Natural Hazards Partnership Surface Water Flooding Hazard Impact Model: Phase 2 Final Report

Natural Hazards PARTNERSHIP

The NHP is a consortium of 17 public bodies (mainly government departments and agencies, trading funds and public sector research establishments) which aims to build on partners' existing natural hazard science, expertise and services to deliver fully coordinated impact-based natural hazard advice provided for civil contingencies and responder communities and governments across the UK.

naturalhazardspartnership.org.uk

Prepared on behalf of the consortium by:
Centre for Ecology & Hydrology
Wallingford, Oxon, OX10 8BB

Research Contractors:

Centre for Ecology & Hydrology (CEH),
Wallingford, Oxon, OX10 8BB

Health & Safety Laboratory (HSL),
Buxton, Derbyshire, SK17 9JN

JBA Consulting,
Skipton, North Yorkshire, BD23 3AE

Collaborators:

Met Office,
Exeter, Devon, EX1 3PB

Flood Forecasting Centre,
Exeter, Devon, EX1 3PB

King's College London,
London, WC2R 2LS

Author(s):
Cole, S.J., Moore, R.J., Aldridge, T.A., Gunawan, O., Balmforth, H., Hunter, N., Mooney, J., Lee, D., Fenwick, K., Price, D., Demeritt, D.

Dissemination Status:
Publically Available

Keywords:
Forecasting, Hazard, Impact, Natural Hazards Partnership (NHP), Pluvial, Surface Water Flooding

Customers:
Graeme Boyce, Flood Forecasting Centre, Fitzroy Road, Exeter, EX1 3PB

Adam Baylis, Environment Agency (on behalf of the Flood and Coastal Erosion Risk Management Programme), Evidence Directorate, Horizon House, Deanery Road, Bristol, BS1 5AH

Executive summary

Phase 2 of the Natural Hazards Partnership Surface Water Hazard Flooding Impact Model (NHP SWF HIM) project has moved the hazard and impact model from a proof-of-concept to implementation as an integrated system running in real-time and ready for trialling by the Flood Forecasting Centre (FFC). Further developments to improve both the SWF hazard footprint and the modelling of impacts, with a sensitivity analysis of key parameters, have been investigated and incorporated into the trial system where appropriate. The results of further verification and assessment of the SWF HIM methodology are reported, as is the structure of the trial system which delivers web map services visualised via configured displays on Visual Weather, the Met Office forecaster visualisation system. The work has been delivered under a collaborative arrangement between Met Office, Environment Agency, CEH and HSL, steered by the FFC, and supported by JBA and King's College London. The collaborative parties are all members of the NHP.

Surface runoff accumulation estimates derived from the countrywide G2G distributed hydrological model were evaluated using eleven case studies and a longer five-year record with potential improvements identified. In addition, using a high-resolution 2D inundation model, CEH and JBA tested the assumption within the SWF HIM that G2G surface runoff estimates can reasonably be equated to the "effective rainfall" estimates used to derive the updated Flood Map for Surface Water (uFMfSW) dataset.

Refinements were made to the impact modelling methodology by HSL and the impact library code was tuned to run within the trial system. Analysis of key impact parameters was undertaken which identified model sensitivities in the summary methods. The analysis did not indicate a pressing need to change the impact severity thresholds derived in Phase1. Property and transport were found to be the most influential impact criteria. Due to limitations in the processed datasets available under Phase 1, the SWF HIM only utilised the three uFMfSW Maximum Output maps for 30, 100, 1000 year return period and an assumed critical storm duration (CSD) of 3 hours. As a known limitation, further analysis has supported the recommendation to incorporate the nine (1, 3, 6 hours CSD and 30, 100, 1000 year return period) uFMfSW scenarios into the SWF HIM when updated datasets are available.

Further verification of the SWF HIM methodology was completed and the findings presented. For those events where there was confidence in the accuracy of the observed rainfall, the SWF HIM demonstrated promising performance. For a number of the case studies the localised pattern of rainfall was not effectively detected by the raingauge network and the SWF HIM was not able to model the subsequent impacts. King's College London, working with the FFC, ensured the robustness of the observed impacts dataset used in verification.

The Met Office led the development of the SWF HIM trial system, designed to be compatible with the NHP's emerging Hazard Impact Framework (HIF) specification, and to communicate via web services conforming to Open Geospatial Consortium (OGC) standards. Following testing, the trial system was available to the FFC to commence end-to-end trials from June 2016 (not part of Phase 2).

Achievements, conclusions and recommendations have been recorded and have informed a recommended set of next steps including trialling and further development prior to full operational implementation.

Acknowledgements

The Project Team would like to thank the following for their contribution to the work reported here. Graeme Boyce (Flood Forecasting Centre) for chairing the Project Board and providing strategic guidance; Stefan Laeger and Adam Baylis (Environment Agency) for providing the Environment Agency perspective to the project and access to the updated Flood Map for Surface Water (uFMfSW) dataset; Flood Forecasting Centre representatives (Keith Fenwick, David Price, Andy Lane) for their inputs and ensuring end-user requirements were met; Gavin Evans, Bruce Wright and Dan Grey from the Met Office for providing historical and live ensemble rainfall forecasts; Maria Escobar-Tello from King's College London (KCL) for providing the impact validation data; Steven Wells and Gianni Vesuviano (CEH) for undertaking runs of the G2G model and supporting analyses.

In addition to funding from the joint Department for Environment, Food and Rural Affairs and Environment Agency Flood and Coastal Erosion Risk Management programme and Met Office, the project also benefited from GoScience funding which is gratefully acknowledged. CEH, HSL and Met Office have also contributed significant in-kind support. King's College London wish to acknowledge funding support from the Natural Environment Research Council (NERC) Susceptibility of catchments to INTense RAinfall and flooding (SINATRA) project.

Contents

1 Introduction

Flooding from surface water continues to be a hazard for potentially up to 4 million properties in the UK with its impact also felt across vulnerable stretches of national and local rail and road infrastructure. With around 36 thousand properties flooded from surface water during the summer of 2007, the Flood Forecasting Centre (FFC) was set-up by the Met Office and the Environment Agency with a remit to forecast flood risk for all natural sources of flooding, including that from surface water. Further instances of surface water flooding (SWF) across England and Wales since 2007 has not diminished the risk presented from this source. The challenge for the FFC has been to provide real-time national SWF guidance for a phenomenon that (i) has complex modelling requirements to be forecast well in an urban environment and (ii) is generally driven by convective rainfall events that are difficult to predict at a spatial scale helpful to local responders. The Natural Hazards Partnership Surface Water Flooding Hazard Impact Model (NHP SWF HIM) project has brought together a number of public sector parties with experience in forecasting flooding and modelling the impacts of hazards to develop a technical capability that will support the FFC in improving its real-time national SWF guidance.

Phase 2 of the NHP SWF HIM project has moved the SWF HIM from Proof-of-Concept under Phase 1 to a real-time end-to-end trial system. Phase 1 provided experience of working collaboratively and developed a rational methodology, built around best available datasets, supported by a limited set of three case studies in two regions of England (the Northeast and Southwest). The Proof-of-Concept SWF HIM approach uses real-time ‘dynamic’ SWF Hazard Footprints based on G2G surface runoff forecasts at a 1km scale combined with a ‘static’ Impact Library based on the updated Flood Map for Surface Water (uFMfSW) dataset for three return periods and supporting impact datasets. This approach, supported by Peer Review, was seen as a pragmatic step forward.

The Phase 2 work reported here concerned further “Development and testing” of the SWF HIM. It comprised a number of work packages that included further investigations aimed at improving both the SWF hazard footprint and the link with impacts, along with sensitivity analyses on key parameters within the impact methodology. Further evidence has been collated to support the SWF HIM capability. A real-time end-to-end trial system has been implemented within Met Office systems with the primary visualisation provided through Visual Weather. An operational trial (Phase 3) by the FFC with support from the project team has been detailed and initial discussions concerning full operational implementation (Phase 4) have commenced.

This report provides a brief overview of the Phase 2 outcomes and references more detailed project notes and reports.

2 Planning for Phase 2

A key element of the Phase 2 work was to extend the case study evidence-base to include more events used for testing and validating the SWF HIM. Lead by the FFC, Phase 2 first selected 11 case studies from the period 2012 to 2014 following a series of self-imposed protocols. These protocols were designed to ensure that the selected case studies produced a representative range of potential SWF scenarios spanning impact severity intensities, land cover types, seasons of the year and relevant spatial domains (Flood Forecasting Centre, 2015a). The actual impact severity for each event and county was also assessed by the FFC.

Ensemble rainfall forecasts from the Met Office were required for the case studies. Under Phase 1, 12-member blended ensembles were generated from deterministic UK4 Numerical Weather Prediction (NWP) output using STEPS-2 methodology. These forecasts were used for some of the June 2012 Phase 2 case study work. During Phase 2 the Met Office Gridded Post Processing (GPP) Team produced 24-member blended ensembles from 12-member MOGREPS-UK (Met Office Global and Regional Ensemble Prediction System) data (Pierce, 2015; Met Office, 2016) for all case studies as being representative of the latest operational ensemble generation process as at 2015.

3 Further development of the SWF Hazard Footprint

The Proof-of-Concept SWF HIM developed under Phase 1 used dynamic 1km accumulations of Grid-to-Grid (G2G) surface runoff estimates as the basis for the approach (Cole *et al*., 2015). Further development of the SWF Hazard Footprint under Phase 2 focussed on evaluating and improving the methodology along with learning from the experience of running a similar operational SWF pilot system in support of the 2014 Glasgow Commonwealth Games. In addition, a range of G2G outputs have been supplied to support other activities under Phase 2 concerned with end-to-end evaluation, development and sensitivity testing of the SWF HIM (Sections 4 and 5) and testing of the operational trial system (Section 6).

A key assumption within the SWF HIM is that the G2G surface runoff accumulation estimates can be reasonably equated to the uFMfSW “effective rainfall” estimates in real-time. This has been tested in a joint CEH and JBA case study investigation: details are reported in Warren *et al.* (2016) and summarised here in Section 3.1.

Learning from the Commonwealth Games pilot system for a 10 by 10 km domain over Glasgow (Centre for Ecology & Hydrology, 2015) highlighted the key similarities and differences with the NHP SWF HIM approach and provided valuable input into the NHP SWF HIM visualisation specification (Section 6). Both approaches use similar detailed inundation maps for the off-line impact assessments. However, the Glasgow Pilot employed more return periods (seven compared to three for the NHP SWF HIM) but used a simpler receptor attribution approach. A major strength of the Glasgow Pilot is that it ran as a real-time trial during a high-profile event. This ensured wide stakeholder and responder community engagement and meant that outputs were really used in earnest within the operational procedures and timelines of the Glasgow 2014 Multi Agency Control Centre. Although rigorous verification of the Glasgow Pilot’s probabilistic impact outputs were not possible, the general feedback from users is that the system performed well and was useful. SEPA are still operating the trial system two years on and gathering more evidence of performance.

Phase 1 had included a limited investigation of the surface runoff estimates of G2G focussed on two regions of England. Under Phase 2, a more thorough analysis was undertaken (Cole *et al.*, 2016) using 11 cases studies (Section 5) along with an analysis over England & Wales for a longer five-year period encompassing all seasons and soil types. This more extensive analysis has allowed assessment of the frequency of runoff accumulation threshold exceedences. In addition, specific G2G process representation and resolution improvements have been investigated including incorporation of infiltration-excess runoff generation and trial of a 250m resolution configuration of G2G over England & Wales. A trial of the infiltration-excess formulation for the 28 June 2012 case study has revealed intuitively sensible changes in surface runoff volumes. In both investigations it has been challenging to evaluate the potential benefits as there are no actual observations of surface runoff. There has been limited comparison with river flow observations which act as space-time integrated estimates of surface runoff.

It is recommended that both the process representation and resolution estimates are assessed in the wider context of national river flow observations with a focus on smaller, urbanised catchments (for example, through national G2G recalibration trials). Operational upgrades to the 1km G2G model can be accommodated in future G2G releases. The move to a 250m resolution model would require more strategic planning by FFC. However, outputs could still be aggregated to a 1km level to work within the current SWF HIM framework. It is worth noting that the 250m resolution G2G configuration does increase run-times significantly when run as a single process. Opportunities for optimisation and parallelisation would need to be explored to meet operational run-time requirements.

3.1 G2G linked with JFlow+ case study

This case study, undertaken by CEH and JBA, aimed to increase understanding of the SWF hazard footprint assumptions within the SWF HIM: details are reported in Warren *et al*. (2016). Here, instead of selecting the most appropriate footprint from pre-computed flood mapping as done in the SWF HIM, G2G surface runoff outputs (15 minute, 1km resolution) produced using observed and forecast rainfall as input served as boundary conditions to high-resolution 2D hydraulic models (JFlow). Two case study flood events - Newcastle upon Tyne (28 June 2012) and Canvey Island (20 July 2014) - were used. The hydraulic models produced flood outlines associated with these events for the G2G outputs employing either observed (simulation) or forecast rainfall (ensemble).

This work therefore sought to compare the hazard footprints produced using the current SWF HIM "look up" approach with those produced from pseudo real-time hydraulic modelling. The aim being to confirm (or not) the suitability of the current approach for deriving SWF hazard footprints without recourse to event-specific flood modelling. The simulation flood outlines were compared with observed flood data, including observed depths, impacted areas and photographs, to validate the linked G2G-JFlow model outputs. The model was found to perform well for Newcastle but less so for Canvey Island due to the effect of a faulty raingauge on rainfall inputs.

The simulation flood outlines were also compared with three Maximum Output scenarios ("worst case" scenarios for return periods of 30, 100 and 1,000 years), and the nine scenarios (durations of 1, 3 and 6 hours for each of the return periods of 30, 100 and 1,000 years) from the uFMfSW to determine which scenario matched the simulation outline most closely in each 1km grid square over the two case study areas. The three Maximum Output scenarios are what are used in the Phase 1 SWF HIM whilst the nine scenarios are a subject of the SWF HIM sensitivity analysis reported in Section 4.2.

The patterns of the best fit scenarios were logical given the distribution of surface runoff with the 1,000-year scenarios (greater flood extents) fitting best in locations with high runoff and the 30-year scenarios (smaller flood extents) fitting best in areas with little runoff. Overall, the best-fit scenarios had high "measure of fit" (MoF) values for Newcastle but fitted less well for Canvey Island. This could be a result of poorly recorded rainfall for Canvey Island which resulted in little flooding in the simulation flood outlines. Fits were generally worst over areas with little flooding as there is increased sensitivity to drainage modelling assumptions during less extreme events. Both case studies produced a better overall fit ("total MoF") when comparing the simulation outlines with the nine scenarios than with the three scenarios.

The final part of the analysis compared the “best fit” scenarios from the simulation comparison with the uFMfSW scenarios for the Newcastle case study with the equivalent identified by CEH using the approach currently implemented within the SWF HIM. CEH compared maximum surface runoffs for 1, 3 and 6 hour durations over their respective whole forecast time-windows with the corresponding 1, 3 or 6 hour uFMfSW effective rainfall grids for the return periods of 30, 100 and 1,000 years, noting the highest threshold crossed. Error matrices were created to compare the grids and an overall accuracy and Kappa statistic were calculated. It was found that this G2G-only method generally showed lower return periods than the linked G2G-JFlow method in many grid squares. This is probably because the G2G-only method requires the threshold to be crossed whilst the linked G2G-JFlow method finds the uFMfSW scenario with the closest match. Some analysis of the sensitivity to the uFMfSW thresholds used was carried out by repeating these comparisons with relaxed thresholds for the G2G-only method. As expected, lowering the thresholds did increase the similarity of the grids produced by the two methods.

4 Further development of the impact assessment and sensitivity analysis

The FFC uses a risk-based forecasting approach based on a Flood Risk Matrix that requires an understanding of the impacts of flooding events. Within Phase 1, HSL led development of a Proof-of-Concept impact assessment approach (Aldridge & Gorce, 2014). This involved creating an off-line Impact Library formed using the uFMfSW dataset combined with receptor datasets from the National Receptor Database (NRD) and the National Population Database (NPD), along with Ordnance Survey transport datasets. Phase 1 of the SWF HIM development highlighted a number of areas for refinement including application of more sophisticated methodologies and use of updated receptor datasets.

Gunawan & Aldridge (2016a) details the refinements completed under the Phase 2 development and the creation of an updated Impact Library. It also documents the findings of sensitivity testing undertaken to evaluate the suitability of parameters and thresholds used in the Proof-of-Concept SWF HIM, identifies potential limitations and model sensitivities, and highlights dependencies for the four selected Impact Library criteria of Population, Property, Key Sites and Infrastructure, and Transport. Findings and recommendations from the Phase 2 sensitivity work were used within the case study analyses reported here in Section 5.

Further to this, the scope of the Phase 2 work was extended to include an analysis of the uFMfSW data for modelling impacts, comparing the nine uFMfSW rainfall scenarios (three durations and three return periods) with the three uFMfSW Maximum Output maps (three return periods) used in Phase 1. The uFMfSW assessment also tested the assumptions for the use of depth-based modelling to evaluate impacts to people. This is reported in Gunawan & Aldridge (2016b).

4.1 Sensitivity Analysis

Property impacts are the most numerous, the most stable and the most influential for subsequent processing steps within the SWF HIM. Less numerous impact types such as key sites and infrastructure are more sensitive to changing impact thresholds and less influential in subsequent classifications. Transport impacts are relatively small, but they occupy an alternative spatial distribution and therefore exert a large influence between urban centres. Analysis of further case studies would help to provide evidence for decision-making over areas that may benefit from refinement. When impact thresholds are changed, property and day population layers are the most resistant to variation, while rail, infrastructure and roads layers are the most sensitive.

For some impact types, low severity cell scores are based on small threshold values which means that they become very sensitive to change because small absolute changes are proportionally larger. This is a heightened issue for non-residential properties (where the Minor impact threshold is set at one property) because this category represents the largest count of impacted cells.

The sensitivity of the method for evaluating impact at the county level has identified the existence of hotspot cells that are classified as Severe for the lowest return period (30 years) and, as a consequence, are sensitive to the forecast input. Many of these cells are located in Central London and are likely to be relevant priority sites for emergency response. Hotspot cells are also found to be influential for 13 of the smaller counties, where a single 1 km^2 hotspot cell being impacted will result in the entire county being assigned an impact level of Severe. Current evidence points to trends between rural and urban areas, but is not sufficient to support amendment of thresholds to mitigate the influence of hotspots; further case studies are required to investigate this further. End-users need to understand why such hotspots exist and their potential impact on results and summaries.

Property and transport were found to be the most influential impact criteria. This is encouraging as they often occupy different spatial domains: property impacts typically dominate urban areas, while transport layers are located between these centres. Population layers had a relatively low influence, particularly in relation to Property criteria which occupy the same locations. This has raised questions as to whether this is an accurate reflection of impact criteria prioritisation, or whether thresholds need to be refined to reduce Property impact dominance, or increase Population impact influence. Infrastructure and key sites are the least influential as there are fewer such locations when compared to the other criteria. These suggestions for refinement of the impact assessment methodology were considered against the evidence collected during the case studies reported in Section 5 along with the priorities of the FFC as end-users.

Recommendations

The findings of Gunawan & Aldridge (2016a) suggest that there is no pressing requirement to change the current impact severity threshold values that are based on existing FFC tools and discussion with stakeholders. However, the work has highlighted some particular sensitivities that need to borne in mind by end-users regarding the influence of impact criteria and summarising impact at the county level. The analysis points towards a requirement for further validation of the SWF HIM processes and outputs, particularly at county level for impact and risk measurements. This will assist in ensuring that the SWF HIM is a relevant and robust model and allow calibration against real-world events.

4.2 uFMfSW assessment

The uFMfSW assessment (Gunawan & Aldridge, 2016b) investigated, from an impact modelling perspective: (i) the additional variability that nine uFMfSW scenarios might offer over the three uFMfSW Maximum Output maps, and (ii) the potential benefits of using a hazard-based definition of property flooding to identify people at risk rather than the depth-based approach used in Phase 1 and 2. Two sample OS 50 km tiles (NZ05 (North East), TQ55 (South East)) from the uFMfSW dataset were provided by JBA in support of this investigation.

Main findings

In general, the evidence from the Phase 2 work indicates differences in the flood impact information provided by the three-scenario and nine-scenario versions of the Impact Library, particularly at the level of the 1 km grid cell. Differences are also demonstrated when considering counts of flooded properties.

For flooded properties, counts based on the 1-hr duration flooding in the uFMfSW nine-scenario data are generally the most severe, and the closest match to the counts for the three-scenario data, except for 10% of properties where this is not the case. This suggests that in the impacts component of the three-scenario SWF HIM, the selection of the 1-hr duration flood severity outputs of G2G is appropriate.

For 1 km grid cell differences, the lowest return period (30 years), and the infrastructure and key sites impact criteria present the most divergence between the three-scenario and nine-scenario impact data. More tellingly, the property criteria (which have a wide spatial spread and have been previously shown to be more influential) are also divergent. Transport (road) and population layers are typically more similar between the three-scenario and nine-scenario data. The average magnitude of per-cell impact severity difference is 20% for tile NZ05 and 34% for tile TQ55.

Case study analysis suggests that the nine-scenario Impact Library captures a larger number and a higher severity of impacts across all impact criteria. This may have an impact of slightly increasing county-level impact scores, which may in turn increase the risk ratings.

The comparison of flood depth and flood hazard rating-based property counts shows that the two approaches produce similar counts and high Goodness of Fit scores for the Moderate flood hazard rating used for assessing vulnerable populations. However, the differences are more marked for the Significant flood hazard rating (used for wider populations), with fit values below what is typically considered reasonable. Both of the depth-based thresholds show a tendency for overestimating the number of properties compared to the equivalent hazard rating. The analysis highlights the limitations of using depth to represent hazard rating.

If we are to assume that the direct hazard rating-based method is the most appropriate, then the analysis reported here suggests that the two approaches produce different outputs. However, another approach may be to apply alternative assumptions for the velocity and alternative depth proxy thresholds that provide more confidence at larger scales. This however may require additional processing of data sources to generate additional depth thresholds, with no guarantee of improved functionality.

Recommendations

- The evidence in this report supports uptake of the nine-scenario Impact Library, to capture more of the variability in flooding and its impacts from different flood durations. Support for the nine-scenario SWF HIM also requires evidence from the flood forecasting component, to further test the use of flood severity information (from G2G) accounting for storm duration.

- The evidence in this report provides tentative support for the uptake of flood hazard rating-based thresholds over depth-based thresholds for the modelling of population impacts. However, the added value that the hazard rating approach might provide should be considered in context with the relative weight of the population impacts in the SWF HIM and subsequent advice.
- It should be acknowledged that the method for counting properties affected by levels of hazard rating has not been tested outside this project. It is advisable to seek further feedback on the strength of the hazard rating-based property count model.

5 Verification and evaluation

The verification analysis (Aldridge & Gunawan, 2016) provides an assessment of the SWF HIM using historical post-event data for 11 case studies from the period 2012 to 2014, as discussed in Section 2. The validation work takes advantage of flood impact observation data taken from media sources, sourced and developed by project partners at King's College London (Escobar *et al.*, 2016). This analysis aimed to answer the following four questions:

1. How well were the observed impacts modelled by the SWF HIM?
2. Does the SWF HIM produce any false alarms?
3. If an observed impact is undetected based on location, what was the reason?
4. How do the county summaries compare?

Matches of modelled and observed impacts were identified based on location, impact criteria and impact severity using G2G outputs obtained from observed and forecast rainfall as inputs.

Main findings

The results of the verification analysis are shown to be very sensitive to the quality of the rainfall data. Initially raingauge rainfall data interpolated onto a 1 km grid were used as G2G input to derive "observed" SWF HIM outputs. For many of the case studies, the localised patterns of rainfall that caused the flooding were not effectively detected by the raingauge network, which means that the observed levels of flooding and subsequent impact were not modelled by the SWF HIM. For three of the case studies, radar rainfall data were also used as this data source better identified the location of the localised rainfall.

Where the rainfall data was associated with greater confidence, the SWF HIM demonstrated promising performance. For Case Study 1 (Newcastle, 28 June 2012) the probability of detecting an observed impact (question 1) peaks at 64% (based on 80 detected out of 125 total impacts), with other case studies producing values of 54 and 33%. This indicates a reasonable likelihood of detecting impacts. False alarms (question 2) were generally high with the lowest being recorded for Case Study 1 (65%). The high occurrence of false alarms implies that the SWF HIM may overestimate the impacts. However, the quality of the observed impact data used for verification (which is influenced by under-reporting and incompleteness of information) and the difficulty in representing it spatially will contribute to the number of false alarms.

Where observed impacts were not identified, it was shown that the location and magnitude of the observed rainfall and resulting G2G flood response was the main factor. The potential for the Impact Library to identify impacts equivalent to those observed is high for all case studies, and typically close to 100% for most impact criteria. This suggests that the SWF HIM has the potential to perform well in assessment, provided both rainfall for input and flooding information for verification are optimal.

For the county summaries, Case Study 1 provides a good match for impact severity level. But Case Study 2 & 3 (North East and East Anglia, 5-6 July 2012) is the only other one that produces a county level assessment that isn't Minimal. The use of radar rainfall data provides an additional match for Case Study 10 (South and East, 14 August 2014) of Significant impact over Greater London.

The ensemble forecast data were used to provide county impact summaries, and outputs that can be used to assess the risk, based on likelihood and impact. Matches of the modal impact severity are found with the FFC's post-event impact assessments.

The county maps also highlight the potential for further development of the thresholds used for upscaling impacts to the county level, with smaller counties highlighted as 'hotspots' for some events and larger counties demonstrating a tendency for underestimating the level of impact.

Recommendations

The limitations of the observed radar and/or raingauge rainfall data for case study events with intense localised rainfall prevent strong conclusions and recommendations being made but do highlight areas for further investigation. The following recommendations can be made:

- Testing of the model using alternative, more robust rainfall and flooding information is a priority. Use of radar rainfall data (as demonstrated for case studies 8, 9 and 10) to complement the raingauge information seems a sensible way forward.
- The value of the *p* percentage threshold used to assign county level impact severity is a sensitivity that could benefit from further case study testing. This could be achieved through a calibration employing ensemble forecast information.
- Most of the verification analysis was based on model simulations employing gridded raingauge rainfall as input. Ensemble rainfall forecasts have also been used (and with more emphasis once limitations in the raingauge rainfall were identified) but the results could be explored in more detail. Further analysis will help build the evidence-base and present options for calibration.
- The spatial uncertainty of the radar rainfall data used for model input and flooding data used in verification was demonstrated in Case Study 10 over Greater London. Methods to account for this in further validation analyses may provide benefits and give further confidence in results. This would complement the flexibility implemented in this analysis for handling the uncertainty of impact severity and impact criteria. Neighbourhood (window-based) verification analysis developed in partnership with the flood hazard modellers at CEH could achieve this.

6 End-to-end trial system

The Met Office led development of a real-time end-to-end trial SWF HIM system that operates within Met Office infrastructure. The work is described more fully in a Project Note entitled 'Structure and Configuration of End-to-End Trial System' (Mooney and Lane, 2016). The FFC consulted with the project team to create a Visualisation Specification document (Flood Forecasting Centre, 2015b) that captured the user requirements and informed the subsequent implementation.

The system has been designed to be compatible with the NHP's emerging Hazard Impact Framework (HIF) specification. This framework envisages a future Hazard Impact Production System (HIPS) composed of loosely coupled components (primarily impact models and user interfaces) communicating via Web Services conforming to Open Geospatial Consortium (OGC) standards. Standardising around interfaces gives great latitude to those building HIMs to select appropriate hardware, software and internal data formats. OGC standards are supported by a range of client software applications and programming libraries, giving flexibility in selection/customisation of the user interface. The HIF approach will make it possible to easily add or upgrade any given component of the HIPS without major impacts on the rest.

During the development, User Acceptance Testing of the web services and the front-end visualisation system was undertaken by FFC. Detailed Quality Assurance by Met Office, HSL and CEH was undertaken using historical case study data to ensure the real-time SWF HIM trial system outputs matched the off-line CEH/HSL analysis. Real-time SWF HIM outputs have been available via the Visual Weather desktop software since April 2016 and these are supporting an ongoing operational trial by FFC users.

System overview

Figure 1 gives an overview of the SWF HIM trial system. The system is spread across three hardware platforms: the Met Office supercomputer (HPC: High Performance Computer), a virtual server (which was the focus of most of the development effort in this activity) and a number of desktop PCs. On the HPC, CEH's G2G software models the likely SWF using a 24-member ensemble of rainfall forecasts as input . Data on predicted flooding and rainfall patterns is passed to the server where it is processed further. The key tasks undertaken on the server are:

- Impact scenario selection
- Detailed impact determination
- Area impact summation
- Generation and publication of cartographic and graphical outputs

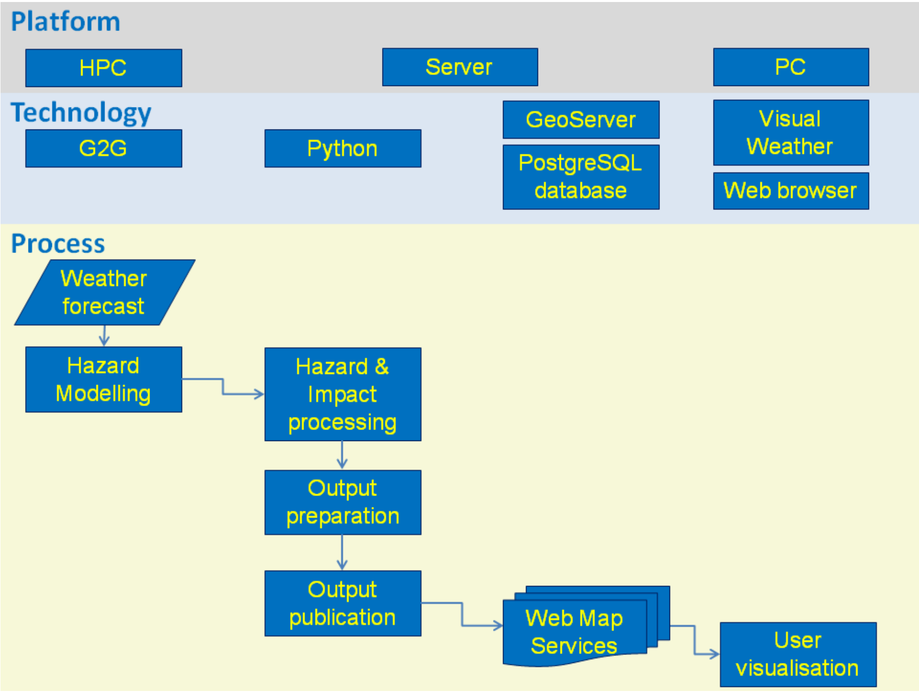

Figure 1 Overview of the SWF HIM end-to-end trial system

There are two parallel implementations of the trial system: one processing very short-range “nowcast” data every hour, the other short-range forecast data every six hours. Each publishes more than 60 web services on each run. The outputs can be broadly categorised as follows:

- Probabilistic mapping of rainfall and surface runoff
- Maximum surface runoff hydrographs
- Maximum surface runoff hydrographs
- Probabilistic mapping of impacts
- Mapping of maximum impact
- County level impact severity and probability tables

The final element of the system is client software on the users’ desktop PCs. For the trial period, FFC have selected Visual Weather as the main client software, taking advantage of its familiarity to users. Due to a minor shortcoming in the way Visual Weather interacts with OGC web services, the hydrographs are viewed via a web page, as are the system’s status messages. Configuration of Visual Weather was led by FFC, with web services organised into page views that fit the Centre’s workflows. The top level summary of 1km and county level flood impacts is displayed in Figure 2.

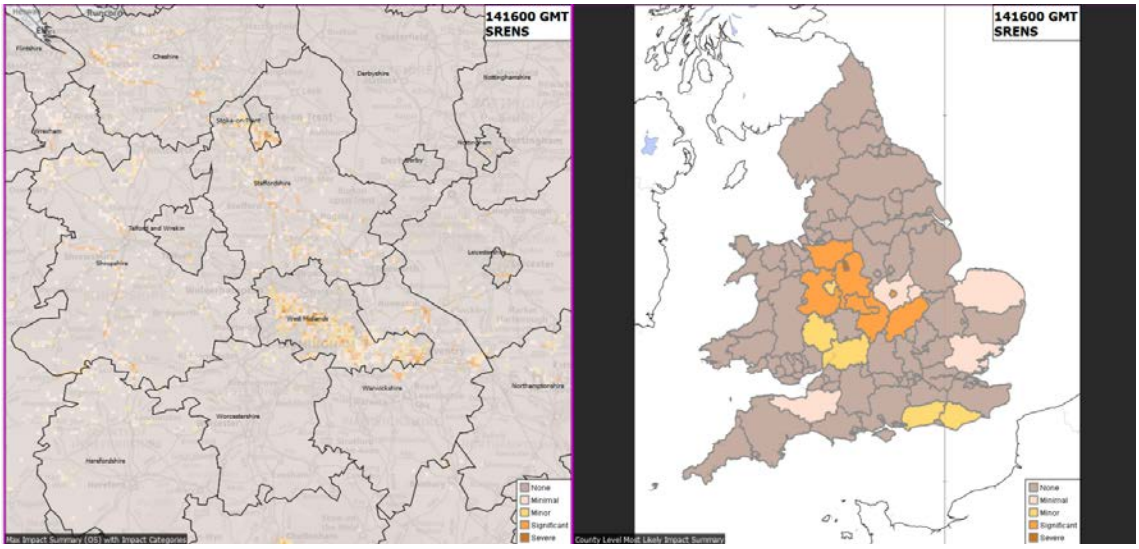

Figure 2 Top level summary of 1km and county level impacts

The decision to incorporate a prototype SWF HIM end-to-end trial system into Phase 2, rather than wait for the work described in previous sections to be completed has been vindicated. The work has proven the practicality of the web service based approach to delivering outputs, has provided FFC with the means to conduct an operational trial and has motivated all of those involved to move beyond the science and the raw outputs to consider the content and design of the displays that will appear before end users.

7 Achievements and Conclusions

The primary achievement from Phase 2 is the collaborative development and implementation of an integrated real-time end-to-end trial SWF HIM system enabling initiation of a pre-operational trial within the FFC. The real-time system has been verified against off-line development systems. Sensitivities and limitations of the SWF HIM have been investigated and recorded to support the FFC trial.

The main conclusions are succinctly captured below.

- Whilst sensitivities and potential limitations of the Impact Library and operational methods were identified, the additional analysis and case study verification undertaken in Phase 2 has not indicated a need to change model parameters and thresholds derived in Phase 1.
- Comparative analysis of SWF hazard footprints, produced using the SWF HIM and a pseudo real-time hydraulic modelling approach, supports continued evaluation of the current methodology without recourse to event-specific flood modelling.
- Further cases study validation has provided additional evidence of model effectiveness but has also signposted areas of uncertainty.
- Development of an integrated real-time end-to-end trial system has proven the practicality of implementing a web service based approach with outputs delivered to OGC standards and supplied via a web map service.
- A comparison between the three-scenario and the nine-scenario uFMfSW approaches for assessing impacts supported the uptake of the latter approach to capture more of the variability in flooding impacts from different storm durations.
- The potential value of media information was demonstrated as a source of impact observations alongside expert assessment for validation purposes.

8 Recommended next steps

During Phase 2, planning for the eventual operational implementation of the SWF HIM has been undertaken, taking into account emerging findings and recommendations listed in this report. It is proposed that Phase 3 should have two parallel strands: (i) an Operational Trial, and (ii) SWF HIM developments. These will be followed by Phase 4 - Operational Implementation.

Phase 3. Operational Trial

The plans for the Operational Trial phase include the following activities.

Live operational system testing
Detailed plans of how the Operational Trial would be undertaken by operational FFC hydrometeorologists has been developed and this commenced in June 2016. This aims to assess the benefits of the SWF HIM relative to the existing SWF Decision Support Tool approach.

Post-event analysis
This will involve HSL, CEH, KCL and FFC and entail offline re-running of the SWF HIM procedure on identified events using observed rainfall data and assessing them against observed impact information derived by KCL. This will improve the evidence-base for the SWF HIM.

Embedding the SWF HIM into FFC operations
This will consider the FFC requirements for a fully operational SWF HIM tool and identify any recommendations for improvements prior to operational implementation. It will include consideration of the Visual Weather interface.

Communication
Training of operational hydrometeorologists within the SWF HIM project was recognised as a priority. The SWF HIM team will be available on a consultancy basis to provide knowledge and information on how the system works. A wider communication plan of the SWF HIM project across the FFC, Met Office and Environment Agency will be developed.

Phase 3. SWF HIM Developments

The precise scope of developments is still to be agreed but consideration of implementing a nine-scenario uFMfSW approach will be considered. The Impact Library also requires extension to Wales following development of a property flooding dataset equivalent to that already in place for England.

Phase 4. Operational Implementation

Moving the real-time end-to-end SWF HIM trial system to be fully within the suite of operational systems at the Met Office is not a trivial activity. During Phase 2, there has been early and positive engagement with the Met Office Tech Gate team aimed at easing the transition. Discussions and planning for operational implementation will continue within Phase 3.

References

Aldridge, T. and Gorce, J.-P. (2014) Surface Water Flooding: Hazard Impact Model Research. Technical Report MSU2014/14 for the Flood Forecasting Centre, Research Contractor: Health & Safety Laboratory, 43pp + appendices.

Aldridge, T. and Gunawan, O. (2016) Surface Water Flooding Hazard Impact Model: Model Validation. Technical Report MSU/2016/17 for the Flood Forecasting Centre, Research Contractor: Health & Safety Laboratory, 91pp.

Centre for Ecology & Hydrology (2015) *D2.4: Learning from Glasgow 2014 Commonwealth Games SWF Trial.* Surface Water Flooding Hazard Impact Model – Phase 2 Project Note, 8pp.

Cole, S.J., Moore, R.J., Wells, S.C. and Vesuviano, G. (2016) *Further Development of the SWF Hazard Footprint*. Surface Water Flooding Hazard Impact Model – Phase 2 Report to the Flood Forecasting Centre, Research Contractor: Centre for Ecology & Hydrology.

Cole, S.J., Moore, R.J. and Mattingley, P.S. (2015) Surface Water Flooding Component for NHP HIM: Phase 1 Report. Contract Report to the Environment Agency, Research Contractor: Centre for Ecology & Hydrology, Environment Agency, Bristol, UK, 55pp.

Escobar, M.P., Vinogradova, M. and Demeritt, D (2016) *Using newspapers as a source of data to assess flood impacts: methodology note.* Department of Geography, King's College London.

Flood Forecasting Centre (2015a) *Case Study Selection – Summary*. Surface Water Flooding Hazard Impact Model – Phase 2 Project Note, 8pp.

Flood Forecasting Centre (2015b) *User specification for the visualisation of the Surface Water Flooding Hazard Impact Model*. Surface Water Flooding Hazard Impact Model – Phase 2 Project Note, 8pp.

Gunawan, O. and Aldridge, T. (2016a) Surface Water Flooding Hazard Impact Model – Impact Library refinement and testing: Phase 2 Report. Technical Report MSU/2016/20 for the Flood Forecasting Centre, Research Contractor: Health & Safety Laboratory, 35pp.

Gunawan, O. and Aldridge, T. (2016b) Surface Water Flooding Hazard Impact Model: uFMfSW Assessment. Technical Report MSU/2016/22 for the Flood Forecasting Centre, Research Contractor: Health & Safety Laboratory, 29pp.

Met Office (2016) The Met Office ensemble system (MOGREPS). http://www.metoffice.gov.uk/research/areas/data-assimilation-and-ensembles/ensemble-forecasting/MOGREPS

Mooney, J. and Lane, A. (2016) Structure and configuration of end-to-end trial system. Surface Water Flooding Hazard Impact Model Project Note, Met Office, Exeter, UK, 10pp.

Pierce, C. (2015) Briefing on the implementation of MOGREPS-UK based blended ensemble precipitation forecasts. Met Office briefing note, 14 May 2015, 7pp.

Warren, S., Hunter, N. and Revilla-Romero, B. (2016) *Linking G2G to JFlow Inundation Model: SWF Hazard Footprint*. Surface Water Flooding Hazard Impact Model – Phase 2 Report to the Flood Forecasting Centre, Research Contractor: JBA, 72pp.