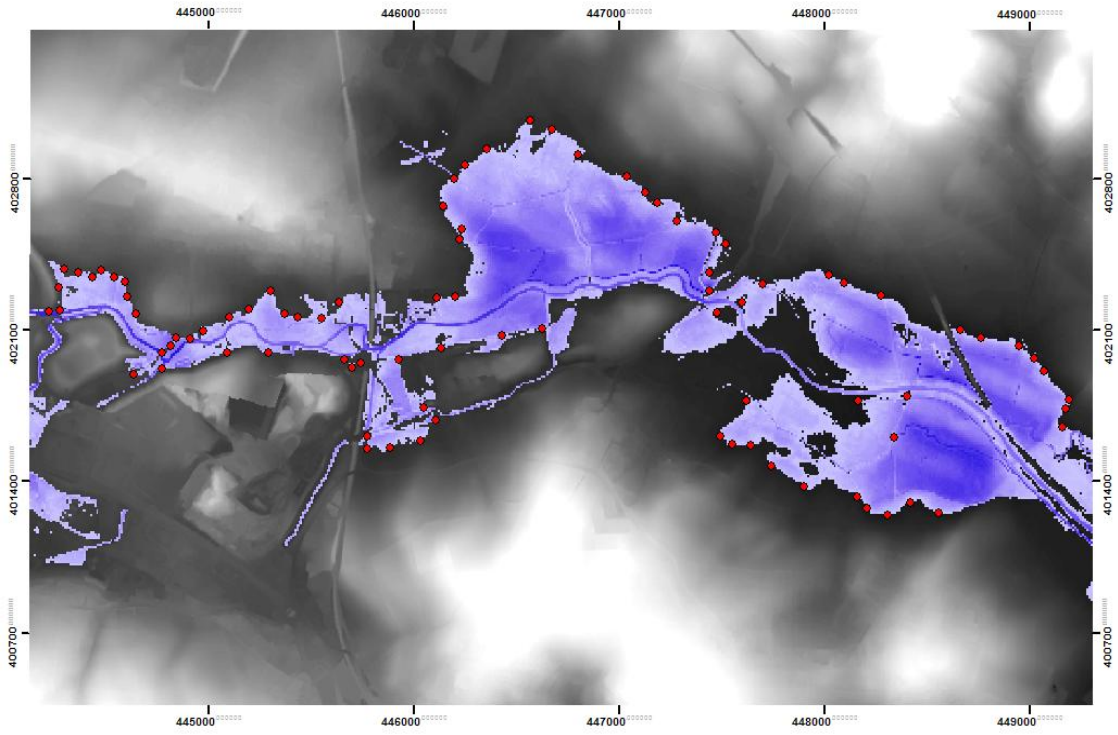

Figure 6.4: Digital Elevation Model (DEM) of Mexborough model, with the surveyed wrack mark points and the calculated observed water surface

The model realisations were evaluated using measured maximum extent data points. A water surface was created from this data, using the 5m DEM model to determine water depths across the model domain, which provided the means for evaluating model results using F^2 (Figure 6.4). The model simulations were also evaluated by comparing output water depths with the observed water surface and

depths, and using the Root Mean Square Error (RMSE) of the variations to describe the overall model performance similar to the approach by Mason et al (2009) and Stephens et al (2012). In order to ensure that the water surface was not affected by the error of the LiDAR DEM, the elevation of the wrack marks was based on the DEM value rather than the recorded value, which had variations of 10cm to the DEM. Further evaluation techniques could be undertaken with this data set, such as distance of the model extent to the measured extent points as used by Pappenberger et al (2006) and Neal et al (2005), but the use of the RMSE evaluation measure is used here as a means of identifying trends between water elevations and the F^2 evaluation technique, as well as investigating calibration methods, and the dependency between important parameter sets and model evaluation technique. These evaluations methods were carried out for both the systematic analysis of parameters and the Sensitivity Analysis techniques.

6.1.2. Exposure Based Evaluation

The rural nature of the model domain requires a different approach to the application of determining areas of significant risk and vulnerability, and applying the risk methods in an appropriate way. The majority of residential units, buildings and significant population areas within the model domain are located outside of the recorded flood extent, and at an elevation significantly above the flood plain. A number of building units to the south of the river are located at the edge of the recorded flood extent, which are more likely to be affected by the flood, but the main impact of a flood in this region is the potential of travel disruption caused by flooded roads, and more significantly damaged bridges. Consequently, the weighted vulnerability approach was designed to account for this, with higher weighting values assigned to cells that represent roads and to cells adjacent to bridges (Figure 6.4). The estimated cost is also anticipated to be lower in this test case than either of the urban based test cases, where the key areas that will contribute to the estimated cost are located to the south of Bolton upon Dearne that include a Water treatment centre, a bridge and a number of residential buildings. South of the river at this point include a number of warehouse and office facilities. The Water treatment centre has been identified as at risk by previous studies (Leedal et al 2010, Beven et al 2009), and depth damage curve relating to Electrical Warehouses was used as the means for determining the estimated cost from model output. The other categories of building in the model domain identified in the OS data, were residential housing and office/warehouse which were evaluated with the corresponding depth curves from the MCM manual (Penning-Rowsell 2010).

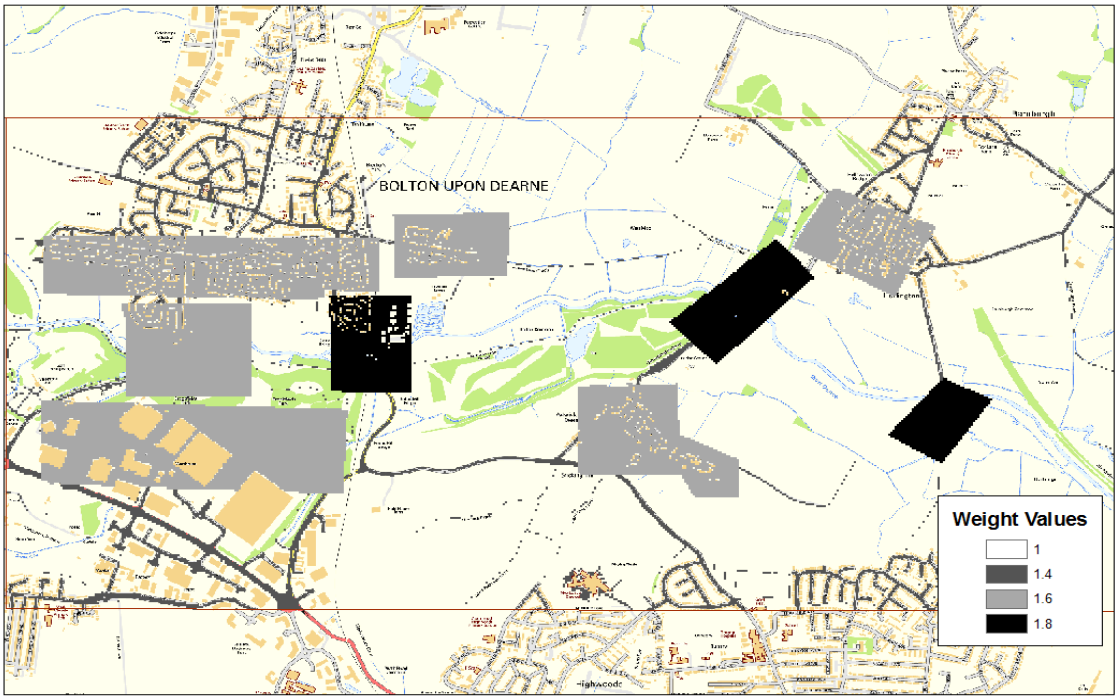

Figure 6.5: Vulnerability Weighted cell values and locations

Regional F^2 analysis was based on dividing the domain up into sections based on the direction of water flow, and building types. For both Coventry and Glasgow the main division had been based on nearby building types, grouping the regions into sections based on common proprieties types. Here, this is extended to also explore spatial dependency of factors and evaluation functions.

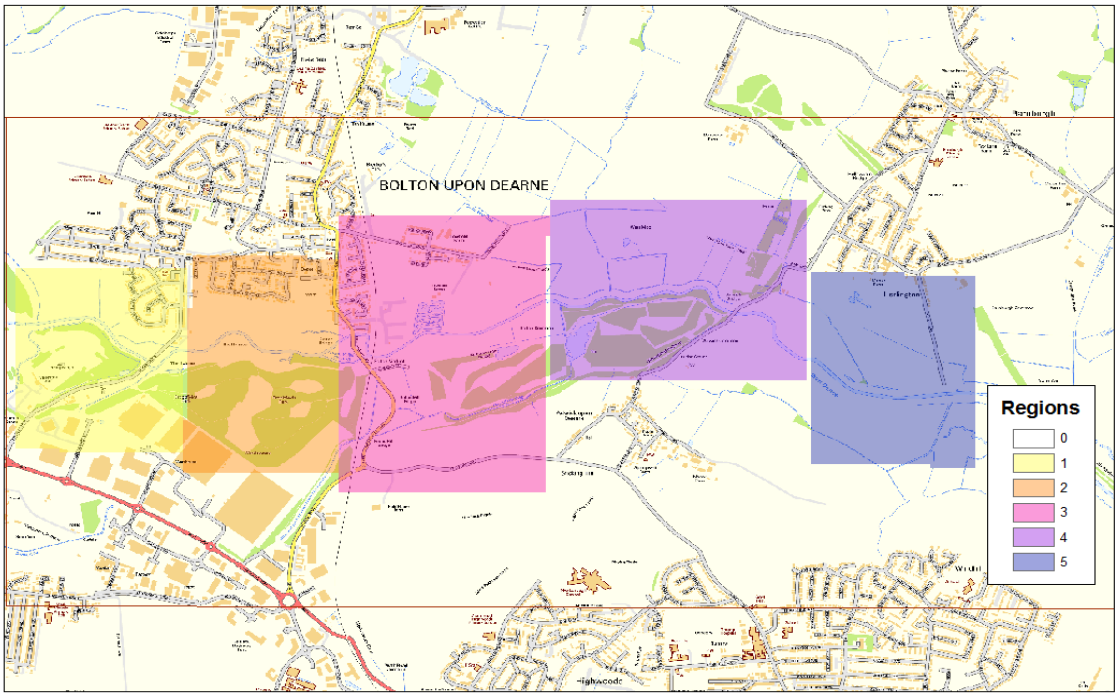

Figure 6.6: Location and value of Regions for Regional F^2 analysis

The F² regional areas, summarised above in Figure 6.6, represent the initial inundated regions (1 and 2), the region south of Bolton Upon Dearne, which includes a significant number of residential buildings as well as the location of a number of critical infrastructure, including a railway bridge, and the Water Treatment plant (regions 3). Regions 4 and 5 also a smaller number of buildings and bridges, than region 3 but provide more critical insight into model performance at the downstream edge of the model domain. Region 3 represents a critical component of this modelling exercise, with a significant number of buildings and infrastructure components that are essential to the wider region.

The duration of the model run was restricted for this test case, due to the time required for the ATS model to run. Model run time was limited in order to allow the ATS simulation to finish simulations within the time requirements of the ARC1 HPC system. An adverse effect of this was to bias the results against the ACC code, which typically produces a lower arrival time for the flood wave, which become accentuated the further away from the inflow point. This impacted the model results in this test case, by not providing sufficient time for the flood wave to fully pass through the model domain in the ACC model runs. This effect is considered in the analysis.

6.2. Results

This section provides an overview of the model results for this test case, with an evaluation of the immediate model outputs, which include simulated extent and depth of water at control points. This is followed by analysis of model evaluation functions, analysis of exposure based methods, and Sensitivity Analysis.

In comparison to the previous test cases, the variation between model codes is smaller in terms of model output, function results, and sensitivity to parameters. However, there are spatial variations of modelled depths and velocities between each code that indicate local impacts related to the level of physical representation in the model code. A summary of the model output will be followed by an analysis of the model functions and the risk evaluation methods to identify the overall trends of this test case.

The hydraulic characteristics of the Mexborough test case are different to both the Coventry and Glasgow cases. Flow across the model domain is typically sub critical, although supercritical conditions are apparent in the initial inundation phase at locations close to the input boundary. The main process of the event involve a flood wave progressing from west to east along the direction of river flow, with the

depth of water near the channel reaching equilibrium with the inflow hydrograph before dispersing during the drying phase at the end of the simulation period. At the centre section of the model domain, at the widest point of the flood plain, flood water also disperses in a north south direction (CP-10, Figure 6.2). A number of additional flow characteristics and process occur which impact the inundation extent, including the influence of bridge embankments which redirect flow from the main channel and floodplain flow path. The bridges are not explicitly represented in the spatial domain, but the influence of the embankments can still be observed. The influence of back water effect is not observable at the nearby control points due to the absence of detailed representation of the bridge structure. The DEM contains part of the embankment related to the bridge within the channel, but not the channel remains unconstrained. A secondary effect on the inundation extent is the influence of additional tributary channels and storm drainage channels. These secondary channels are significant in directing river flow from the main direction of flow, towards the floodplain. This effect is particularly noticeable around CP-3 (Figure 6.2), which directs channel water towards a number of office and warehouse facilities. The impact of these secondary channels changes as a direct result of the level of physical representation in the model, as can be seen in the uncertainty flood inundation extent plots, based on the ensemble of model runs. The final inundation extents for the 4 modules are displayed in Figure 6.7, and show broadly similar extents with high frequency values for the extents. This indicates that each module can reproduce the flooded extent reasonably accurately and consistently, across the parameter range.

The uncertainty extent plots, Figure 6.7, provide a guide to the impact of physical representation on inundation extent. A number of small variations provide insight into impact of the level of physical representation on the potential inundation extent. The ATS module (top left Figure 6.7) contains very few regions of uncertainty (inundation frequency values >0.7), in comparison to the other modules, which all display larger regions of low frequency values. The ACC module (top right Figure 6.7) has produced a region of high uncertainty at the eastern boundary of the model domain (at 4500m, 1000m). The elevation of this region rises steeply from the direction of the river, before a gradual downhill section and a second small rise in elevation, creating a channel, which allows water to flow in a west east direction. The area inundated by a small break in an embankment, indicating that a certain level of physical representation is required to overcome these obstacles. It is noticeable that the other modules contain high inundation frequency values for this area. This is caused by the ACC producing a slower flood wave front in comparison to the other modules and is explored in more detail in the comparison of depths. A

noticeable variation in this region can also be identified in the ATS solver, where a small band of high value inundation frequency cells are present, where as the rest of this region has a lower frequnecy level. The implication here is that both the simplified modules lack either the process required to regularly inundate this region, for the ATS it is a product of the lack of representation, where as the slower flood wave produced with the ACC module fails to reach the lower level of the module in sufficient time. The ACC module has however produced similar extents to the other higher level modules over the rest of the model domain, with similar regions of uncertainty. The impact of secondary channels and module choice on simulated extent is evident by a region of uncertainty at the western end of the model domain.

The 3 higher level modules have the additional process in the numerical model which allows the water from the main channel to flow along the storm channels to inundate this region. The low level of inundation frequency values indicate high levels of uncertainty and the influence of other parameters in affecting how this region is inundated. Typically, models that use lower friction values and higher hydrograph values from the parameter set inundate this region. Each module produces a distinctive inundation pattern in this region, with the ACC the smallest area and higher frequency values (≈0.4), where as the Rusanov produces the widest area, with the lowest frequency values (≈0.2), indicating infrequent inundation from the model set. The Roe solver produces a similar extent as the Rusanov solver but with a similar range of frequency values to the ACC module. The indication here is that with increasing representation leads to a significant increase in the amount of water able to move along these channels. The higher level of numerical solution from the Roe solver creates a region with a higher frequency value, although still with a large amount of uncertainty.

Another variation between all the modules is the flooding of a separate drainage channel which is located in the centre of the model domain, at the 2000 to 2750 meters and at the 1100 meters region (Figure 6.7). For the ATS, ACC and Roe modules, a distinct increase in the frequency value can be observed. This channel is protected from the main flow path by a ridge of high land. Water inundates this region by moving up a drainage channel, and can be observed in the observed flood extent as flooded (Figure 6.2). This indicates that increased physical representation is a significant factor in modelling smaller scale features even in large rural scale events.

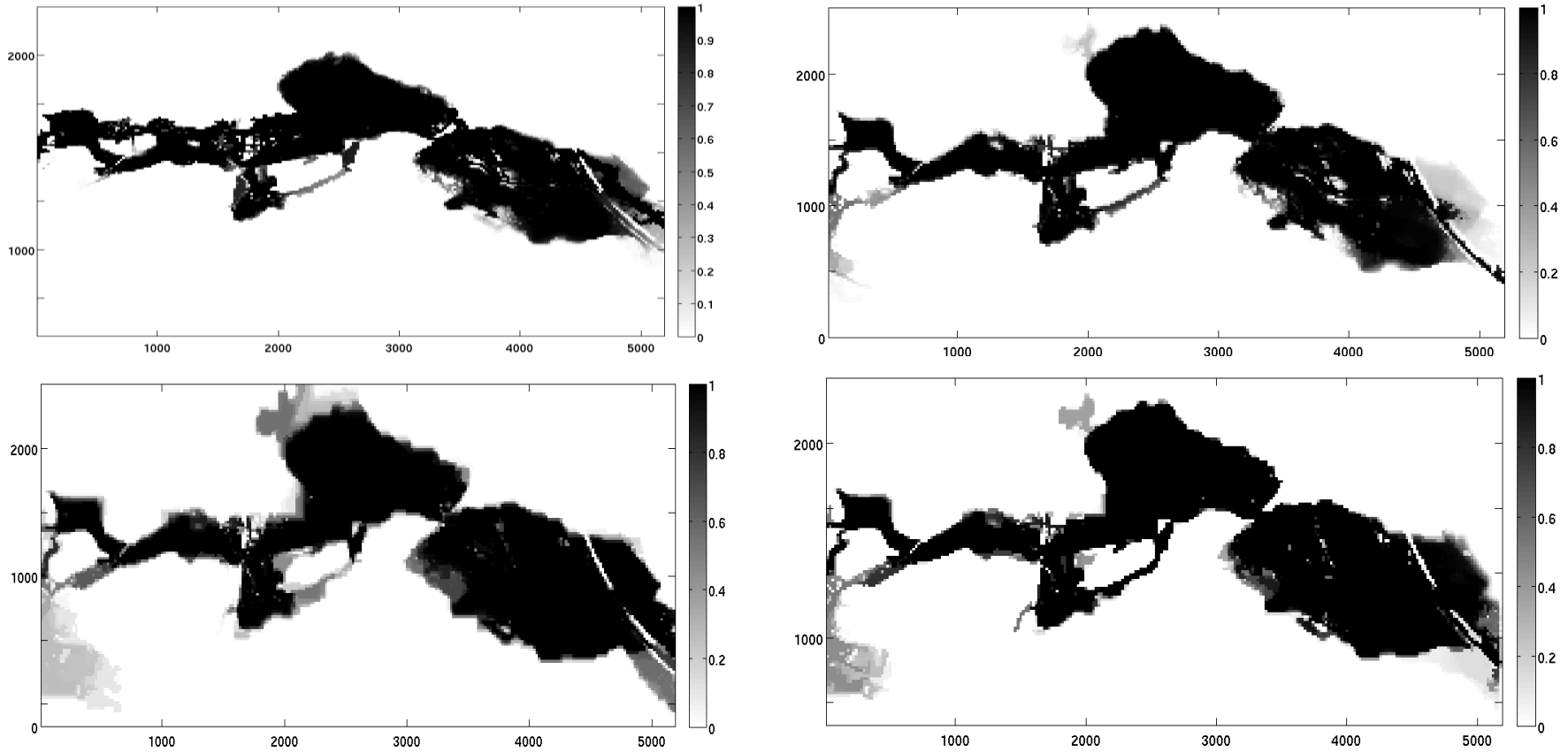

Figure 6.7. Uncertainty flood extent plots for the 4 LISFLOOD modules over all test case results (top left ATS, top right ACC, bottom left Rusanov, bottom right Roe), where dark regions represent cells that have flooded in all model realisations.

The Rusanov solver by comparison also records this as flooded, albeit with a lower frequency value than any of the other 3 modules. Both the Rusanov solver (bottom left, Figure 6.7) and the Roe solver (bottom right Figure 6.7), have produced similar flood extents. Similar patterns identified in the previous test cases appear here. Firstly, the Rusanov solver produces a wider flood extent than the Roe solver. These regions on the edge of the extent tend to be of low frequency value. This indicates a strong interaction with other parameters, in particular the value of Manning's friction, which at low values causes the Rusanov solver to create a wider flood extent. Secondly, the boundary of the Rusanov uncertainty flood extent is not as clearly defined as the Roe solver, with increased fuzziness of the flood extent boundary at the edges and between zones of greater uncertainty. This is similar to the Coventry test case, indicating that the use of simplified numerical approaches increases the potential of uncertain parameters to influence model results.

Analysis of water depths at key points through the model domain also provides insight into model performance, and also the influence of the value of the friction parameter as a control on water depths. CP-4 in Figure 6.7 is located near the railway bridge and near residential buildings, and the impacts of the flood event on these features should be clear from the depths at this location. Figure 6.8 is the water depths for each module over the time of the model run, where each line represents a model from the parameter set. Each module produces a distinctive range of depths for the total model ensemble, which indicate that the relative influence of the level of physical representation is important, as well as indicate the level of interaction with other parameters for each module. The influence of the input hydrograph can clearly be seen in all the water depth graphs, where the inflow reaches a peak value which is maintained until the last 2 hours of the simulation time. At this control point, each module replicates this effect, with a drying period also noticeable at the end of the simulation. The ACC (top right, Figure 6.8) module produces a narrow range of maximum depths across the parameter space, which indicates a relative insensitivity to other parameters, which has been seen in the Coventry test case, and the Glasgow test case. In comparison, the ATS module (lop left Figure 6.8) creates a wide range of maximum depth levels. The main control of this wide range of depths is the friction value, and the hydrograph, where low Manning's friction results produce the lowest modelled depth values. This appears to counter previous evidence that suggests the ATS solver is relatively insensitive to floodplain friction parameters (Horritt and Bates 2001). The same pattern of sensitivity to other parameters emerges in analysis of the SWE modules, although interaction between the value of Manning's friction, the model code and

other factors also appears to influence the maximum depth value and the timing of initial inundation. The Roe and Rusanov solvers produce similar hydrographs, although the range of depths produced by the Roe solver are larger and feature higher peak values, indicating a greater sensitivity to other factors and parameters. The depths at this control point indicate not only a significant variation between levels of physical representation and interactions between the modules and the other input factors; it also indicates a variation between local model outputs and global outputs.

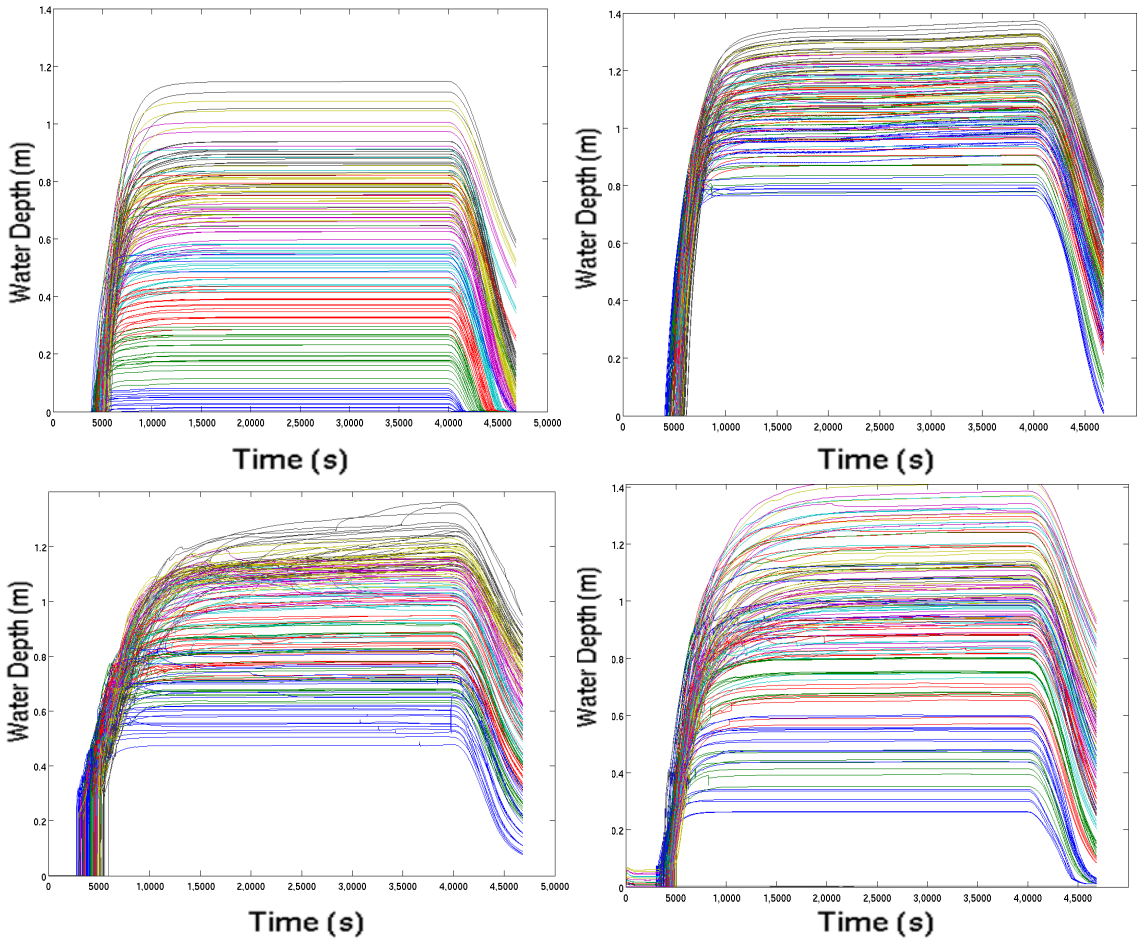

Figure 6.8: Water depths for the 4 modules (top left ATS, top right ACC, bottom left Rusanov, bottom right Roe) over the range of tests at the control point 4, all plots scaled to 0 and 1.4m across the entire model run time. Each line represents a model run

Analysis of control points further into the model domain provides insight into the speed of the flood wave front produced by each module, the influence of other factors, and the performance of each module with increased distance from the inflow boundary. CP-8 (Figure 6.2) is located in the centre section of the model domain, and the depths for each module display similar characteristics as had been seen at CP-4. The ATS module (top left Figure 6.9) has the widest range of depths

across the model ensemble. ACC (top right Figure 6.9) produces the narrowest range, and also the highest depths. Both Roe and Rusanov produce similar shapes to the depth graphs, a sudden initial inundation phase with a significant depth of water, rather than a gradual inundation phase demonstrated by the simplified approaches. This is followed by a short inundation to peak time, with a lower maximum depth level than the ATS and ACC modules, followed by a gradual draining phase until the end of the model simulation, confirming some of the findings from analysis of CP-4.

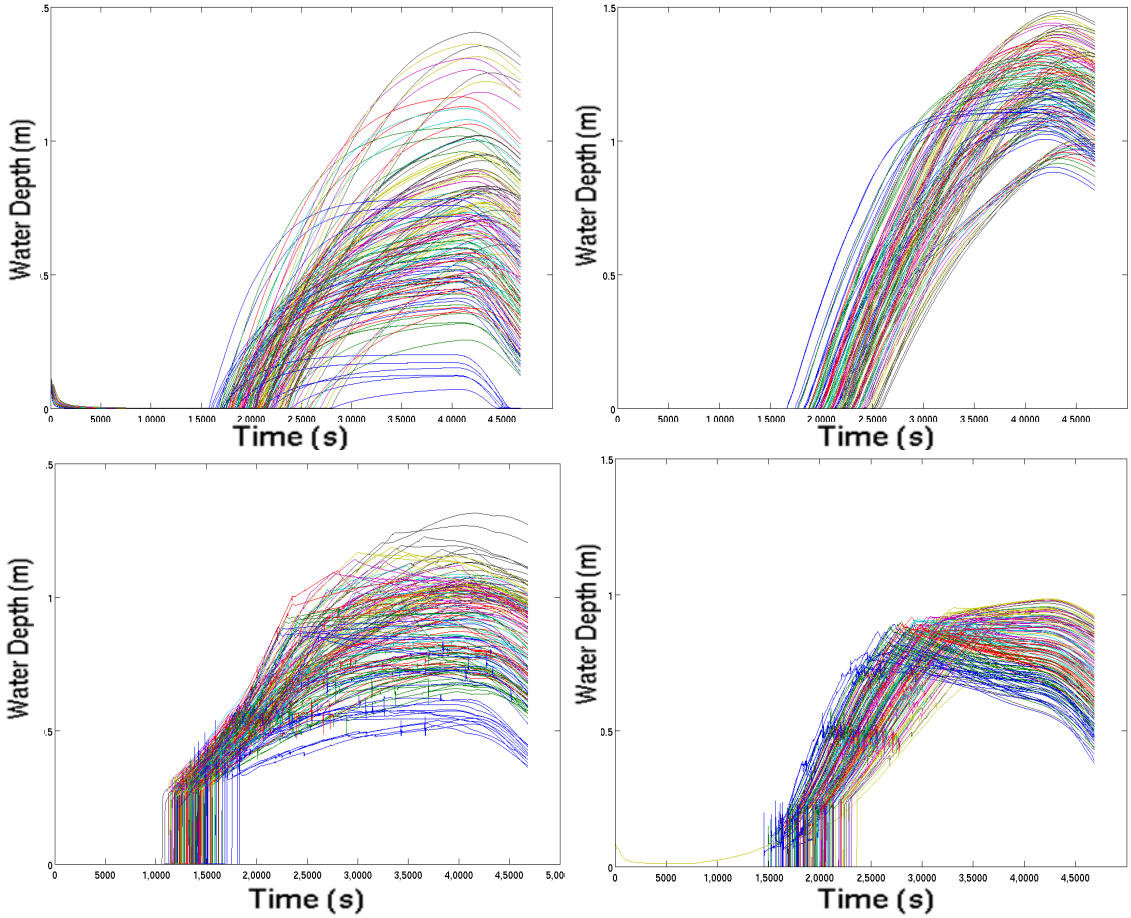

Figure 6.9: Water depths for the 4 modules (top left ATS, top right ACC, bottom left Rusanov, bottom right Roe) over the range of tests at the control point 8, all plots scaled to 0 and 1.5m across the entire model run time. Each line represents a model run from the parameter set

Both the Rusanov and Roe solvers create an earlier inundation period than either of the simplified models, indicating that these produce a faster moving wave front. In contrast to CP-4 however, the Rusanov solver now produces a wider range of maximum depth values and final depth values in comparison to the Roe solver. The final peak level maximum for the Roe solver is approximately 1m compared to 1.3 meters for the Roe solver. This distinct spatial variation in the module results

indicates that each module responds differently with increasing distance from the inflow source. In comparison the simplified approaches appear to retain the characteristics seen at CP-4, and these compare to what had previously been seen in the Coventry test case, that the ACC produces a later arrival time compared to either of the other modules and a longer initial inundation to peak value time. As with CP-4, a significant control on the water depth level at this control point is the value of friction coefficient for all modules, indicating again that the value of water depth maybe strongly influence by this factor.

Control point CP-10 is located at a further distance from both the river and the inflow boundary than CP-8, but due to the lower elevation is actually inundated before CP-8. The simulated depth is greater than at the previous locations, with a peak value of 2.5meters modelled by the ACC module (top right Figure 6.10).

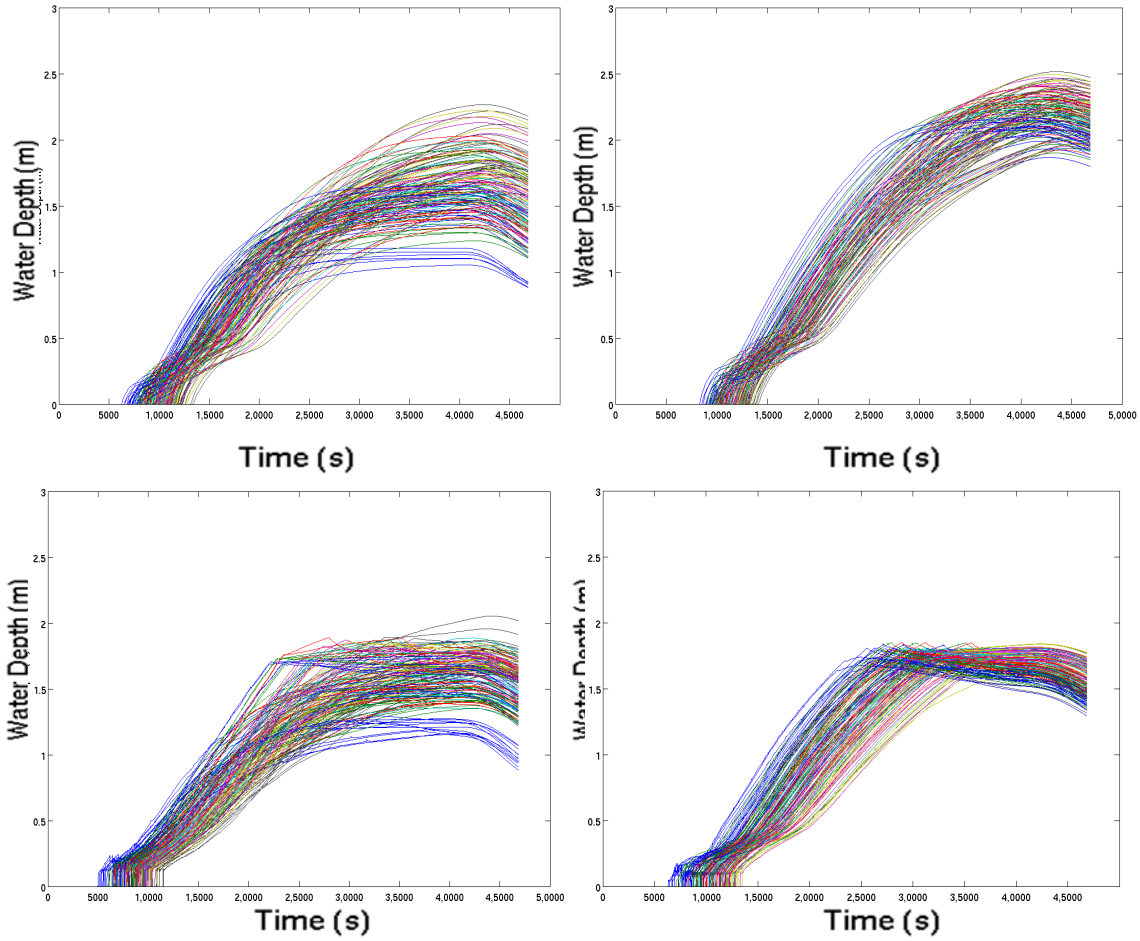

Figure 6.10: Water depths for the 4 modules (top left ATS, top right ACC, bottom left Rusanov, bottom right Roe)over the range of tests at the control point 10, all plots scaled to 0 and 3m across the entire model run time. Each line represents a model run from the parameter set

Similar patterns are also observable at this location as has previously been observed in this test case and the previous test cases. The ACC module produces a later arrival time, a longer initial phase to peak value time and a narrower and

higher range of final depth values. The ATS module produces a wider range of final values that are dependent on the friction value and both Roe and Rusanov produce a similar shape depth graph, but the Roe solver produces a narrower range of final depth values and maximum depth values. Again this indicates the relative variation between the simplified approaches compared to the full dynamic wave approach, the sensitivity of the ATS module to friction values, and the relative insensitivity of the ACC module to other factors. The arrival time of the flood wave also varies across the module range. Both the Roe and Rusanov solver produce earlier arrival times, and a wider range of arrival times than the ACC and ATS modules. These variations indicate that the choice of physical representation is significant in impacting model results.

Overall, the analysis of stage points indicates that a significant variation is apparent between each module and the input factors. This appears to contradict the frequency extent results, which indicate that modules are relatively insensitive to other parameters. This indicates that the relationship between extent and water depths, is not clear, and can impact the evaluation of risk from model results. Despite the different hydraulic properties of the Coventry test, including a lower underlying slope value and the predominate flow characteristics being sub critical, the variations between water depth outputs of the modules is similar to that test case, such as the later arrival time and higher peak time of the ACC solver, in comparison to the ATS and full SWE modules. The implication is that for modelling scenarios where diffusion wave hydraulic conditions predominate and that can be modelled with larger cell sizes, the inclusion of additional terms to the computational code, can be significant in affecting water depths and the timing of the wave front. The similarities between the 2 full dynamic codes also indicate that the inclusion of the full terms from the governing equations will provide similar results, regardless of the numerical solution to the equations. The main issue in the use of a simplified numerical approach to the full governing equations is one of uncertainty and the influence of other factors on the final computational solution. In order to account for this uncertainty in this and across the input factor space, a detailed calibration process with adequate model performance criteria is required.

6.2.1 Model Evaluation

Each model is assessed by comparing the model output with observed extent and evaluating this comparison with two functions; a comparison of binary extent, with a goodness of fit measure, and a comparison of depths with a Root Mean Square error measure to generalise this comparison (RMSE). The model evaluation results

for the goodness of fit measure (F^2), the comparison of depth measure (RMSE) and the vulnerability weighted enhanced measure (F^2VW) are summarised in Table 6.3.

Broadly, the ranges of results are similar across the 4 module types, and the values could be described as acceptable or behavioural if the use of a GLUE methodology approach was used. In this test case, each evaluation technique provides insight into both relative module performance, and parameter influence. The highest value of goodness of fit (F^2) is modelled by the ATS module (0.7896) with the parameter set of a 20m grid resolution non degraded DEM, a spatial distributed friction surface, represented by the lowest value of friction surface - with the road surface 0.008 and the vegetation surface a friction value of 0.015, with a peak hydrograph value of 98m^3/s. Buildings and surface topography are represented with the building block (BB) method. Each module can produce a similar peak value, although with different sets of parameters. The ACC module produce a maximum F^2 value of 0.78359, with a parameter set of 20m, a single friction value of 0.01 and a peak hydrograph value of 86m^3/s. Here, the building representation is the porosity model (BP). Both the Rusanov and Roe models produce lower F^2 values, (0.7792 and 0.7772 respectively), but by a relatively insignificant margin. Again both parameters sets are based on a 20m grid resolution model, with the BB building representation method. Each uses a single friction value (0.01, 0.02), but different peak values in the input factors (0.96m^3/s, 0.76m^3/s). The similarity between the maximum values indicates that the relative significance of the module is perhaps less critical in replicating the observed extent. This would appear to confirm previous evidence (EA Benchmarking 2010), which indicates that with larger scale model domains, and in rural conditions, diffusion wave conditions dominate. However, a key point in the analysis of F^2 values is the level of equifinality between model results. A number of models from across the total factor space are capable of recreating a level of F^2 similar to the maximum values noted above. These higher F^2 values also contain slightly different parameters sets, with different peak hydrograph values, friction values and building representation values. This is a good example of the issue with calibration methods, the requirement to distinguish between large parameter spaces and number of model runs with sufficient clarity to determine a single model run as adequate.

Model(function)	ATS	ACC	Rusanov	Roe
F^2 (Mean)	0.7348	0.6908	0.6826	0.7287
F^2 (Max)	0.7896	0.5545	0.4176	0.7792
F^2 (Min)	0.6325	0.7835	0.7772	0.6691
RMSE(Mean)	0.3751	0.4714	0.3975	0.3654
RMSE(Max)	0.5222	0.6336	0.8312	0.5226
RMSE(min)	0.2437	0.3330	0.2695	0.2956
F^2 VW(Mean)	0.7384	0.7180	0.6957	0.7340
F^2 VW(Max)	0.7927	0.7866	0.7844	0.7846
F^2 VW(Min)	0.6296	0.6007	0.4351	0.6824

Table 6.3: Summary of Objective Evaluation functions for the Mexborough Test Case

The range of each module are subtle different, and are summarised visually in a box and whisker plot of model type against F^2 value in Figure 6.11. This Figure is scaled to the same degree as the corresponding Figure from the previous test cases, to emphasis the relative insensitivity of this test case to the parameter range and the level of equifinality in the results in comparison to the other urban test cases. A number of distinctive variations from those test cases are noticeable. As with the Coventry test case, the highest mean value is produced by the ATS module, this module also has the largest range of model results, although the upper and lower quartile values are higher than any other module. Analysis of the individual results indicates that the ATS module provides a consistent range of results regardless of the parameter dataset. This insensitivity increases at the larger grid cell size of 40meters. The ACC module has a lower mean value, and a wider range of total results and a greater distance between the upper and lower quartile than any of the other modules. As with the ATS module, at 20m, the ACC module produces a consistent range of results, with fewer model runs with a peak value above this range. However, at the larger grid resolution this range increase, leading to the relatively high range of model results seen in the box and whisker plot of Figure 6.11 indicating that the simplified approaches are relatively sensitive to grid cell size.

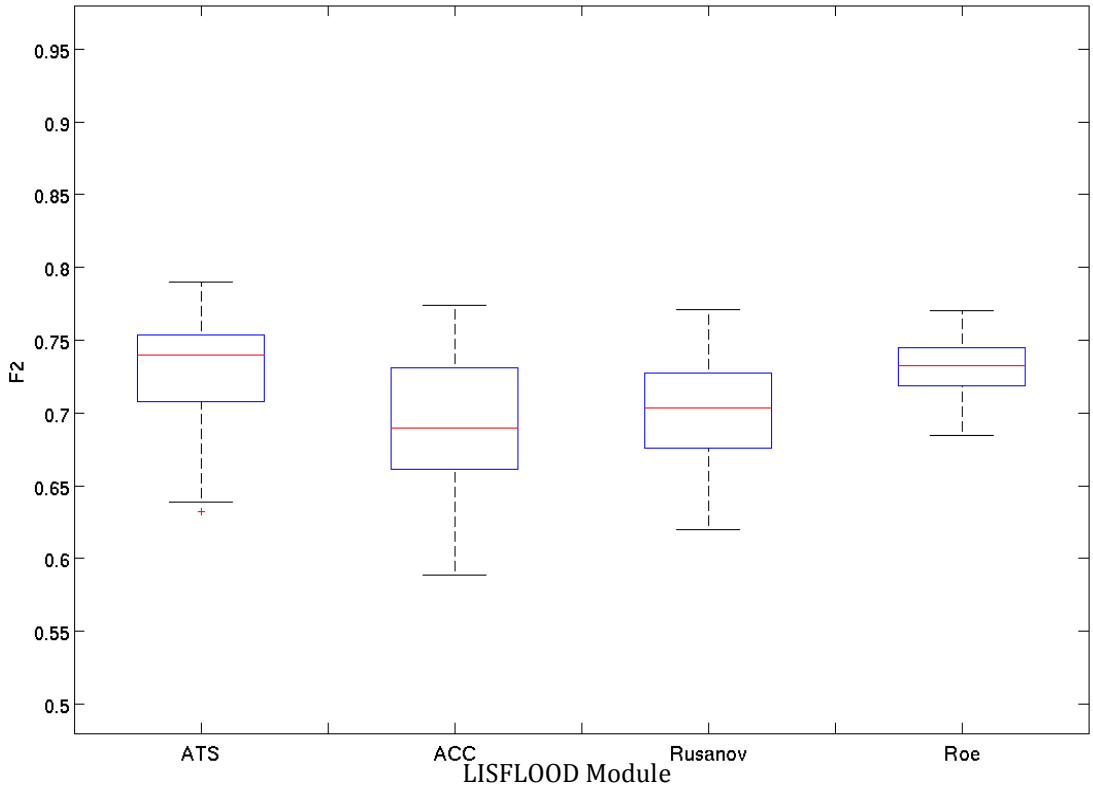

Figure 6.11: Box and whisker plots of the F^2 value for the LISFLOOD modules

The full SWE modules display similar results in comparison to previous test cases. The Rusanov solver shows a wide range of results in comparison to the Roe solver, but with a similar mean value. The cause of this is the Rusanov codes' higher level of sensitivity to the friction parameter. A reduction of sensitivity to the friction parameter is notable at the higher grid cell size, which also indicates a level of parameter interaction relative to the grid cell size, and therefore indicates a level of uncertainty associated with this factor relative to the use of a simplified numerical approach to the full SWE. This again furthers points from this test case and previous test cases about the relative benefits of using reduced computational costs numerical methods in assessing inundation problems. The Roe solver, by comparison, shows a high range of insensitivity to other parameters, and a reduced range of uncertainty. Here the conclusion is that, a higher degree of confidence can be given to the results from the Roe solver. What is more interesting with this module, is the relationship between F^2 and the uncertainty plot of Figure 6.7, which demonstrates that the Roe solver creates larger regions of uncertainty than the ATS solver, that do not necessarily impact the goodness of fit measure. This is in part explained by a region of the model domain (between 3000 and 4000m, and 1000-1200m, Figure 6.7, located around CP-12 and CP-13, Figure 6.2), which is consistently flooded in the Roe solver (bottom right Figure 6.7), but contains smaller regions of low frequency values in the ATS solver. This region is inundated entirely in the observed data set, which in the Roe model results will create a higher number of matched observed and modelled cells, which reduced the impact of incorrect cells in the F^2 calculation. Again this brings into question the use of extent comparison as a means to evaluate inundation models, where the number of cells correctly predicted may have a greater influence on the model results than those which are incorrectly predicted, but may be of greater interest to model developers and decision makers. A further point to note is the range of depths previously seen in the analysis of control points, which does not relate to a wide range of F^2 values. This again indicates not only the implication of using extent comparison methods, but also that local effects are not captured with the global model evaluation that may be critical to determining consequence of flooding or taking further action from model results.

The role of other factors in controlling model output appears to be greater in this test case than other test cases. The relative significance of the friction parameter appears to be high in this test case. This is further confirmed by plotting F^2 against Manning's *n* for each module (Figure 6.12). The friction value is based on the spatial distributed friction parameter set, with the value of friction for the vegetation area displayed on the x axis. The other parameters are fixed for this Figure. The

scale of the y axis is based on the same scale as the Glasgow test case to demonstrate the relative insensitivity of this tests case to the factor space.

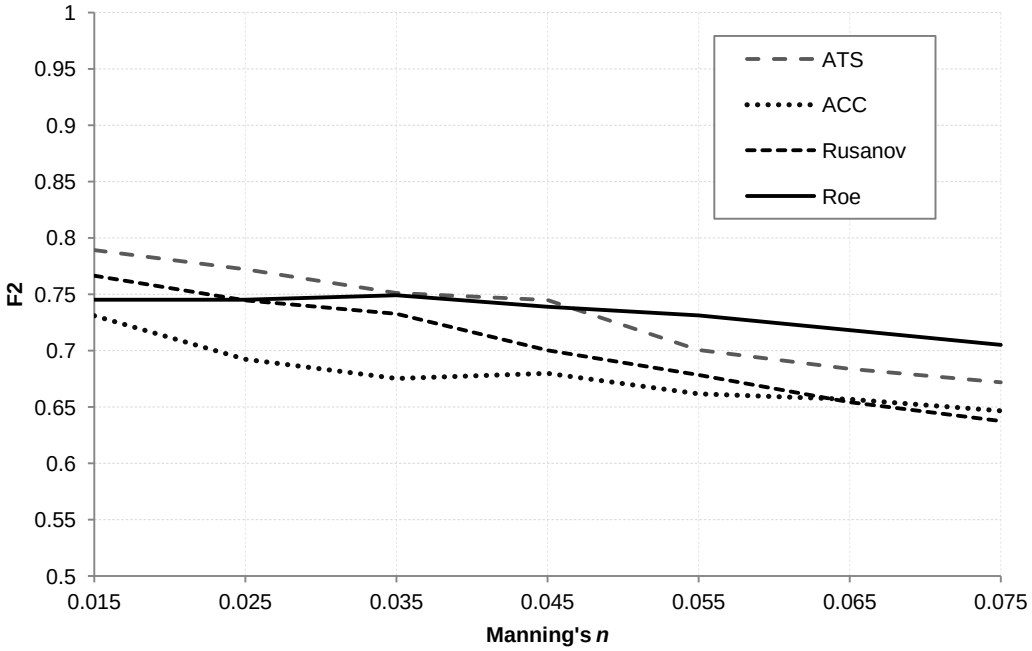

Figure 6.12: The F^2 value plotted over the friction value for each module

The peak values can be seen for ATS and Rusanov. the variation between each module is greater between at lower friction values, although this reduces in the centre of the figure, before increasing at higher friction values. This indicates that both choices are significant in influencing model results, and that some level of interaction is occurring between the parameter sets, similar to the Glasgow test case. In order to further explore the causes in variations of model results, the use of an interaction plot (Figure 6.13) is used to identify the role of other parameters in affecting model results. The Figure has been scaled to the same level as in the previous two test cases as a point of comparison, and, as with Figure 6.12, shows that in comparison to those cases, there is a significant insensitivity to the ensemble of parameters in this test case.

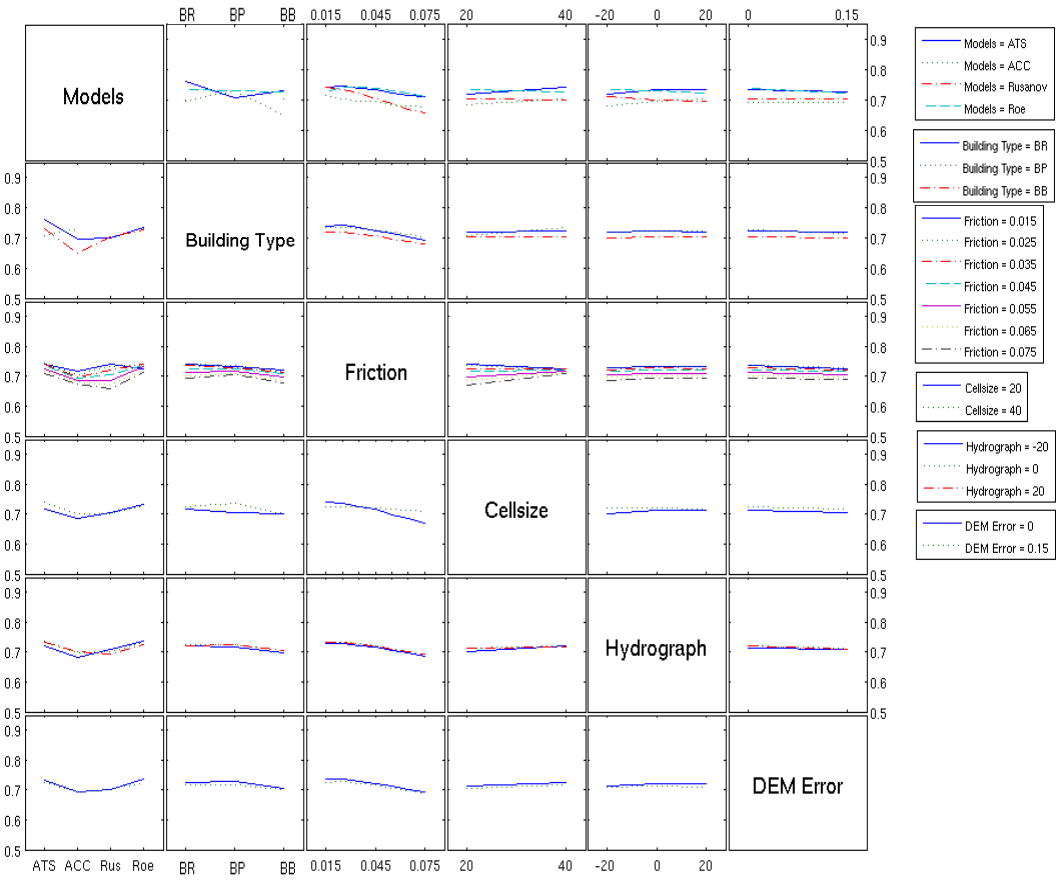

Figure 6.13: Interaction plot for each input factor comparing mean value of F^2 per level of factor other levels of factors.

As with the Coventry test case, there appears to be small amount of variation across the parameter range. This is caused in part by the large number of easy to predict wet cells. The majority of the domain is the inundated valley floor – smaller variations produce localised effects that the global objective function of comparison of extent ignores. The largest variation is caused by the choice of model (first column), although the influence of other parameters is noticeable. The choice of building representation appears to be influential as an interactive factor with the choice of module (column 1, row 2). The significance of the value of friction is also noticeable, particularly in comparison to cell size (column 4, row 3). As cell size increases (x axis) the mean value of F^2 converges for all levels of friction, indicating that at increased cell size, the significance of the value of friction reduces. However, at the lower cell value, the variation is noticeable greater. Analysis of column 3 which relates to friction values shows a distinctive gradient for this parameter, with reducing levels of model performance with increasing values of friction. This indicates a number of higher order interactions between parameters and friction value. The significance of these factors is only relative however, in comparison to the role of modular choice, which although significantly less responsive than in

previous test cases, still demonstrates a significant control on F^2 based model results.

Both the value of hydrograph and the DEM error are also of low significance and display little interaction with other parameters. This is not surprising given that the cell size is relatively large in comparison to the DEM error (+-15cm), and the issues relating depth of water to flood extent (Mason et al 2009, Stephens et al 2012), which impact the significance of the hydrograph in valley filling exercises, such as in this test case. The use of model evaluation functions that incorporate depth may provide more relevant insight into this aspect of model input.

6.2.1.1. RMSE

The use of using observed water depths to evaluate models, and the use of the root mean square error RMSE values of the difference between simulated and observed water depths has been evaluated previously (Stephens et al 2012). The advantages include a broader range of results and a more relevant function to determining the risk associated with flood inundation. In this evaluation technique, a value closer to zero indicates a better model fit.

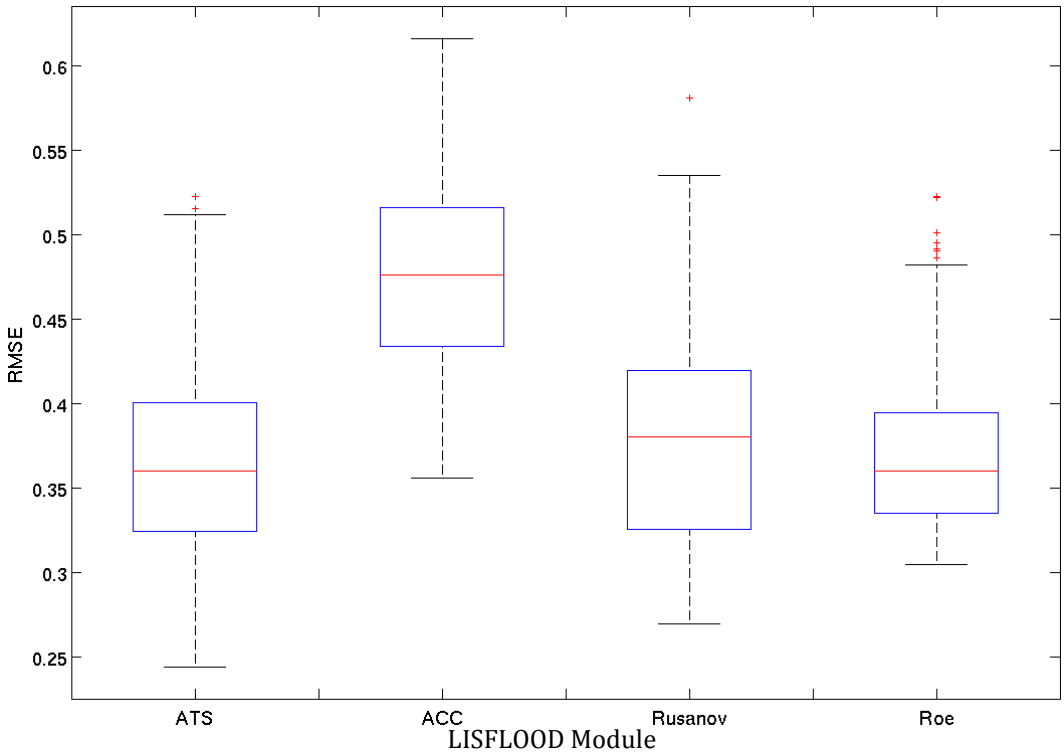

Figure 6.14: Box and whisker plots of RMSE of observed depth vs. modelled depth for each module

A number of patterns observed in the analysis of the goodness of fit function, are observable here. As with the F^2 value, the ATS module produces the lowest value, the ACC the widest range of results and highest mean, whilst the Roe solver has

the smallest range of results. The ATS module produces a best value of 0.2437 RMSE, which relates to parameter sets of a 20m non degraded grid resolution using the BP method of building representation. The standard hydrograph value was used and the value of friction was based on a distributed range with a low value of 0.015 for vegetated areas, the lowest value in the range of parameters. For the other 3 modules, the lowest values (0.3303 for ACC, 0.2695 for Rusanov and 0.2905 for Roe) are achieved using the lower value of hydrograph, a low friction surface, although with a distributed value for the Rusanov solver, and a single value for ACC and Roe. Both Roe and ACC had BP method as the building representation, whilst Rusanov used the Building Resistance (BR) method. The issue of equifinality is reduced with this function with a slight increase in the difference between absolute peak model, however it still remains an issue that other high performing models come from different sections of the parameter space. Whilst this again indicates the relative insensitivity of the parameter data set, the issue of determining critical factors becomes increasingly difficult. Further analysis of this space, does reveal some wider patterns between factors and model performance.

Using the same box and whisker plot as Figure 6.11, but using the RMSE value, Figure 6.14 shows the range of results for each module, and provides further insight into the relative significance of module performance. The most noticeable variation between the modules is that the ACC module has a lower level of performance relative to the other modules. This is a product of creating higher water elevations across the model domain as a result of the slower moving flood wave front seen in the analysis of Control points. The variation between the other modules is significantly lower than has been seen with the F^2 functions, despite the range of results greater for each individual module. Despite not being able to produce the same high value as the ATS code, the Roe module, produces a narrower range of model results, and a higher mean value (ATS 0.373, Roe 0.363). The Roe solver is also less sensitive to the value and distribution of friction value, and is also less sensitive to choice of cell size. The Rusanov solver follows a familiar pattern of providing a similar mean value, albeit with a greater range of water depth results, and a higher mean value. The implication of the results related to the Roe solver, is that the depths at CP-4 are local, with limited impact on the wider depth values. The indication from this Figure is that other parameters maybe more influential in determining model results and can be further understood in the interaction plot for RMSE (Figure 6.15)

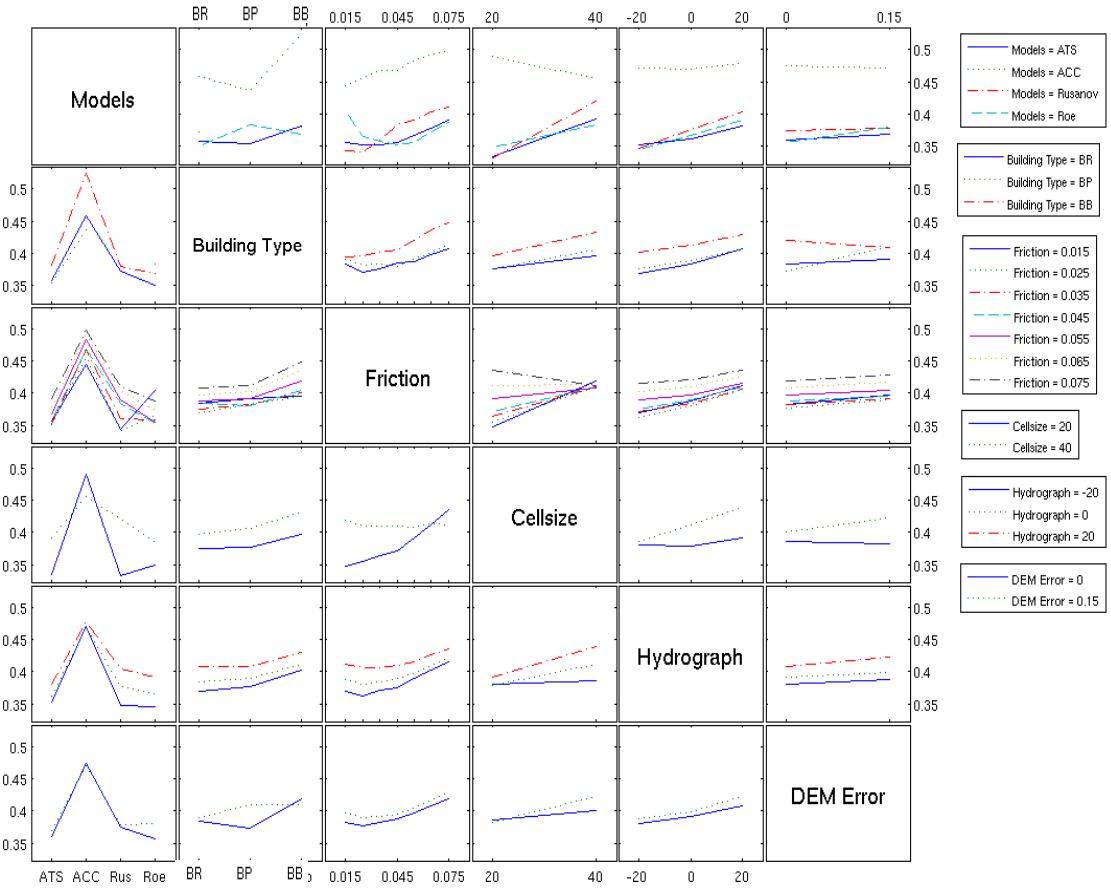

Figure 6.15: Interaction plot for each input factor comparing mean value of RMSE per level of factor other levels of factors.

The level of physical representation again appears to be the most significance factor in influencing model results, based on the gradients of the line in column 1. However, examination of row 1 shows that the large part of the variation in this column is caused by the ACC module. The other 3 modules appear to produce roughly the same mean value for every level of factor and follow similar interactions with other parameters. This indicates that the significance of this choice is being affected by the results of one module and not as a result of overall variations related to the level of physical representation.

Based on the gradients and interaction between lines, the level of friction value, cell size, and hydrograph all appear to be highly significant in determining the water depth value. The value of friction choice, again displays a strong gradient of increasing RMSE towards higher friction values (column 3), which confirms the significance of the findings from the control point analysis. Analysis of row 3 shows that the higher friction line (grey dashed line) is significantly above the rest of the in comparison with other factors. This indicates that a large amount of the uncertainty of the choice of friction value is determined by this extreme value. As with the F^2 value a strong interaction occurs with increasing cell size, where higher cell size

value creates a convergence to a similar mean level for all range of parameter (column 3, row 4). The significance of this is that with increasing cell size the uncertainty associated with friction value decreases, and therefore requires less justification in determining an appropriate value. That this is noticeable in F^2 also, indicates that this a key conclusion for rural based river overtopping models. Further analysis of cell size value (row 4) shows that at higher cell size value, a significant increase in the mean RMSE occurs, and a level of interaction can also be observed. This would indicate the significance of this factor in determining model results, and is responsible for the majority of the variation between module types in Figure 6.12. This may be in part due to the large increase in cell size between the two levels of factor used here, but does indicate that this factor should be well considered before modelling is undertaken. A noticeable gradient can also be observed with increasing hydrograph value and increasing RMSE value. Interestingly, the mean RMSE value is lowest at the lower peak value, indicating that the event may have been closer to a 1 in 100 year event. The linear relationship between increasing levels of hydrograph and increasing levels of water depth is also a logical one, but demonstrates the need for model evaluation methods to consider more than the inundation extent. By comparison with these 3 factors, the choice of building representation is less significant, although small variation do emphasis the significance of this choice. This is not surprising considering that a smaller percentage of the surface of the model domain is occupied by buildings, with the majority of the buildings occurring outside of the observed flood extent. A point that should also be considered in this analysis is that the RMSE for all models is outside the range of DEM error, and whilst the significance of this factor has proved to be low, indicates that the values of RMSE could be improved for this modelling test case.

The variations between the parameters and the higher peak value of RMSE provide a more robust model evaluation tool than the F^2 approach. The RMSE reduces the level of equifinaility between models of different parameter sets to provide an enhanced method of evaluating models. What this approach also identifies is the increased interaction between level of physical representation and factor choice, which is critical to further model development as well as increasing confidence in model results, through further understanding of the sources of uncertainty. A comparison of the two methods is undertaken in Figure 6.16. This approach has helped refine the calibration process and reduce the impact of equifinality across both model evaluation techniques. The pareto front of higher performance models can be identified at the lower right hand side of the Figure. As with other test cases

in this research and previous papers, the best overall model is a compromise between the mode functions F^2 and RMSE (Dung et al 2011).

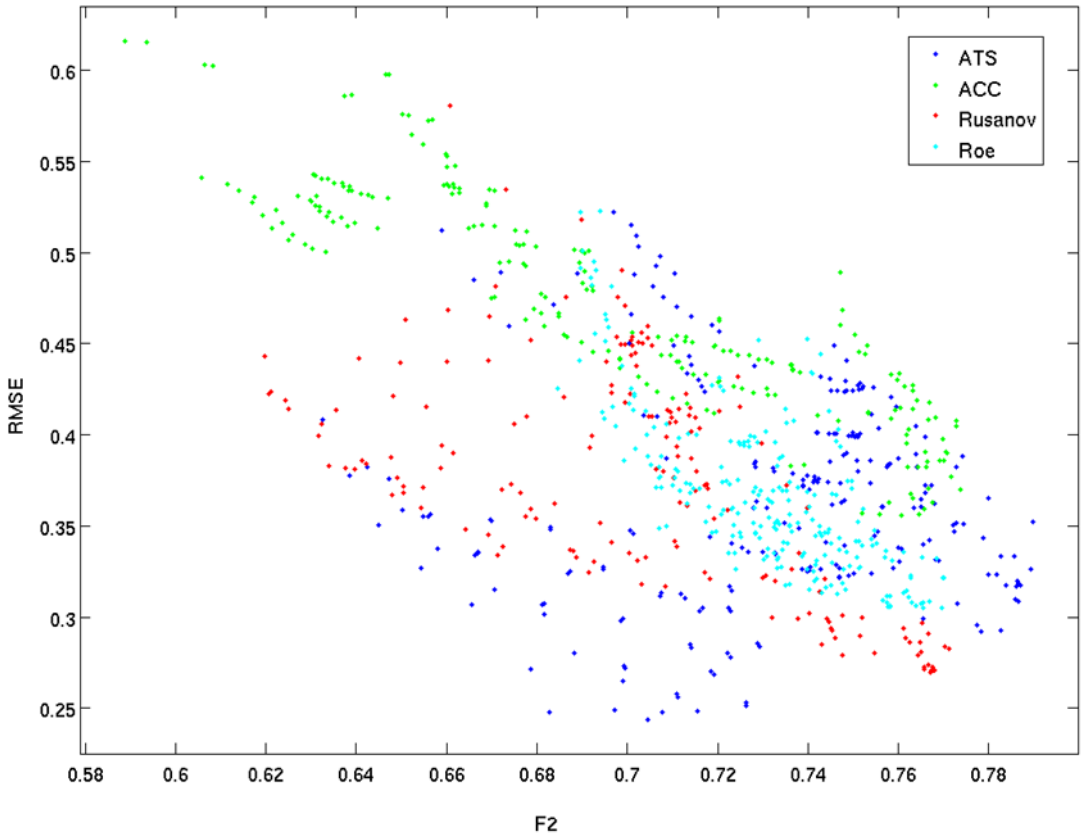

Figure 6.16: Combined results for F^2(x-axis) and RMSE (y axis), where each dot represents a model run. The pareto front occurs in the bottom right section of the graph

The pareto front is not dominated by one particular code, unlike the previous urban tests. Each module is represented within this lower portion of the graph and no clear pattern between the code emerges. The Rusanov code and the ATS code have a number of model runs that occupy this region, although occupy a wide part of this graph, further confirming the high level of uncertainty associated with these codes, but the ability of both to produce simulations with high model values. The relative insensitivity of the Roe solver is apparent with the narrow banding of model runs, whilst the wider range of ACC can also be spotted. Whilst the relative positioning of the pareto front is difficult to determine visually, it appears that a number of models can produce a high goodness of fit level as well as a low RMSE value. This indicates that the use of both approaches in unison can be used to refine the modelling calibration approach as well as highlight parameters with significant controls on overall results.

6.2.1.2 Distributed vs. Single value friction

The model evaluation techniques indicated that the choice of friction is significant, which increases the significance of how this value is implemented in this test case. A comparison of interaction plots related to the two methods provides insight into this. In the previous test cases, this choice created a wider range of model results, although the relative gradients that indicate interaction remain the same.

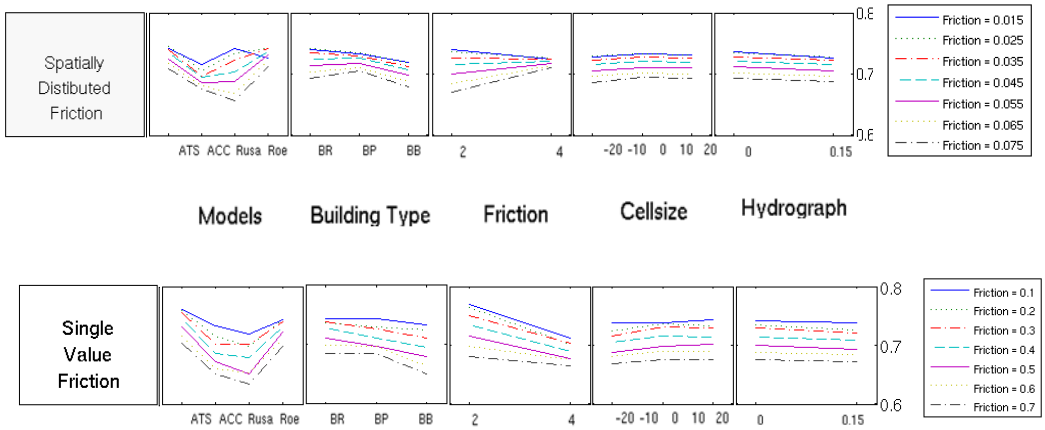

Figure 6.17: Comparison of mean F^2 level per parameter for spatially distributed friction and single value friction. The comparison parameter from left to right is model type, building type, cell size, hydrograph, and DEM error

Figure 6.17 is a comparison of spatial distributed friction (top) versus single value model domain friction (bottom) for F^2. Broadly, the same pattern of a wider range of results, with similar interactions and lower total function values, evident in the previous test cases can be seen here. This indicates that the inclusion of low friction surfaces reduces the uncertainty of this factor by reducing the variations between the higher friction surfaces, even in scenarios where the low friction surface occupy less of the domain, and have less control on the flow paths of the flood water. The increase in F^2 value also suggests that this is significant in improving model performance in this test case. However, a small variation can be seen on two of the interaction squares; square one which compares module choice and friction and square three which compares cell size and module choice. The variation for square one is caused by a decrease in the mean F^2 value associated with the lower friction values and the Rusanov solver. This instability at low friction surface has been noted in the previous test cases, but indicates that this model is particularly sensitive to this factor. The interaction between friction and cell size has been noted previously in this test case. Here the variation is a downward trend in F^2 value for a single friction value, rather than a slight increase in value which the distributed friction value creates. The indication here is that variations in cell size will impact the choice of friction and the distribution technique, leading to a change

in how significant this factor is in reducing modelling uncertainty. The impact is further noted in a comparison of approaches using the RMSE evaluation method (Figure 6.18).

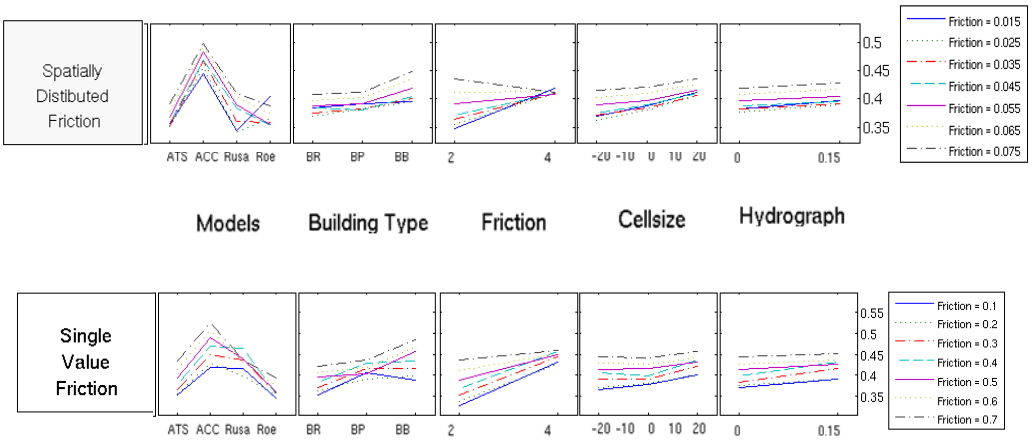

Figure 6.18: Comparison of mean RMSE level per parameter for spatially distributed friction and single value friction. The comparison parameter from left to right is model type, building type, cell size, hydrograph, and DEM error

A number of distinct variations of the model evaluation function between the two friction distributions are evident in Figure 6.17. Square one represents the interaction between model type and friction value. Both the ACC module (represented by the second peak) and the Roe module (represented by the third point on the x axis) have change between distribution types, with the ACC module having a lower mean value and wider range of mean values across the friction ranges for single value friction, and the Rusanov code having a small increase in mean level for friction values below 0.06. A distinct change also occurs between the building representation measure and friction distribution (second square), where for all values excluding the maximum range, a increase in mean RMSE value related to the porosity model occurs (centre peak, on the x-axis). This is related to the Rusanov flux, which at single friction values could not produce stable model runs and was consequently excluded from results, but could create a model runs at the single friction value, which produced a much higher mean value for this factor. This appears to skew the results at this section, although do not change the relative position of the porosity models as significant factor compared to the other building representation techniques.

As has been seen in previous test cases and model evaluation figures, the use of single value friction creates a wider range of model performance values, and a lower mean value. The key implication here is that explicit representation of low

friction surface reduces the uncertainty of the friction value and improving model performance.

6.3. Exposure Based Evaluation Methods

Considering the rural nature of the test case, the use of vulnerability and consequence based measures to evaluate model may not be optimal for model development. The three measures used to determine risk as a model output have produced results which demonstrate the usefulness of these measures even in largely urban areas, and have helped identify approaches that may reduce equifinality based across a single model function. Considering the previous two methods have demonstrated a level of equifinality, a methodology that could demonstrate reduced levels of this affect may be an important approach in future calibration approaches and modelling studies.

6.3.1 Weighted Vulnerability

The summary of vulnerability weighted goodness of fit measure (F^2VW) is given in Table 6.3, and the distribution and weighting values of cell in Figure 6.5. In comparison to the urban based test cases of Glasgow and Coventry and significantly lower portion of the model domain is divided by weighted cells, as a consequence of the rural domain, where less of the region is occupied by infrastructure and other features that will creating long term and short term impacts after the inundation event. As a result, the overall results display little variation compared to the original F^2 method. A key change is a slight increase in all maximum values and minimum values, where the maximum value matches the same parameter data set that had provided the high F^2 value for all modules. Plotting the F^2VW value over the same parameter space as Figure 6.12 and using the same scaling as the Glasgow test case, shows very little difference between the two methods (Figure 6.19).

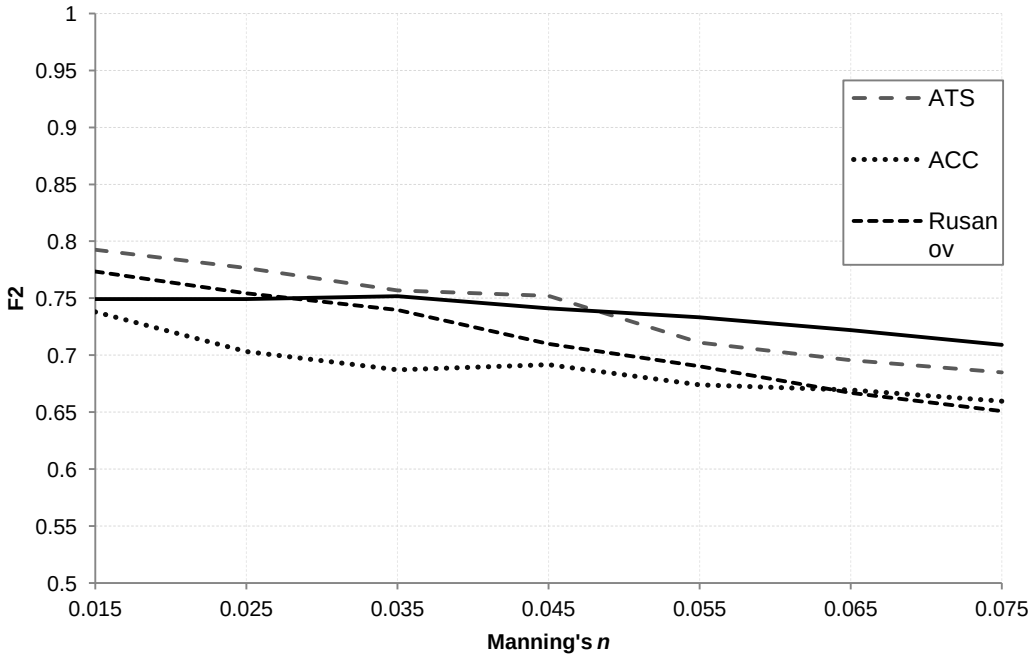

Figure 6.19: The F²VW value plotted against Manning's n for all modules.

A few examples of the F²VW approach in enhancing the overall modelling process are evident. A number of ACC modules have been improved with a higher minimum objective function value than had been seen with the F² value. This is observed for simulations that use the higher grid resolution value of 40m, and has the added impact of increasing the mean value to being at a similar level as the other modules. This is in part due to the majority of the higher weighted cells being close to the inflow boundary, where the ACC produces a consistent flood extent and generally has a similar level of performance as the other modules. This enhances the overall objective function for the ACC module, which is penalised for a poorer model performance at the downstream section of the model domain in the evaluation of F². This also confirms other indications of the sensitivity of the ACC module to cell size variation, which is here further proved by the impact of higher minimum values only occurring for models that use the 20m dataset. The same impact is not noticeable for any of the other modules and is caused by a slower wave front for the ACC module at 20m in comparison to 40m. For the ATS module, realisations using the 20m grid resolution BB methods include a wider range of both peak and minimum values, although by a relatively small amount, where the mean for the F² value for is 0.722 with a standard deviation of 0.04 compared to 0.725 and S.D of 0.05 for the F²VW method. This slight variation indicates that this method provides some small variations that could be enhanced to improve the overall method, but would require a different approach to the weighting values, either in the precise weighting values or the distribution of weighting cells. The fact

that the range remains similar to the non weighted method indicate that in these environments a more detailed approach may be required to create a function less susceptible to equifinality. An interesting point is that a slight divergence has occurred in the centre of the friction space between each modules, indicating that this method could enhance variations of objective function results where results are similar.

6.3.2 Regional F^2

Regional analysis of F^2 values has provided a key insight into model performance and modular performance across the model domain. Here, the regions were determined based on direction of flow, and provide not just information into areas of higher risk but also local hydraulic properties of the flood wave and model results related to this.

The relationship between regional values and the global F^2 value, and module type are broadly similar, with a convergence towards closer range of values apparent with increasing distance from the input boundary. This is evident from column one in the regional analysis figure (Figure 6.20), where the similarity of the spread of results and the positioning of the modular clusters for each row indicate that the relationship between global values and regional value are approximately similar. This is also apparent in the histograms of the total model results for each region at the top of each column, which display similar Gaussian distributed histograms with skewed means.

Further detailed analysis shows that modular performance is similar for each of the regions, hence the linear appearance of each cluster in Figure 6.20. The cluster of results for each section is also broadly similar, indicating that the variations in extents at the eastern extreme end of the model domain have a significant control on F^2 value. It also reveals further important insight into the significance of modular choice. For region 1 (column 2), the ACC module (green points) occupies a similar range as the other modules, despite the lower total global value (column 1, row 1). The ATS module (dark blue points) also displays wider ranges of points in region 1 compared to the regions 4 and 5. This indicates that a large proportion of the variations in ATS results are a product of variations in region 1, which also includes the storm drain channels referred to in the uncertain flood extents analysis. Regions 4 and 5 also demonstrate a linear relationship for the total range of model results. This indicates that with increasing distance from the boundary conditions, variation between model results remain approximately similar, indicating a spatial dependency between model performances.

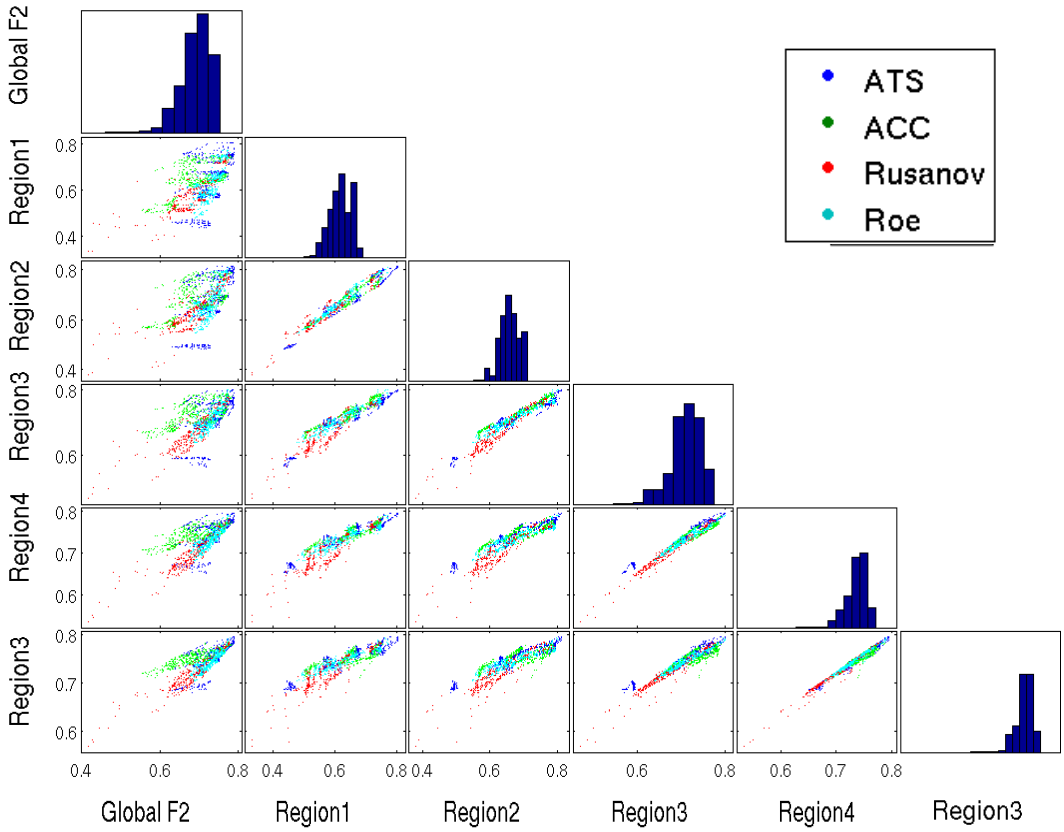

Figure 6.20: Comparison of regional values of F^2 versus Global F^2 value (first column). Each dot represents a model run, the colour relating to the LISFLOOD-module

Regional variations and modular variations converge to a greater extent at the higher level of grid resolution, where the majority of model results converge to a lower range of results than at the 20m grid resolution. A significant factor in creating regional variations is related choice of building representation model, particularly in the regions with larger percentages of urban areas such as region 1, 2 and 3. Here, the combination of this factor and modular type leads to the greatest variation across the input factor ensemble. As these regions contain the most building features this is not surprising, but does indicate even at larger model domains, the choice of these factors is critical to overall model results and a critical to decision makers. The use of this method in evaluating risk helps to focus model development and highlight these factors.

For this test case, the regional analysis has allowed further insight into the relationship between global values of F^2 and level of physical representation. The use of it in determining local risk factors has been less critical than in other test cases, where the relationship between local risk levels and global risk levels have been more varied than evident here. Further work would be required to understand if this a particular factor of either the application of the methodology in this test

case, or the broader interaction between test cases where varied hydraulic conditions are less apparent such as rural test cases with a greater emphasis on diffusion wave properties rather than numerical codes that contain momentum conserving properties or represent transcritical flows. A conclusion that can be made from this is that modular choice and local risk factors are broadly dependent on the level of other factors for test cases where the dominant process is diffusion based. Essentially, when the level of physical representation is less significant, the choice of model is less critical as an uncertain input, even when using the results to evaluate risk.

6.3.3. Analysis of Models using Depth Damage Curves

The majority of buildings are located at the western side of the model domain, and comprise 3 general categories of building type; residential, office and ware house. The water treatment facility is modelled with an electric warehouse depth damage curve. The results for total cost of damage across the module types are given in Table 6.3

Model	Mean	Min	Max	Standard Deviation
ATS	£5,370,398	£2,473,468	£7,117,214	£1,091,567
ACC	£4,927,666	£3,308,568	£7,270,588	£1,036,229
Rusanov	£5,494,452	£3,190,224	£7,013,866	£880,756
Roe	£5,261,410	£3,023,408	£7,095,962	£1,007,163

Table 6.3: Results for analysis of models using depth damage curves

The variation between the ATS code and the SWE modules is low, with a similar means and extreme values. By comparison the ACC has a lower mean value of around £5x10^5less than the other 3 modules. The highest estimated cost of the ACC module used a parameter set of the BP building representation, a degraded 20m grid resolution, with the higher peak value hydrograph and the lowest value distributed friction surface. The ATS module uses the same DEM and hydrograph value with the BR (building resistance) method and a higher value single friction surface, where as the Rusanov solver uses the BB method, the middle value hydrograph and a middle ranked distribution factor of 0.055 for the vegetation surfaces. The Roe code uses a similar parameter set to the ATS, but with a lower friction value of 0.02 (single value). Despite the variation of the ACC module relative to the other modules, the other modules produce a similar range of estimated costs, as demonstrated in the box plot figure of Figure 6.22. A number of key variations occur in further analysis of this evaluation function.

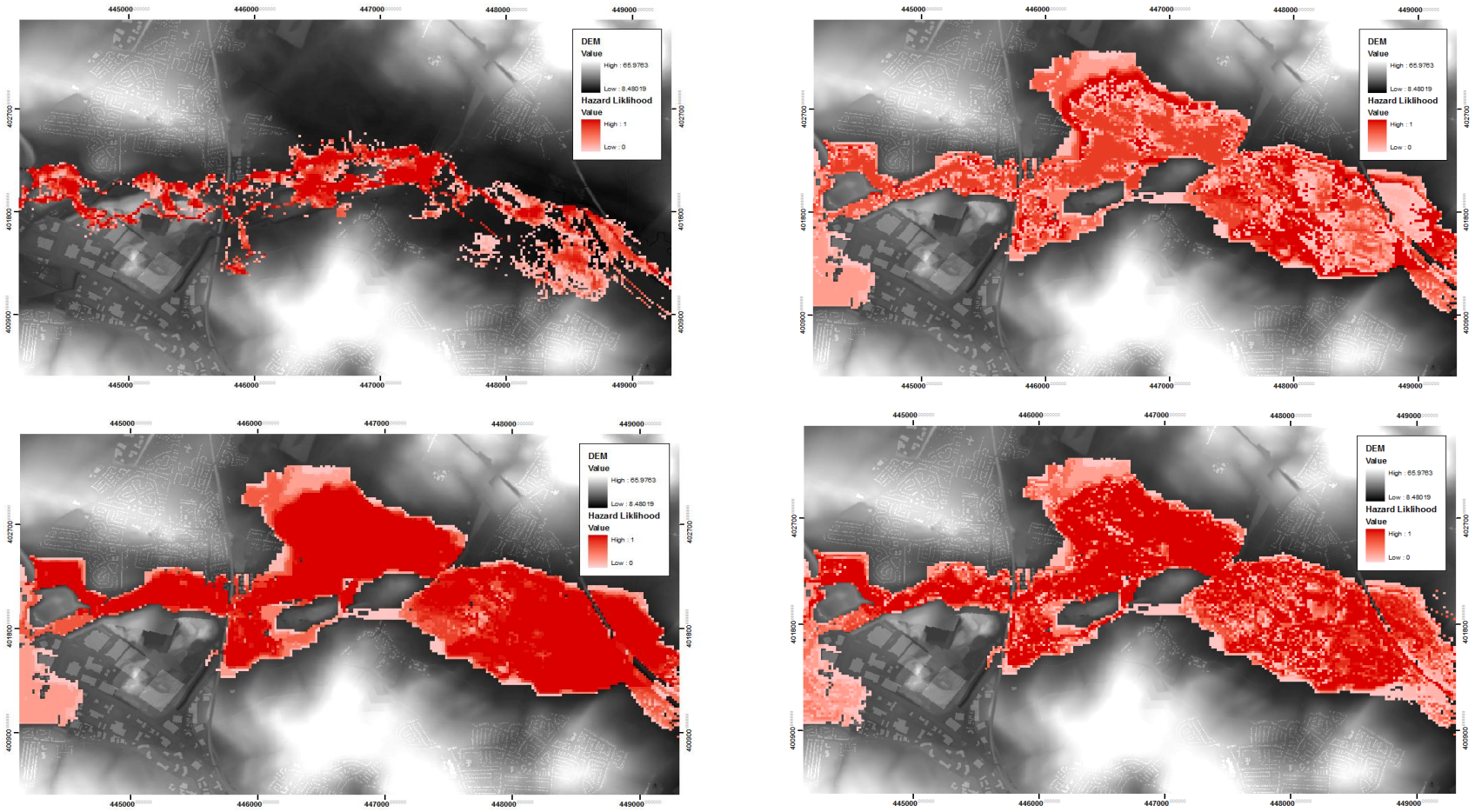

Figure 6.21: Uncertaitny Hazard Maps for the 4 modules where red regions represent high hazard, and the grading of the colour is related to the hazard frequency of the cell based on the total model ensemble. (Top left, ATS, Top right ACC, bottom left Rusanov, bottom right Roe)

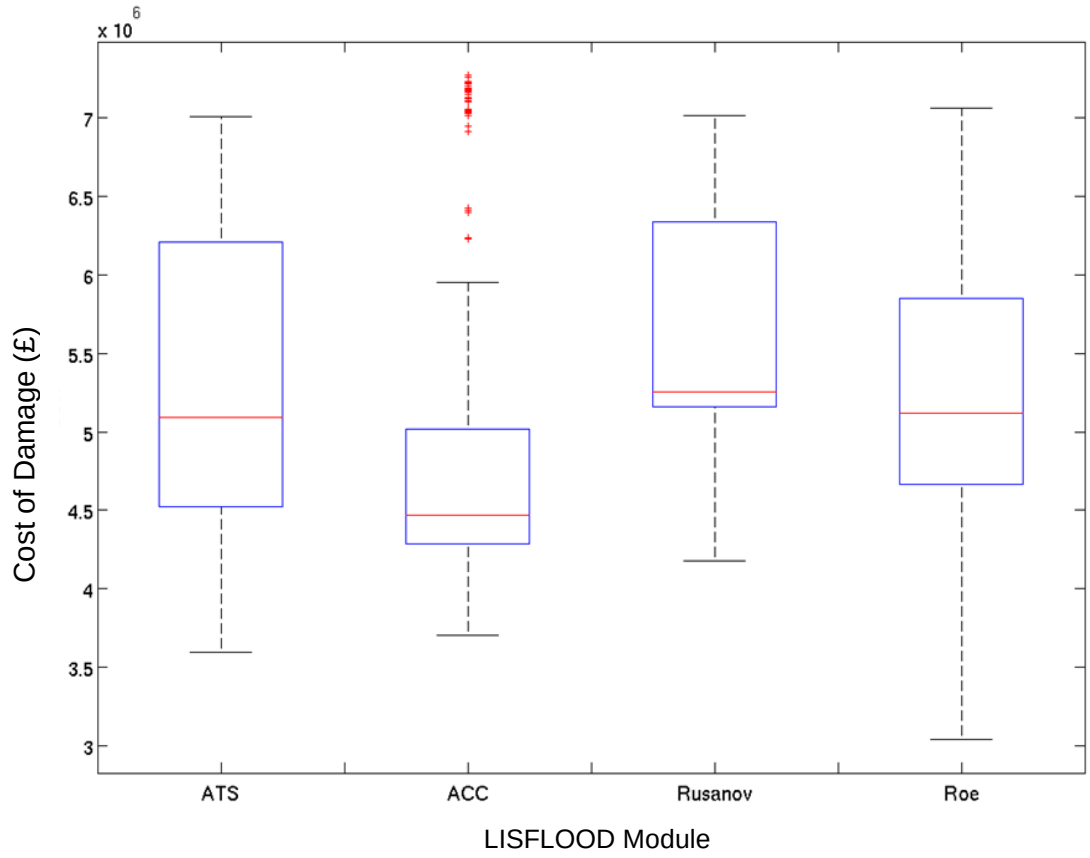

Figure 6.22: Box and Whisker plot of the modelled estimate damage cost of the event for each of the LISFLOOD modules.

In comparison to the previous test cases, the Roe solver now demonstrates the widest range of results, with a similar mean level to the ATS solver. This is a distinct change from the previous evaluation techniques where the Roe solver produced the narrowest range of results in both RMSE and F^2 evaluation techniques. However, analysis of CP-4, near the residential area of Bolton, undertaken in Section 6.2, revealed that the Roe solver had a wider range of depths than the Rusanov solver and the ATS module. The impact of local model results in critical areas becomes more apparent with this evaluation technique. Other patterns evident in the analysis of CP-4 are also evident here, such as the narrower range of depths the Rusanov flux method and ACC module created in comparison to the other two modules, reflected here as a narrower range of total costs. The consistent depth levels reproduced by the ACC module create a lower quartile range of model results. The ACC module does however produce the highest modelled depth values in the analysis of control points section. The location of regions evaluated with the depth damage curves are located slightly further away from these locations (see Figure 6.2) What this indicates is that water depth is distributed differently with this code leading to lower values at critical points in the model domain.

The overall similarity between modules does indicate that other factors are more significant in the total estimated cost. Using a similar interaction plot that is used to determine the effects of other factors on mean level of factors, further insight can be gained into the relevance of factor sets in affecting costs, the interactions between parameters, and the usefulness of this methodology in differentiating between factors and model runs.

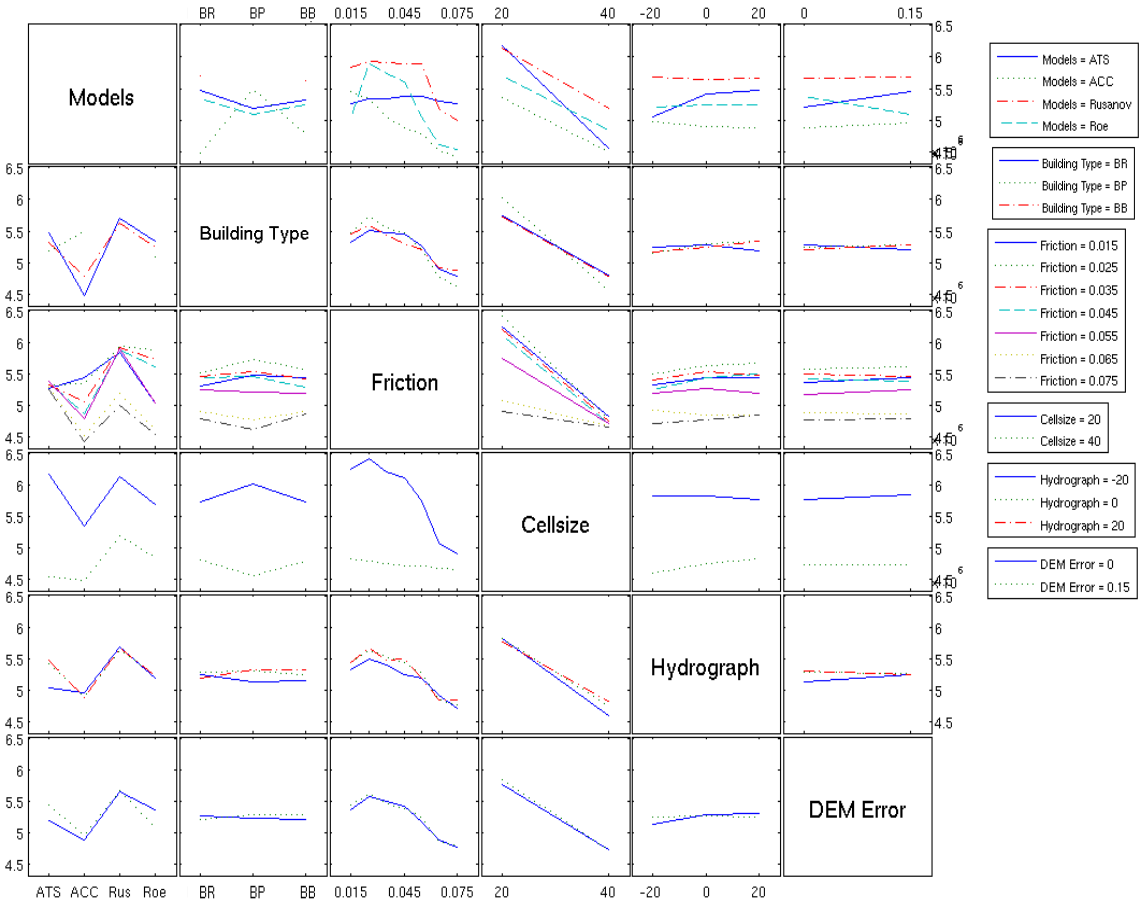

Figure 6.23: Interaction plot for each input factor comparing mean value of RMSE per level of factor other levels of factors.

Further analysis shows that the variation of mean cost level for the cell size appears to be the most influential factor (column 4 and row 4). At the higher gird resolution, a considerable drop occurs with the estimated cost and a reduction in the range of cost values estimated by the same parameter sets as the 20m models. This is evident through the gradient of the grey line in row 4, which remains relatively stable across the other factor levels and indicates that at higher grid resolution the significance of all other factors reduces. The sensitivity of cost to cell size has been noted before. Whilst here the sensitivity is less extreme than had been observed with the Glasgow set, the impact is still clear and indicates that cost based methods are sensitive to the methods used to implement them. To elaborate further, the cell

based method here used to evaluate cost is affected by the resampling process, and leads to increased sensitivity of the gird resolution. Other key interactions include the choice of friction value and the modular choice, which also display a level of interaction. The general trend for reducing estimated damage cost with increasing friction value can be observed in column 3, with the exception of increasing cell size which reduces the significance of this value. This implies that at higher friction values, less water is distributed to the extents of the flood margin, where the buildings that are affected by the inundation occur. Furthermore, the interaction with modular choice (square at column 3, row 1) demonstrates that these two factors in combination provide significant control over estimated damage cost. In this square, the relative insensitivity of the ATS module can be seen (blue line), which remains at a low gradient across the range of friction choice (x axis). By contrast the other modules display a higher level of interaction with the levels of friction choice, which indicates that this factor becomes more critical with increasing level of physical representation. These three factors appear to have the greatest control on the estimated cost. Interestingly the value of hydrograph appears relatively insignificant (column 5), despite the significance of this value to water depths as seen with the evaluation of RMSE.

The advantage of estimated cost as an evaluation method is that it enhances aspects of model output that may be undervalued or ignored by using direct model evaluation techniques. In this test case, the wide range of depths demonstrated by the ATS and Roe codes as a result of friction values at certain points in the domain are not critical in the global model evaluation techniques. By using a methodology that is based on both local factors and global outputs, more information about the simulation can be included in the model output. A key issue with this methodology is the sensitivity of it to cell size. Here, as in previous test cases, the level of cell size is influential on model result, which may require further development in order to ensure that it could be used in a wider application.

6.3.4. Uncertainty Hazard Plots

The Uncertainty Hazard plots in the previous test cases have provided key insight into the different hydraulic characteristics of the flood captured by each LISFLOOD module (top left, Figure 6.22). The ATS module produces a similar outline to previous test cases, with a narrow hazard extent centred on the main flow path, in this test case the River Dearne, with high hazard frequency values. As with the previous test cases, this indicates that the immediate hazard level may be under predicted by this level of physical representation, which may prove critical depending on the decisions taken from these model results. By comparison the

modules with higher levels of physical representation capture a similar hazard extent to their uncertain flood extents (Figure 6.8) but with lower hazard frequency values across the floodplain by comparison. The ACC module contains very few regions of high frequency (top right Figure 6.22), indicating a large degree of uncertainty across the parameter set. Noticeable regions of high frequency include near the margin of the flood plain in the centre of the model domain and the eastern extent of the model. This indicates a certain amount of uncertainty in the ACC module to predict hazard in terms of spatial analysis, but also highlights that some of the characteristics captured by the full SWE modules are also captured here. In comparison with the ATS module, it appears that the inclusion of the convective acceleration term from the SWE reduces the overall velocity of the flood wave, which in turn will impact the associated hazard level, but capture a wider level of processes, in that velocity levels are maintained across the flood extent, unlike the ATS module. The ACC results represents a key trade off between the reduced physics model approaches, and computational efficiency.

The Rusanov solver by comparison produces a large extent, similar to the flood extent plots of Figure 6.8, with high levels of frequency, which relates to the higher velocity speed at the point of initial inundation in comparison to the other modules. The regions of lowest uncertainty also relate to the regions of uncertainty identified by the uncertain flood extent indicating that the frequency values of the two plots are related. This is similar to previous test case results, indicating that the Rusanov solver may over estimate the level of hazard in all scenarios regardless of the hydraulic conditions. The Roe code has also produced a similar outline in comparison to the flood extent plot, but like the ACC module produces a wider range of frequency values across the model domain, than the Rusanov solver, but a greater percentage is of a higher value. A number of regions with greater uncertainty are similar to those identified in the flood extent plot, such as the western margins of the model domain, and contain similar value indicating that these are related. A spatial dependency is also apparent with the Roe solver. Increasing distance from the input boundary appears to increase the regions of frequency values less than 0.9, indicating a gradual diffusion of water and velocity across the site, which is intuitively correct. This same process is not captured by the other solvers, which tend to show a more sporadic arrangement of lower hazard frequency values. The margins of the hazard extent also demonstrate a lower frequency value. This too would appear to be intuitively correct, as water levels and velocities reduce with distance from the source. This provides further insight into the nature of hazard and vulnerability across the model domain, and also indicates the importance of this method in not only determining regions of hazard but also how uncertainty in modelling progresses to other decision making processes.

6.4. Sensitivity Analysis

6.4.1. Morris Method

The F^2 value and RMSE value are both evaluated using a 4 level, 5 factor parameter space, which exclude the building representation factor from the original parameter space, due to instabilities with the Rusanov solver and the porosity model. Distributed friction values were used instead of single friction values as a result of their higher performance level indicated in the systematic analysis section. Considering the level of equifinality demonstrated by the F^2 function the use of a screening method may not be able to distinguish between the different factor levels. The results for F^2 are displayed below and show this impact (Figure 6.24), and the RMSE effects (Figure 6.25) shows a similar pattern

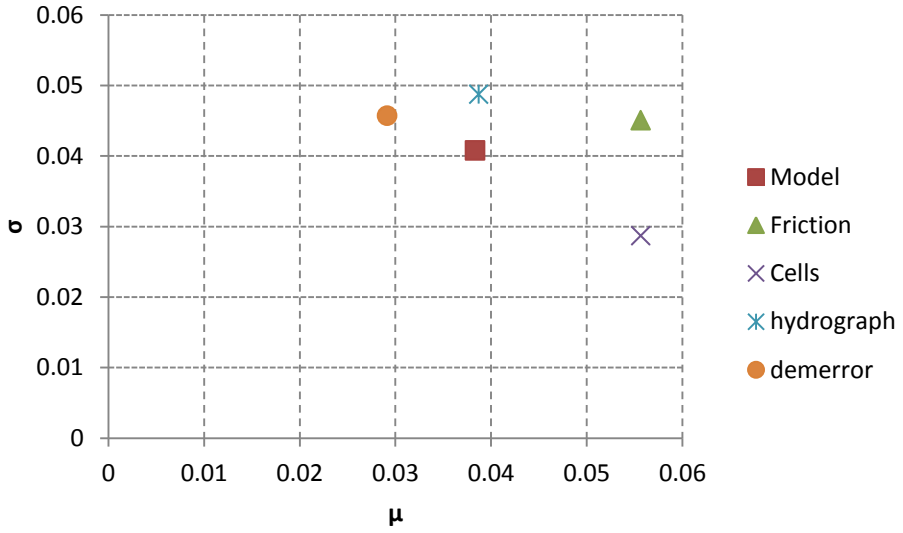

Figure 6.24: F^2 Morris Method results.

The indication from the Morris result is that the whilst the friction and cell factors appear to have similar first order influence and are ranked first from the parameter space, the small variations indicate that each factor is similar in terms of contribution to model results. The ability of the Morris method to discriminate between factors of equal significance has been noted before (Pappenberger et al 2008), and here where the level of equifinaility is relatively high, the same impact is noted. The combination of friction and cell size as an influential factor has been observed in the interaction plot of Figure 6.12, and the Morris method further confirms this factor. The confidence in this result is however reduced by the similarity between factors. A key requirement of Sensitivity analysis method is the well posedness of the questions (Saltelli et al 2000). Here the indication is that due to the level of equifinality, a screening method may not be suitable to distinguish between model simulations.

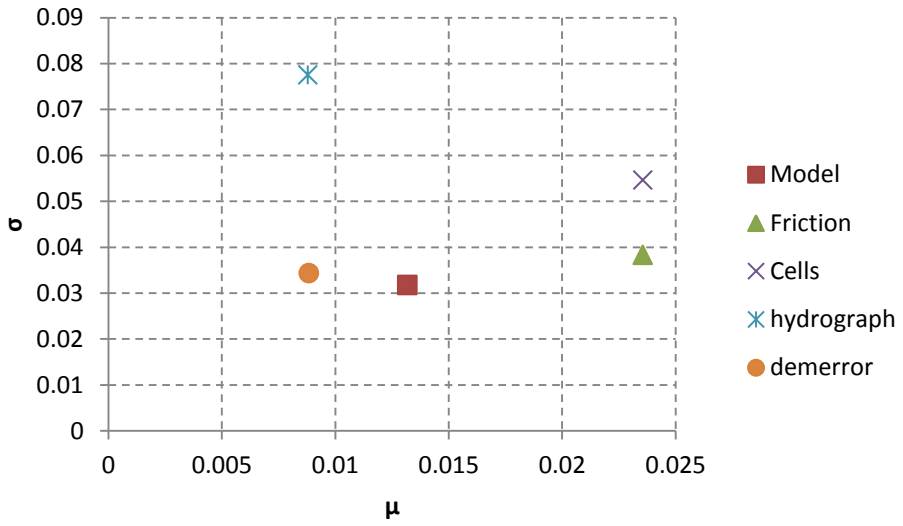

Figure 6.25: F^2 Morris Method results.

A similar pattern also emerges when the RMSE method is also considered. A similar ranking of factors occurs, with both the friction and cell factor indicating high first order influence. The hydrograph appears to contain a low first order but very high interactive component, which can be observed in the interaction plot of Figure 6.14. Again a similar issue occurs, considering the equfinality of model results creates a small range of results with which to rank the factors. The overall indication is that the Morris method may be unsuitable in scenarios where the range of model results is low.

6.4.2. BACCO-GSA analysis

The results from the BACCO Gaussian emulator based Sensitivity Analysis are summarised in Table 6.4. As with previous analysis, a LP-Tau sampling method with 100 training provides the best emulator fit in comparison to a Latin Hypercube approach. As with previous tests the aim here is to investigate further the influence of parameters on a simplified physics model and a full SWE module, rather than the influence of the whole parameter and factor ranged used in the systematic analysis section.

Parameter	Total effect (ACC F^2)	Total effect (ACC RMSE)	Total (Roe F^2)	Total (Roe RMSE)
Cell	56.0	24.2	46.7	36.31
Hydrograph	85.83	91.09	71.2	91.52
LiDAR error	8.11	11.24	56.8	1.13
Friction	42.07	36.63	41.23	35.65

Table 6.5: Results from the BACCO GEM analysis

Each module displays a similar sensitivity to other factors, again demonstrating that each module has a relatively similar output across the parameter range, and is therefore less critical by comparison. The hydrograph appears to be the most significant input factor with a lowest total effect value of 71.2. This appears to differ from the systematic analysis section, which demonstrates a relative lack of influence to the mean level of model performance for F^2 with a higher contribution to the RMSE. Both the friction value and cell size also appear to be significant contributing factors, which has also been seen in the systematic analysis, and the Morris Method analysis. The high level of total effect indicates a strong interaction process between factors, which had also been partly identified in the Morris method analysis. The overall indication is that these factors all influence the modelling output. Reduction of uncertainty for this test case, would require further detail of all these parameters, which indicates the complexity of solving inundation problems.

These results should be considered in relation to variation with the previous test case. The different parameter sampling strategy of the two approaches, two sensitivity analysis sections, and the exclusion of both building representation and physical representation factor would affect the relative significance of the factors. Inclusion of these factors would depend on using a means of describing their distribution in a continuous means, rather than a discrete method that has been used in the analysis in this section. This represents an area of further development for this research.

6.5. Conclusions

The Mexborough test case has provided a useful counter point to the two urban based test cases of the previous chapters. The results indicate that whilst the level of physical representation provides variations in certain aspects of model results, the overall significance of the level of physical representation is less than in urban environments. This compares with previous results which indicate both the dominance of diffusive conditions in rural environments and the similarity in

objective functions between the simplified approaches and full SWE models at larger cell size (Bates et al 2000, EA Benchmarking 2010). Both model evaluation techniques had issues of equifinality. The use of multiple objective function approaches and the use of risk based methods provide approaches for reducing this impact and refining the calibration approach. The estimated cost approach has also highlighted a number of properties displayed by the modules which were not captured by the traditional objective functions, mainly the range of depths displayed by the Roe solver. This indicates the potential usefulness of this approach as well as providing a calibration method with greater differentiation between model results.

The issue of model runtime duration is clearly highlighted in this test case. This can often be an arbitrary value determined by a modeller, but here a clear bias has been highlighted against simplified approaches which may have clear implications in terms of Cost of Damage and economic loss. Here, the reduced model runtime impacted the ACC code by not allowing sufficient time for the model to route the flood water and reach the downstream boundary. This is a potential issue which should be clarified in future modelling of this region, but a number of points emerge from this. First that model runtime is itself an uncertain parameter and second that the ACC code requires a longer simulation time in order to replicate the complete flood processes. The bias in this test case helps to highlight this point. A further implication is that the variation between model codes may be further diminished in terms of modelled extent comparison. The results and analysis from this chapter consider this implication, but it should be considered further in future modelling exercises.

A number of issue and further modelling exercises could be undertaken. First the use of a explicit river model will provide further insight into modelling approaches and to further explore methods for modelling these events. This is particularly key as previous studies have indicated the importance of these factors in model results (Hall et al 2005), and will help determine appropriate approaches to modelling river overtopping events. Secondly, the use of hydrographs that approximate the recorded values from nearby river gauges. The gauged time at peak value (or values over the recorded maximum of 56m^3/s was over 3 days). Due to computational costs, this was reduced to a 15 hour peak time and a shorter model runtime. Although the results indicate a reasonable modelling process has been undertaken, further work could help to investigate the importance of this factor. Thirdly, the methodology behind creating observed water levels could also be explored. Here the methodology was based on a nearest member sampling method based on the GPS positions of the extent marks and depths in the DEM. Further work could investigate the importance of either the sampling approach, the errors

between GPS and DEM elevation and the techniques used to evaluate model performance. This test case has removed these issues due to computational constraints, but a detailed investigation could identify future issues and uncertainties related to the evaluation of modelled depths and observed depths. Finally more direct SA methods could also help to identify key factors. The sensitivity analysis used here is a combination of computational efficient methods and screening based methods. A more detailed approach, that uses direct model results and knowledge of the exact distribution of parameters rather than approximation of the model response surface, may overcome the equifinality issue of the objective functions, to make a clearer distinction between significant factors.

7. DISCUSSION AND CONCLUSIONS

7.1. Overview

The aim of this research has been to investigate the uncertainty associated with the level of physical representation in flood inundation, and to investigate how this impacts on model results and evaluation techniques. In order to provide a comprehensive analysis of this impact, the level of physical representation was compared to parameters and model inputs which are used in inundation modelling problems. This input parameter set was then used in three test cases, each of which had different topographic feature and hydraulic properties, which has reduced the potential of model bias from impacting the results. Each test case used a number of evaluation techniques based on two general approaches; traditional observed data approaches, and modified approaches based on risk, vulnerability and financial consequence. For each test case, the goodness of fit measure F^2 has been used to evaluate the model, which has been consistently used as a evaluation function and provides a means of not only identifying trends between model types and test cases, but also direct, smaller scale model outputs and other evaluation techniques. Finally, two sensitivity analysis techniques were employed to quantify the significance of other factors.

The overall conclusions from each test case indicate that the level of physical representation has a significant control on most of the model functions and model results. For both the Glasgow and Coventry test cases, this factor controlled most of the variance in model results, both in evaluation techniques and model outputs, such as depth and extent. The impact is also monotonic in the Glasgow test case, where levels of increasing physical representation are matched by alterations in the output and evaluation techniques. This confirms previous findings about the significance of the level of physical representation to accurately model inundation in urban environments (Hunter et al 2008, Neal et al 2011). However, the significance of this factor appears to be less in the rural test case of Mexborough, where diffusion wave conditions dominate and transcritical flows do not appear in same manner as in the previous test cases. The choice of physical representation creates numerous local impacts which are often over looked by global model measures, and can impact other measures taken from the modelling results such as exposure. A key impact of the choice of the level of physical representation is the significance of other factors, for example the sensitivity of ACC to cell size, the impact that friction has on ATS, and the sensitivity of the Rusanov solver to all parameters. The

implication is that with changing physical representation and numerical complexity of the code, the significance of other factors changes in a non-linear way.

7.2. Discussion

A number of trends are noticeable across the test cases relating to the precise impact of factor choice on model output. The choice of building representation impacts how well flow paths in the model can be defined, which can impact both extent and the modelled hazard level in key locations. As with the variation of results caused by the level of physical representation, the impact is often localised in the model domain with variations in the water depths and extents occurring near the location of buildings. The global impact is often small, compared to other factors but can impact the level of hazard associated with individual buildings. For all the test case, the use of spatial distributed friction methods and uniform values were compared. The explicit representation of low friction surfaces appears to have two effects. First, it reduced the level of uncertainty associated with the value of friction by reducing the range of model results. Secondly, the model performance was higher where low friction surfaces are represented than where a uniform value is used. The fact this occurred for all three test cases, including Mexborough, which is predominantly rural and contains a lower percentage of surface occupied by road networks, indicates that the inclusions of these surfaces is significant in further modelling exercises. Further work could be undertaken into determining the precise nature of this, and in determining an optimal number of friction surface categories. The impact of inflow hydrograph uncertainty also appears to be low in terms of global effects and relative to other factors. The comparison of extent often displayed little variation to the depth of water in the model and between the different inflow hydrographs, but other methods appeared to show greater response to this factor, such as the RMSE and Nash Sutcliffe coefficient. Overall though, the impact of the inflow appears to be responsible for controlling the shape and hydraulic characteristics, rather than the precise distribution of depth values, which is also controlled by friction values. The overall indication is that interaction between friction, inflow and physical representation is responsible for extent and depths, and where depth of water at specific locations is required these factors must be carefully considered. A detailed calibration process is required to reduce the uncertainty of model results. A further conclusion from this work is the relative insignificance of the error associated with the DEM, which showed little signs of significance contribution to variations in model output. However, the method to assess this factor may be a contributing reason to this, where the error is randomly assigned between the maximum value and zero. Further information on this may allow the DEM to be

degraded in a non random way, to define spatial patterns between error, surface features and measures elevation. This way a further conclusion could be drawn about the significance of this error in a spatial context rather than a broader global level of uncertainty, than has been made here.

The model evaluation techniques also provided further insight into the relative values of parameter significance. In each test case the use of extent measure provided a useful comparison point to previous research, and provides an indication of the response of module performance, particularly in urban areas. Methods that go beyond the approach of binary comparison of extent are increasingly popular (Stephens et al 2012), and provide different properties of the observed data with which to test the model outputs. In terms of evaluating uncertainty in modelling, the use of alternative evaluation functions has demonstrated two points. First, that the method of evaluation has a distinct impact on what factors are identified as critical. For example the use of extent indicates the importance of friction for most test cases, and less critical is the value of inflow. This however changes where depth becomes a key element in the model evaluation technique, such as Nash Sutcliffe in the Glasgow test case or estimation of depth in Coventry. Secondly, that combined approaches can significantly overcome the issues of equifinality, even in scenarios such as Coventry where a single uncertain depth measurement has refined the parameter evaluation method. A combined multi objective function analysis can also refine the number of best performing models down to a smaller selection (as in Dung et al 2011). Using other uncertainty methods could then be employed to further understand the relationship between factors and model outputs as well as best performing models. A modified GLUE approach could use these techniques to create a smaller final sample of behavioural models, which can further reduce uncertainty in inundation modelling

The use of risk based methods to evaluate model output has been undertaken in previous research to create functions with greater differentiations between models, and to provide more direct communication methods with decision makers (Pappenberger et al 2008). In this research, the use of vulnerability weighted approaches was mixed in terms of providing a more refined calibration method than the traditional method. For the Glasgow test case, a clear distinction was made with the overall spread of results, however for both Coventry and Mexborough this was reduced, as a result of either too many or too few weighted cells in critical areas. The issue is one of how to weight the regions in a model domain where model and observed prediction do not match, by using assets which are determined as being of higher consequence as the weighting component. A clearer distinction between

models is required in order to provide a comprehensive approach to modelling, and the subjective nature of the approach may create more uncertainty in the process than reduce it. The regional approach provides not only information about model performance in critical areas; it also provides the opportunity for decision makers to have further input, to the modelling process. In this research this method has provided considerable information on local model variations, as well as provides information concerning the interaction of global evaluation functions with regional results. Finally the use of a cost of damage based approach has provided further detail of model performance, wider ranges of model variation with which to evaluate the models and a means of communicating uncertainty. In the case of Mexborough it provided a form of evaluation that provided more insight than the traditional evaluation methods. Moreover, it linked global model performance to local variations in critical areas in a method that neither of the direct model calibration methods could. Further work could look at where it would be appropriate to use a methodology that also included depth damage approaches.

Broadly, in order to reduce the uncertainty associated with the different levels of physical representation depends greatly on the modeller's ability to determine key hydraulic characteristics of the model, and to chose a model that reflects this. A broad rule from the conclusion of this research would be to base this on the underlying topography and grid resolution. Where complex topography, such as urban environments are prevalent, and where the underlying topography is not complex, the level of physical representation must be carefully justified and full SWE modules should be used. However, where the underlying topography contains steep, well defined bed slope such as present in the Coventry test case, the significance of this factor lessens. For rural domains this reduces further, indicating that underlying topography and not just the representation of surface features is an important control on model results and the choice of input factors.

7.2. Conclusions

The main conclusions can be summarised as following;

- The level of physical representation is a significant factor in all scenarios. However it is less critical in larger scale rural modelling exercises, where the diffusion wave properties common to all the codes tested here become dominant. This is similar to conclusions in previous research, but here it is demonstrated in a more systematic way, in comparison to other parameters that may also contribute to uncertainty.

- The level of physical representation will impact the model in a local manner, such as urban road networks, where conservation of momentum is required. This may not be significant at a global scale.
- Simplified approaches to the governing equations, such as ATS and ACC, and simplified numerical approaches to the full SWE, such as the Rusanov code, can provide a useful and potentially computationally cheaper inundation model. However, the large variation in results with these codes indicates uncertainty, and the reliance upon a reliable calibration process.
- The method of evaluating model performance is critical to determining the significance of parameters, and where possible multiple evaluation functions should be used to reduce the level of equifinality of individual functions and reduce the number of 'best performing' models.
- The use of subjective vulnerability approaches can enhance the calibration process in two ways. First, it can provide more information related to local impacts within the modelling domain either through enhancing regions designated of greater significance (F^2VW) or providing regional breakdowns of global factors (Regional F^2). Secondly, it could be used to provide more direct ways of communicating uncertainty with decision makers whilst enhancing the knowledge of local model output (through depth damage curves).
- Interaction between parameters is strong, and the choice of level of physical representation will impact on what other factors are critical, but the effect is not linear with each code producing a different reaction to the parameter set. In particular the practice of adjusting the value of Manning's *n* friction value to calibrate model may not take into account the insensitivity of the underlying model and may be insufficient to account for the uncertainty in the modelling process.
- Sensitivity analysis has shown that the level of physical representation is a significant factors, when compared to other factors. This must be considered in respect to the assumptions that were made in applying the SA methods.

A number of additional points must be considered with these conclusions. The choice of the three test cases was designed to reduce the potential of model bias occurring in the results. This is where a single test case with prevalent hydraulic conditions may impact the decisions and conclusions made. Here, this has in part been avoided, but a number of issues are raised in the analysis and aspects of the test cases could be improved to provide a more comprehensive analysis. The grid

resolution had been set based on issues of computational costs as well as feature representation. For both urban test cases, this could be further refined to a 1m grid level, at which level further features can be included such as road crown features which impact model results (Fewtrell et al 2008, Sampson et al 2012). This would provide insight into how refined the topographic surface needs to be to reduce uncertainty related to topographic features. Moreover, the use of a random degraded DEM to represent the modelling error could also be improved upon by using knowledge about the structure of the DEM to refine how error is attributed to the original LiDAR surface. A methodology that uses knowledge of the distribution of this error will provide further insight into the importance of this factor (Tsubaki et al 2013).

The sensitivity analysis approach used a number of assumptions. First, in relation to the Morris method, the model inputs were divided to create four discrete intervals, so as to allow the level of physical representation and the building representation type to be included as a model input. Parameter independence was assumed, which may not strictly be the case with the grid resolution and DEM error. However, the value of the LiDAR RMSE error is not strictly a product of the cell size, which is adjusted by the modeller, and can be assumed to be related, but independent. Conclusions with this method must be considered within this context. Secondly, the assumptions of both the BACCO-GSA and Morris method are based on inputs having a Gaussian distribution. Again this is not strictly accurate as previous research has found (Hall et al 2005), although calibration methods that have used a Bayesian Gaussian approach have proved successful (Hall et al 2011). As with the Morris method, the conclusions must be appreciated in this context.

Concerning the risk based methods the use of subjective methods to assess model output may increase uncertainty rather than remove it. For the vulnerability approach and the regional analysis there is a large number of assumptions based in defining the weighted areas and values. The concept is to allow decision makers an opportunity to engage in the model building process. The methods should be seen as assisting the calibration process rather than a definite form of evaluation function, a conclusion reached by Pappenberger et al (2007), who developed the concept. Here the use of these techniques has been furthered, by using them in multiple scenarios, comparing them with other standard techniques, and using unique approaches to determining the weight and location of cells that contribute to these exposure based calibration methods. These methods have potential to refine the calibration process, but further work could enhance this by determining the

uncertainty associated with the weighting of cells and the positioning of the cells. The estimation of damage is also subjective; the choice of cells that contribute to the calculation, the choice of depth damage curves and even the methodology used (MCM) are all based on subjective choices. The uncertain boundaries of the depth damages curves were not included in the analysis for ease of communication, but these values would in some of the test cases exceed the total variance caused by the other parameters. Further work could look at using depth damage curves from other techniques and other models, such as in Apel et al (2010). The use of these methods must be considered in this aspect.

7.3. Further Work

Further work could be undertaken by expanding the parameter space. The Mexborough test case has provided some insight into the parameter space of the floodplain in a river topping exercise. This test case could be expanded to include a river channel model, which has shown previous evidence of significantly contributing to model variance (Hall et al 2005). This could also include the use of sub-grid models (Neal et al 2012) and the use of bridge representation (Sampson et al 2012) as well as including channel width and friction as variable parameters. Inclusion of these factors may introduce a set of parameters with significant control on model results. It will also provide a useful point to determining the level of representation required where rivers and channel networks form part of the inflow boundary conditions.

The inflow boundary has been treated with a simple range of error determined from the literature (di Baldassare et al 2009). This could be further explored, by identifying the structure of the uncertainty in the extrapolation methods of depth rating curves, and using that to refine the analysis of uncertainty associated with this factor.

The use of distributed friction surfaces has proved to have a number of advantages over uniform values, not only in terms of model evaluation function but also the representation of flow in critical low friction surfaces and regions such as road networks. In this research the parameter space was based on thirteen friction surfaces, with two surface regions. Consequently regional values were fixed relatively so that the same Manning's values were used for each parameter set. Given that there is a clear reduction in the range of model function, further work could be focused on determining the influence of the relationship between the value for the low friction surface and the high friction surface to output, by comparing all values of each surface to all values of the other surface. This would have increased

the total number of model runs in this research by an order of magnitude, but would provide further insight into the précised nature of Manning's friction and how it should be implemented in inundation models.

In order to provide a comprehensive analysis of the entire flood inundation modelling process, the uncertainty related to the observed data should also be included. Here, due to the increased computational cost it has not been included, but could be included in both the Mexborough test case, and more particularly the Coventry test case, where the uncertainty in the digitisation process could be included, or assumptions concerning the precise nature of the extent limit could be tested. Previous work has introduced the idea of fuzzy extent methodology (Pappenberger et al 2006), or a various approaches to determining observed depth from either satellite images or recorded depth point (Mason et al 2009, Stephens et al 2012). This work would then provide further perspective on how to further reduce uncertainty in modelling, and to focus further work on increasing the knowledge of the modelling building process or to refine the data capturing process.

The use of the Morris Method and BACCO approach, were used to reduce the computational costs of a full variance based quantitative sensitivity analysis method, such as Sobol Indicies. These approaches make a number of assumptions about the input distributions and in the case of the Morris method, can only provide a qualitative level of analysis. For both Coventry and Glasgow, this level of analysis was sufficient to indicate the importance of the level of physical representation compared to other factors. But where parameters were similar in significance such as in the Mexborough test case, the Morris method proved ineffective at differentiating between the factors. The BACCO approach also used a smaller parameter space, excluding both building representation and level of physical representation allowing only a comparison between the ACC and Roe solver could be made in this research. This approach has allowed some further understanding of the how the each code is individually affected by factors. In the Mexborough results, the similarity between the results indicated that the level of physical representation was less critical than in the previous cases. The variation between the significance of factors and approaches is indicative of the issues surrounding the application of sensitivity analysis methods in that each method provides a slight variation in the description of sensitivity. For the BACCO method, this is related to the contribution of a factor to the total model variance, and for the Morris method it is related to the average effect on model output. The implication is that these different definitions will provide variations indentifying key parameters. These conclusions have been made before (Pappenberger et al 2007, Tang et al 2007), and indicate that to further

understand the role of physical representation qualitative methods are required. The role of sensitivity analysis in this research has been to further the findings of the initial analysis, but has also shown evidence of contradictory findings to the systematic analysis approach, such as increased sensitivity to cell size in the Coventry test case and a reduced sensitivity to building representation in the Glasgow test case. A robust sensitivity analysis approach (rather than a split analysis approach used here) which used these test cases as a basis would overcome these issue. Using the method of Sobol or other moment based analysis techniques that evaluate parameter contribution to model output, would provide a more detailed approach, with quantitative results). It has been suggested that a method to including physical representation would be to include a module selection method based on a random switching variable between cells (where the level of physical representation used would be determined by the value of a random variable, as suggested in Hall et al 2005, Saltelli et al 2000) However, this would lead to a parameter space several orders of magnitude greater than the other parameters combined. Utilising the regional analysis approach suggested here to determine model performance in vulnerable areas, the switching method could be employed over model domain regions, thereby reducing the computational cost of the switching method and providing insight into regional response to model variation.

List of References

APEL, H., ARONICA, G. T., KREIBICH, H. & THIEKEN, A. H. 2009. Flood risk analyses-how detailed do we need to be? *Natural Hazards,* 49**,** 79-98.

ARICO, C., SINAGRA, M., BEGNUDELLI, L. & TUCCIARELLI, T. 2011. MAST-2D diffusive model for flood prediction on domains with triangular Delaunay unstructured meshes. *Advances in Water Resources,* 34**,** 1427-1449.

ARONICA, G., BATES, P. D. & HORRITT, M. S. 2002. Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. *Hydrological Processes,* 16**,** 2001-2016.

ARONICA, G., HANKIN, B. & BEVEN, K. 1998. *Topographic sensitivity and parameter uncertainty in the predictions of a 2D inundation model,* Leiden, A a Balkema Publishers.

ARONICA, G. T., FRANZA, F., BATES, P. D. & NEAL, J. C. 2012. Probabilistic evaluation of flood hazard in urban areas using Monte Carlo simulation. *Hydrological Processes,* 26**,** 3962-3972.

BATES, P. D. & DE ROO, A. P. J. 2000. A simple raster-based model for flood inundation simulation. *Journal of Hydrology,* 236**,** 54-77.

BATES, P. D., HORRITT, M. S., ARONICA, G. & BEVEN, K. 2004. Bayesian updating of flood inundation likelihoods conditioned on flood extent data. *Hydrological Processes,* 18**,** 3347-3370.

BATES, P. D., HORRITT, M. S. & FEWTRELL, T. J. 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology,* 387**,** 33-45.

BEGNUDELLI, L. & SANDERS, B. F. 2006. Unstructured grid finite-volume algorithm for shallow-water flow and scalar transport with wetting and drying. *Journal of Hydraulic Engineering-Asce,* 132**,** 371-384.

BEVEN, K. 2007. Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process. *Hydrology and Earth System Sciences,* 11**,** 460-467.

BEVEN, K. & BINLEY, A. 1992. THE FUTURE OF DISTRIBUTED MODELS - MODEL CALIBRATION AND UNCERTAINTY PREDICTION. *Hydrological Processes,* 6**,** 279-298.

BORGONOVO, E. 2006. Measuring uncertainty importance: Investigation and comparison of alternative approaches. *Risk Analysis,* 26**,** 1349-1361.

BRANDIMARTE, L. & WOLDEYES, M. K. 2013. Uncertainty in the estimation of backwater effects at bridge crossings. *Hydrological Processes,* 27**,** 1292-1300.

CHADWICK, A. M., J; BORTHWICK, M 2004. *Hydraulics in Civil and Environmental Engineering*.

CHOW, V. E., MAIDMENT, D.R., MAYS, L.M. 1988. *Applied Hydrology*, McGraw Hill Book Company.

CLOKE, H. L. & PAPPENBERGER, F. 2009. Ensemble flood forecasting: A review. *Journal of Hydrology,* 375**,** 613-626.

COBBY, D. M., MASON, D. C. & DAVENPORT, I. J. 2001. Image processing of airborne scanning laser altimetry data for improved river flood modelling. *Isprs Journal of Photogrammetry and Remote Sensing,* 56**,** 121-138.

CUNGE, J. A., HOLLY, F. M. & VERWY, A. 1976. Pracitcal Aspects of computational river hydraulics. *In:* PITMAN (ed.). London.

DE ALMEIDA, G. A. M., BATES, P., FREER, J. E. & SOUVIGNET, M. 2012. Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling. *Water Resources Research,* 48.

DEFRA/ENIVRONMENT AGENCY (2006) R&D Outputs:Flood Risks to People FD2321/TR2 Guidance Document

DI BALDASSARRE, G., LAIO, F. & MONTANARI, A. 2009. Design flood estimation using model selection criteria. *Physics and Chemistry of the Earth,* 34**,** 606-611.

DOMENEGHETTI, A., CASTELLARIN, A. & BRATH, A. 2012. Assessing rating-curve uncertainty and its effects on hydraulic model calibration. *Hydrology and Earth System Sciences,* 16**,** 1191-1202.

DUN, R. W., WICKS, J.M., 2013. Canal breach risk assessment for improved asset management *ICE Water Management* 167,1, 5-16

DUNG, N. V., MERZ, B., BARDOSSY, A., THANG, T. D. & APEL, H. 2011. Multi-objective automatic calibration of hydrodynamic models utilizing inundation maps and gauge data. *Hydrology and Earth System Sciences,* 15**,** 1339-1354.

ENVIRONMENT AGENCY SCIENCE REPORT (2009) SC080035/SR, Desktop review of 2D hydraulic modelling packages

FEWTRELL, T. J., BATES, P. D., HORRITT, M. & HUNTER, N. M. 2008. Evaluating the effect of scale in flood inundation modelling in urban environments. *Hydrological Processes,* 22**,** 5107-5118.

FEWTRELL, T. J., DUNCAN, A., SAMPSON, C. C., NEAL, J. C. & BATES, P. D. 2011. Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data. *Physics and Chemistry of the Earth,* 36**,** 281-291.

GUPTA, H. V., CLARK, M. P., VRUGT, J. A., ABRAMOWITZ, G. & YE, M. 2012. Towards a comprehensive assessment of model structural adequacy. *Water Resources Research,* 48.

HALL, J. W., BOYCE, S. A., WANG, Y. L., DAWSON, R. J., TARANTOLA, S. & SALTELLI, A. 2009. Sensitivity Analysis for Hydraulic Models. *Journal of Hydraulic Engineering-Asce,* 135**,** 959-969.

HALL, J. W., MANNING, L. J. & HANKIN, R. K. S. 2011. Bayesian calibration of a flood inundation model using spatial data. *Water Resources Research,* 47.

HALL, J. W., TARANTOLA, S., BATES, P. D. & HORRITT, M. S. 2005. Distributed sensitivity analysis of flood inundation model calibration. *Journal of Hydraulic Engineering-Asce,* 131**,** 117-126.

HARTMANN, A., WAGENER, T., RIMMER, A., LANGE, J., BRIELMANN, H. & WEILER, M. 2013. Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resources Research,* 49**,** 3345-3358.

HERMAN, J. D., KOLLAT, J. B., REED, P. M. & WAGENER, T. 2013. Technical Note: Method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models. *Hydrology and Earth System Sciences,* 17**,** 2893-2903.

HORRITT, M. S. 2002. Evaluating wetting and drying algorithms for finite element models of shallow water flow. *International Journal for Numerical Methods in Engineering,* 55**,** 835-851.

HORRITT, M. S. 2006. A methodology for the validation of uncertain flood inundation models. *Journal of Hydrology,* 326**,** 153-165.

HORRITT, M. S. & BATES, P. D. 2001. Predicting floodplain inundation: raster-based modelling versus the finite-element approach. *Hydrological Processes,* 15**,** 825-842.

HORRITT, M. S. & BATES, P. D. 2002. Evaluation of 1D and 2D numerical models for predicting river flood inundation. *Journal of Hydrology,* 268**,** 87-99.

HORRITT, M. S., DI BALDASSARRE, G., BATES, P. D. & BRATH, A. 2007. Comparing the performance of a 2-D finite element and a 2-D finite volume model of floodplain inundation using airborne SAR imagery. *Hydrological Processes,* 21**,** 2745-2759.

HUNTER, N. M., BATES, P. D., HORRITT, M. S. & WILSON, M. D. 2007. Simple spatially-distributed models for predicting flood inundation: A review. *Geomorphology,* 90**,** 208-225.

HUNTER, N. M., BATES, P. D., NEELZ, S., PENDER, G., VILLANUEVA, I., WRIGHT, N. G., LIANG, D., FALCONER, R. A., LIN, B., WALLER, S., CROSSLEY, A. J. & MASON, D. C. 2008. Benchmarking 2D hydraulic models for urban flooding. *Proceedings of the Institution of Civil Engineers-Water Management,* 161**,** 13-30.

HUNTER, N. M., HORRITT, M. S., BATES, P. D., WILSON, M. D. & WERNER, M. G. F. 2005. An adaptive time step solution for raster-based storage cell modelling of floodplain inundation. *Advances in Water Resources,* 28**,** 975-991.

JUNG, H. C., JASINSKI, M., KIM, J. W., SHUM, C. K., BATES, P., NEAL, J., LEE, H. & ALSDORF, D. 2012. Calibration of two-dimensional floodplain modeling in the central Atchafalaya Basin Floodway System using SAR interferometry. *Water Resources Research,* 48.

KENNEDY,M. & O'HAGAN,A. (2001) Bayesian Calibration of computer models *Journal of the Royal Society,Series B* 63,425-464

KRZYSZTOFOWICZ, R. & KELLY,S.(2000) Hydrological uncertainty processor for probabilistic river stage forecasting *Water Resources Research* 36, 11

LEVEQUE, R. J. 2002. *Finite Volume Methods for Hyperbolic Problems*, Cambridge Text in Applied Mathematics.

LIANG, Q. H. & BORTHWICK, A. G. L. 2008. Adaptive quadtree simulation of shallow flows with wet-dry fronts over complex topography. *Computers & Fluids,* 38**,** 221-234.

MANTOVAN, P. & TODINI, E. 2006. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of Hydrology,* 330**,** 368-381.

MARKS, K. & BATES, P. 2000. Integration of high-resolution topographic data with floodplain flow models. *Hydrological Processes,* 14**,** 2109-2122.

MASON, D. C., HORRITT, M. S., DALL'AMICO, J. T., SCOTT, T. R. & BATES, P. D. 2007. Improving river flood extent delineation from synthetic aperture radar using airborne laser altimetry. *Ieee Transactions on Geoscience and Remote Sensing,* 45**,** 3932-3943.

MASON, D. C., SPECK, R., DEVEREUX, B., SCHUMANN, G. J. P., NEAL, J. C. & BATES, P. D. 2010. Flood Detection in Urban Areas Using TerraSAR-X. *Ieee Transactions on Geoscience and Remote Sensing,* 48**,** 882-894.

MERZ, B., KREIBICH, H., SCHWARZE, R. & THIEKEN, A. 2010. Review article 'Assessment of economic flood damage'. *Natural Hazards and Earth System Sciences,* 10**,** 1697-1724.

NEAL, J C., BATES, P.D., FEWTRELL, T.J,. HUNTER, N.M., WILSON, M.D., HORRITT, M.S., (2009) Distributed whole city water level measurements from the Carlisle 2005 urban flood evenet and comparison with hydraulic model simulations *Jounral of Hydrology* 368 42-55.

NEAL, J., VILLANUEVA, I., WRIGHT, N., WILLIS, T., FEWTRELL, T. & BATES, P. 2011. How much physical complexity is needed to model flood inundation? *Hydrological Processes***,** n/a-n/a.

NEAL, J., SCHUMANN, G., BATES, P., (2012) A subgrid channel model for sumulating river hydraulics and floodplain inundation over large and data sparse areas *Water Resources Research* 48, 11.

NOTT, D. J., MARSHALL, L. & BROWN, J. 2012. Generalized likelihood uncertainty estimation (GLUE) and approximate Bayesian computation: What's the connection? *Water Resources Research,* 48.

OAKLEY,J & O'HAGEN,A (2002) Bayesian inference for uncertainty distribtuion of computer model outputs *Biometrika* 89 769-784.

PAPPENBERGER, F., BEVEN, K., FRODSHAM, K., ROMANOWICZ, R. & MATGEN, P. 2007a. Grasping the unavoidable subjectivity in calibration of flood inundation models: A vulnerability weighted approach. *Journal of Hydrology,* 333**,** 275-287.

PAPPENBERGER, F., BEVEN, K., HORRITT, M. & BLAZKOVA, S. 2005. Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. *Journal of Hydrology,* 302**,** 46-69.

PAPPENBERGER, F., BEVEN, K. J., RATTO, M. & MATGEN, P. 2008. Multi-method global sensitivity analysis of flood inundation models. *Advances in Water Resources,* 31**,** 1-14.

PAPPENBERGER, F., FRODSHAM, K., BEVEN, K., ROMANOWICZ, R. & MATGEN, P. 2007b. Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. *Hydrology and Earth System Sciences,* 11**,** 739-752.

PAPPENBERGER, F., MATGEN, P., BEVEN, K. J., HENRY, J. B., PFISTER, L. & FRAIPONT DE, P. 2006. Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advances in Water Resources,* 29**,** 1430-1449.

PAPPENBERGER, F., STEPHENS, E., THIELEN, J., SALAMON, P., DEMERITT, D., VANANDEL, S. J., WETTERHALL, F. & ALFIERI, L. 2013. Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication. *Hydrological Processes,* 27**,** 132-146.

PENNING-ROWSELL, E., PRIEST ,S. PARKER, D. MORRIS, J., TUNSTALL, S., VIAVATTENE, C., CHATTERTON, J., OWEN, D., (2010) *Flood and Coastal Erosion Risk Management: A Manual for Economic Appraisal* Routledge

PRESTININZI, P., DI BALDASSARRE, G., SCHUMANN, G. & BATES, P. D. 2011. Selecting the appropriate hydraulic model structure using low-resolution satellite imagery. *Advances in Water Resources,* 34**,** 38-46.

ROE, P. L. 1981. APPROXIMATE RIEMANN SOLVERS, PARAMETER VECTORS, AND DIFFERENCE-SCHEMES. *Journal of Computational Physics,* 43**,** 357-372.

ROMANOWICZ, R. & BEVEN, K. 2003. Estimation of flood inundation probabilities as conditioned on event inundation maps. *Water Resources Research,* 39.

SALTELLI, A., CHAN, K., SCOTT,E.M., eds. 2000 Sesnitivity Analysis *Wiley Publishers*

SAMPSON, C. C., FEWTRELL, T. J., DUNCAN, A., SHAAD, K., HORRITT, M. S. & BATES, P. D. 2012. Use of terrestrial laser scanning data to drive decimetric resolution urban inundation models. *Advances in Water Resources,* 41**,** 1-17.

SANDERS, B. F., SCHUBERT, J. E. & GALLEGOS, H. A. 2008. Integral formulation of shallow-water equations with anisotropic porosity for urban flood modeling. *Journal of Hydrology,* 362**,** 19-38.

SCHUBERT, J. E. & SANDERS, B. F. 2012. Building treatments for urban flood inundation models and implications for predictive skill and modeling efficiency. *Advances in Water Resources,* 41**,** 49-64.

SCHUBERT, J. E., SANDERS, B. F., SMITH, M. J. & WRIGHT, N. G. 2008. Unstructured mesh generation and landcover-based resistance for hydrodynamic modeling of urban flooding. *Advances in Water Resources,* 31**,** 1603-1621.

SIMOES, F. J. M. 2011. Finite Volume Model for Two-Dimensional Shallow Environmental Flow. *Journal of Hydraulic Engineering-Asce,* 137**,** 173-182.

SLEIGH, P. A., GASKELL, P. H., BERZINS, M. & WRIGHT, N. G. 1998. An unstructured finite-volume algorithm for predicting flow in rivers and estuaries. *Computers & Fluids,* 27**,** 479-508.

TANG, Y., REED, P., WAGENER, T. & VAN WERKHOVEN, K. 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrology and Earth System Sciences,* 11**,** 793-817.

TORO, E. F. 2001. *Shock Capturing Methods for Free-Surface Shallow Flows*, Wiley and Sons Ltd.

TRIGG, M. A., WILSON, M. D., BATES, P. D., HORRITT, M. S., ALSDORF, D. E., FORSBERG, B. R. & VEGA, M. C. 2009. Amazon flood wave hydraulics. *Journal of Hydrology,* 374**,** 92-105.

TSUBAKI, R. & KAWAHARA, Y. 2013. The uncertainty of local flow parameters during inundation flow over complex topographies with elevation errors. *Journal of Hydrology,* 486**,** 71-87.

VILLANUEVA, I. & WRIGHT, N. G. 2006. Linking Riemann and storage cell models for flood prediction. *Proceedings of the Institution of Civil Engineers-Water Management,* 159**,** 27-33.

WANG, J. P. & LIANG, Q. 2011. Testing a new adaptive grid-based shallow flow model for different types of flood simulations. *Journal of Flood Risk Management,* 4**,** 96-103.

WARMINK, J. J., VAN DER KLIS, H., BOOIJ, M. J. & HULSCHER, S. 2011. Identification and Quantification of Uncertainties in a Hydrodynamic River Model Using Expert Opinions. *Water Resources Management,* 2**5,** 601-622.

WILSON, M., BATES, P., ALSDORF, D., FORSBERG, B., HORRITT, M., MELACK, J., FRAPPART, F. & FAMIGLIETTI, J. 2007. Modeling large-scale inundation of Amazonian seasonally flooded wetlands. *Geophysical Research Letters,* 34**,** -.

ZHAN, C.-S., SONG, X.-M., XIA, J. & TONG, C. 2013. An efficient integrated approach for global sensitivity analysis of hydrological model parameters. *Environmental Modelling & Software,* 41**,** 39-52.

List of Notation

$\boldsymbol{F}$	x axis flux vector
$\boldsymbol{G}$	y axis flux vector
$\boldsymbol{U}$	vector of conserved variables
$\boldsymbol{S}$	vector of source terms
x	x spatial node
y	y spatial node
A	channel cross section area
R	channel hydraulic radius
P	channel perimeter
Q	total discharge
S_o	bed slope source
S_f	friction slope source
q	unit discharge
h	flow depth
u	x axis velocity component
v	y axis velocity component
n	Manning's friction coefficient
g	acceleration due to gravity
z	bed elevation
$\Delta x, \Delta y$	Cell indexes
i, j	Spatial node indexes
Δt	Timestep
D_i	Parameter Variance
Θ	Theta – numerical diffusion coefficient for LISFLOOD-ACC
c	wave celerity
C_{cfl}	CFL condition
S^+	Wave Speed
S_i	First order sensitivity factor
S_{Ti}	Total sensitivity factor
X	Parameter set
Y	Model output in sensitivity analysis
k	number of parameters in sensitivity analysis test

n	number of iterations in test
p_y	Conditional Probability distribution of output, dependent on the parameters
P_y	Unconditional probability distribution of output
PX	Probability distribution of parameter set x
$L(v)$	Likelihood value
S	similarity factor
d_i	Parameter elementary effect
df	debris factor
DR	Damage ratio
μ	Mean elementary effect
σ	standard deviation of the elementary effect

List of Abbreviations

BACCO – Bayesian Analysis of Computer Code Output

BB -Building Block

BP- Building Porosity model

BR- Building resistance model

BH- Building hole

CFL – Courant-Freidrichs-Lewy number

GEM – Gaussian Estimation Model

GSA- Global Sensitivity Analysis

DTM – Digital Terrain Model

F^2 - Goodness of fit measure

GSA- Global Sensitivity Analysis

LIDAR – Light Detecting Aperture Radar

MC- Monte Carlo Modelling

MCM – Multi Coloured Manual, a handbook for determining depth damage curves

NS- Nash Sutcliffe

RMSE- Root Mean Square Error

SA- Sensitivity Analysis

SAR- Synthetic Aperture Radar

SWE – Shallow Water Equations