everest

D3.4
Natural Hazard Use Case Demonstration Report

Workpackage	**WP 3**	**VRE User Cases**
Task	3.4	Natural Hazard Use Case Demonstration Report
Author (s)	WP3	NERC
Reviewer (s)	Stefano Salvi	INGV
	Ugo Di Gianmatteo	ACS
	Sergio Ferraresi	MEEO
Approver (s)	Rosemarie Leone	ESA
	Cristiano Silvagni	ESA
Authorizer	Mirko Albani	ESA
Document ID	EVER-EST DEL WP3-D3.4	
Dissemination Level	Public	
Status	Approved by the EC	
Version	1.1	
Date of issue	14 December 2018	

Document Log

Date	Author	Changes	Version	Status
01/06/2018	ESA	ToC definition	0.1	Table of Content
15/07/2018	NERC	First draft	0.2	Draft for Review
28/09/2018	NERC	First internal review	0.3	Under Review
16/10/2018	NERC	Internal review consolidation	1.0	Draft to be approved by the EC
14/12/2018	Cristiano Silvagni	Document status	1.1	Approved by the EC

Table of contents

List of Figures

Definitions and Acronyms

Acronym	Description
COTS	Components Off-the-Shelf
DHA	Daily Hazard Assessment
DOI	Digital Object Identifier
DoW	Description of Work
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable Data
FFC	Flood Forecasting Centre
G2G	Grid-to-Grid
HIF	Hazard Impact Framework
HIM	Hazard Impact Model
HIPS	Hazard Impact Production System
HSL	Health and Safety Laboratory
IP	Intellectual Property
NHP	Natural Hazards Partnership
PID	Persistent Identifier
RTDI	Research, Technology Development and Innovation
R&D	Research & Development
SA	Service Activities
SWF	Surface Water Flooding
TBC	To Be Confirmed
TBD	To Be Defined
VOT	Vehicle Overturning
VRC	Virtual Research Community
VRE	Virtual Research Environment
WFS	Web Feature Services
WMS	Web Map Services
WP	Work Package

Applicable and Reference Documents

Document ID	Document Title
[1]	European Virtual Environment for Research - Earth Science Themes, Grant Agreement N° 674906 – Available on the EVER-EST Alfresco Platform under WP1 Management Folder
[2]	EVER-EST Description Of Work – Available on the EVER-EST Alfresco Platform under WP1 Management Folder
[3]	EVER-EST Consortium Agreement – Available on EVER-EST Alfresco platform under WP1 Management Folder
[4]	European Commission Research and Innovation Participant Portal, http://ec.europa.eu/research/participants/portal/desktop/en/home.html
[5]	EVER-EST Project Web Site, www.ever-est.eu
[6]	D. 3.1.1 Deliverable: EVER-EST VRE Use Case Description and User Needs
[7]	D. 3.2 Deliverable: EVER-EST VRE Validation Plan
[8]	D. 2.5 Deliverable: Report on training activities
[9]	D. 6.2 Deliverable: Report on VRE population and testing
[10]	D. 2.9 Deliverable: Report on dissemination activities
[11]	D. 2.8 Deliverable: Sustainability Development Plan
[12]	D. 3.7 EVER-EST VRE Overall Impact Assessment Report

Abstract

This report describes the validation of the implementation of the Natural Hazards use case in the EVER-EST VRE solution deployed by WP6 against the user requirements documented in AD [6]. It has been assessed for enhancing collaborative working between all partners of the Natural Hazard Partnership for the purposes of further developing the Hazard Impact Modelling process relating to specific natural hazard forecasting for surface water flooding. In addition, its value in archiving related documents together in Research Objects in large volumes plus the shared storage of important datasets whilst preserving their security.

1 Introduction

1.1 Document scope

This document describes the validation of the implementation of the Natural Hazards use case in the EVER-EST VRE solution deployed by WP6 against specific user requirements. It has been assessed against specific requirements that have deviated to some extent from the original user requirements as documented in AD [6]. These include the development of workflows and associated code used to develop Hazard Impact Models on the VRE platform, specifically for surface water flooding, and the subsequent sharing of that code and workflows with members of a small development team. Additional functionality tested includes the archiving of related documents as Research Objects for the purposes of archiving and sharing plus the storage of a range of datasets, some of which require additional security settings to restrict use. An overall requirement for the VRE was to enhance collaborative working across the NHP between its 17 partners.

1.2 Relationships with other EVER-EST work packages and deliverables

The overall objectives of WP 3 are:

- To liaise with the virtual research communities (VRCs) participating in the project and detail typical use cases, starting from and further detailing the ones identified at proposal time and specify the common user requirements arising from the analysis and refinement of VRCs use cases (described in [AD 6]);
- To define user requirement validation methodology and success criteria (described in [AD6], [AD7]);
- To validate and assess, against the above mentioned user requirements, the implemented use cases, through the deployed VRE solution (this report);
- To perform and report on the overall assessment, in terms of user uptake and societal benefit, of the overall infrastructure [AD10].

In line with EVER-EST VRE Validation Plan [AD 7], Natural Hazard VRC validation activities have been carried out in strict alignment with WP6. During the VRE validation, research objects were generated, user communities trained, and dissemination events attended and carried out.

To avoid duplication of input and assure coherency of content among the various documents, the relationship between WP deliverables is as follows:

- Validation activities: use of the VRE to assess whether it meets the needs of the Natural Hazard user communities and use case identified [AD6] are reported in this document.
- Verification activities: to verify whether EVER-EST VRE components and systems are compliant with design, well-engineered, error-free and properly integrated. These activities are out of the scope of this validation report document and reported in [AD9]. Coordination with the verification activities under WP6 for the specific Natural Hazard use case testing is reported in [AD9].
- Population: Natural Hazard VRC use case scenarios research objects generated during the project are listed in [AD9]
- Training: Natural Hazard VRC training activities and capacity building events detailed list is provided in [AD8]
- Dissemination: Natural Hazard VRC dissemination activities and awareness raising to new users in order to enlarge current user base and attract new communities. List of activities is detailed in [AD10]

- Sustainability: Natural Hazard VRC value proposition statements as derived from VRE validation and usage and inputs to business model is addressed in [AD11]
- User Uptake and Key Performance indicators: Natural Hazard user uptake assessment has been provided as input for the relevant description in [AD12]

2 Natural Hazard Virtual Research Communities

The Natural Hazards Partnership (NHP) is a group of 17 collaborating public sector organisations comprising government departments, agencies and research organisations. The NHP provides a mechanism for providing co-ordinated advice to government and those agencies responsible for civil contingency and emergency response during natural hazard events. The NHP provides daily assessments of hazard status via the Daily Hazard Assessment (DHA) to the UK responder and resilience communities, pre-prepared science notes providing descriptions of all relevant UK hazards and input to the National Risk Assessment. In addition, the NHP has set up a Hazard Impact Model (HIM) group tasked with modelling the impact of a range of UK hazards within a common framework and operational delivery of the model outputs. The partnership is currently UK-based with offices located across the UK. The work of the partnership is monitored by a steering group comprising a representative from all partner organisations.

The NHP use case to be validated by use of the VRE constitutes the following:

1. Establishment of a forum for the exchange of knowledge, ideas, expertise, intelligence and best practice in relation to natural hazards and their impacts;
2. Provide a timely, common and consistent source of advice to government and emergency responders for civil contingencies and disaster response;
3. Create an environment for the development of new services to assist in disaster response;
4. Availability of services improving collaboration among participants and paving the way for a more effective support;
5. Provide a research environment for the development, testing and sharing of new methods.

2.1 Natural Hazard VRC user requirements

2.1.1 Natural Hazard VRC Users Operational Scenarios

The NHP partners recognised the following operational scenarios that could be tested and used to validate use of the VRE:

- The integration and homogenisation of available data from different sources including in-situ data or other external sources which is searchable and provides easy access for a specific need;
- Linking scientific experts and institutional/stakeholder communities allowing exchange of information;
- Design and implementation of a repository for sharing analytical methods, protocols and algorithms for hazard impact modelling assessments and forecasts;
- Implementation of Research Objects including all the necessary info, targets and monitoring protocols to be shared among the user communities.

2.1.2 Natural Hazard VRC Service Requirements

The service requirements identified by the VRC are as follows. In addition, in this section, more detail is given to any original requirements that have been met, partially met, not met, or no longer needed.

- E-collaboration services; the community needs to communicate and share knowledge, ideas and protocols and have quick communication with institutional entities and agencies;
- E-research services; services needed to share data, protocols and advice. The community needs to guarantee provenance and traceability of data and advice;
- E-learning services; for educational activities, and for internal training purposes;
- Digital Information services; to achieve an easy acquisition and discovery of data, to create and access data and text mining tools as well as automatic tools for data processing and relevant relationships models.

E-collaboration services – Services that deal with collaboration between users

The following e-collaboration services were included in the NHP's initial user requirement. However, the NHP partners have not taken advantage of the services offered. As the project progressed, the VRC partners decided to give priority to using the VRE as a research and development tool that would enable more efficient collaboration on this important aspect of model development. It was also felt that existing messaging and chat facilities such as Skype, Webex™ and telephone conference calls provided adequate communication options enabling the VRC to concentrate on its priority requirements. However, to have all these facilities on one platform in the future would provide advantages beyond those already offered by existing systems.

Hazard Impact Model (HIM) progress reporting

- A space/forum for HIM members to post and briefly describe progression points as they occur.
- A single forum where anyone in the HIM can view the current state of affairs with technical and scientific progress.
- This feature would allow us to reduce frequency or refine the format in which work packages are updated in face-to-face meetings, thus freeing up time for more productive activities.

An instant messenger/chat service

- HIM members log in at the start of a working day therefore becoming 'available' on the system, allowing improved instant communications that could be useful for arranging last minute meetings, clarifications, or other collaborative discussions. This could then serve as a log of these communications.

Discussion Forums

- Separated into themes so could be used by different groups within the NHP for organisational and/or technical/scientific conversations.
- To include keyword tagging to relate conversations to specific topics which are searchable.
- Some form of capture and archiving of discussions would be useful. This should be supported by a simple cataloguing function.

Conferencing

- Tools for webinar/video conferences.
- Screen sharing capabilities would be useful.
- The ability to conduct and present demonstrations using shared data/tools.
- Requires easy visualisation options and information exchange.
- Should show who is online and where and enable simple click-and-connect functions

The following e-collaboration services requirements have been met by the VRE. Whilst a dedicated document storage and sharing system has not been specifically provided for documents, any documents can easily be stored on the Seafile storage system or packaged up as Research Objects and stored on ROHub. Folders and Research Objects can be shared with VRC partners and different security settings can be attributed to each storage option.

Document Storage and Sharing

- A document repository with editing capability.

- Needs to be accessible by different partner organisations with different internal security issues.

The following project management tools originally requested were not highlighted as a priority for the VRC. The VRC has access to its own file storage for archiving project reports, minutes and actions from meetings.

Project Management Tools

- A central site for project management tools would assist in increasing efficiencies within project management including facilitating a greater level of communication.
- Tools would include presentation and editing of Gantt charts, and reporting (e.g. by-monthly reports).
- Secure storage of minutes and actions, where access is limited to core project team, would be useful.

The following general e-collaboration services have largely been met. The interface is largely engaging and user friendly although there are some small improvements that could be made. However, it is generally acknowledged that the perfect system does not exist and users' preferences can make designing the 'perfect' system challenging. Although not deemed a priority, as mentioned above, providing a range of e-collaboration services alongside a development environment would enhance collaboration amongst the VRC members. This would need a much larger investment therefore is not an expected requirement. Single sign-on has been provided and works well.

General

- The service must be engaging and user friendly.
- Many tools exist e.g. Facetime, Skype, web-based meeting services. However, our experience suggests that few have access to all these tools at any one time. Providing a service that is readily available to all participants is beneficial.
- Needs to be readily visible and easy to access and visually pleasing.
- Single sign-on preferable.
- Should be available on a range of platforms.
- Should help service partner organisations in developing joint messages for various stakeholder communities.

<u>E-Research services – Services which support research processes and tools</u>

All the following requirements have been successfully met and work well. However, work done during the development stages have not yet been transferred to the operational environment as development work is still ongoing. There is currently no capacity to deliver outputs directly from the VRE to the UK responder community and it is unclear as yet whether this will be a feature in the future.

Online Development Environment

- Collaborative Hazard Impact Model development could be undertaken on a development platform hosted on EVER-EST.
- Case studies could be re-run with updated datasets, configurations and parameters.
- When development stages are completed the work would be transferred to the operational environment.

- The VRE would need to enable partners to interact regardless of their preferred platform of use e.g. Linux and Windows.

The following services have been successfully met. The development team has successfully deployed R scripts into Taverna workflows which have then been packaged up as workflow-type Research Objects. It is clear that other scripting languages (e.g. python) could also be used in this way therefore enhancing opportunities to collaborate over technical developments. Whilst modules of generic model building tools have not been set up, it is clear that this is a possibility with a Research Object. This would enable all relevant tools to be housed in a Research Object together with associated metadata and user guidance. A library in Seafile is available to store all available workflows and data. Whilst this is not a visual library (see below), the interface is straightforward and requires very little user instruction. Data and documents can be uploaded to Seafile and included in Research Objects via a library of tools (including drag and drop functionality) that is available on the VRE platform; this works well.

Facility for working code

- Somewhere to place workable scripts (R, Python etc) for ourselves and others to use, develop, copy and amend as necessary. This would enhance collaborative research efforts between organisations.
- Provision of modules of generic model building tools, with suitable metadata and user guidelines and accessible via simple search queries would be useful.
- Modules could be user generated and sit in a visual library of tools that is easy for other users to pick from (users should be able to contribute to and pick from the library).
- This facility should include proper code versioning and archiving where necessary.
- Needs to allow for multiple scripting languages.
- Library of tools could be provided by a smart user interface providing easy drag and drop facility.

Use of Taverna has allowed workflows implementing different tasks to be compiled. This includes complex multi-stage scientific processing as well as simpler input output functions. Access to off-the-shelf software has not been provided on the platform, however access can be provided through the use of Web Processing Services hosted by users. On further defining the user requirements prior to commencement of the project, a virtual machine giving access to additional software was not requested. However, an updated request has now been made to the technical team for access to R on a virtual machine via the VRE. This is in progress therefore cannot be validated in this report.

- The VRE could provide access to modules at different levels of complexity (from detailed process models to simple I/O structures).
- Access to off-the-shelf software would also be valuable, especially ESRI tools, which are commonly used – licensing, may need to be arranged from one of the partners.

The following requirements have been satisfactorily met. Code and tools are stored within workflow Research Objects and can be shared and re-used easily by other NHP partners. However, hazard impact modelling is still in the research and development phase and therefore full testing in an operational setting is not currently possible. It is expected that with time, this requirement will be fully tested by a wider range of stakeholders.

Model Testing/Validation

- Functionality to run release versions of code/tools could open up the work to partners within NHP.
- This could be used for validation and testing of models by stakeholders (e.g. Flood Forecasting Centre (FFC)), project partners (e.g. Kings College London (KCL)), or wider – into other related projects such as FFIR (Flooding From Intense Rainfall) or other academic organisations.

The following requirement was considered outside of the VRC's immediate user requirements.

- Technology to automatically populate an Impacts database e.g. social media 'scraping' tool. This would be bespoke to HIM models and the impacts they forecast enabling access to 'easy to process' impact verification data.

The following requirements have been fully met. Taverna workflows have been successfully incorporated into Research Objects which can be versioned using Fork and Snapshot tools.

Workflow

- Ability to link Research Objects and work-flows/thought-flows together.
- Function to enable documentation of model development e.g. model history – what worked/what didn't (lessons learned).
- Facility to provide ideas at each stage of model development e.g. ideas/thoughts added to Research Object

The following requirement has been met. Visualisation tools within the VRE allow the components of ROs and the results of case study workflow runs to be visualised. The tools enable users to view the data in different ways

Visualisation

- Innovative, interactive and flexible visualisation tools for model outputs and data.

<u>Digital Information services</u>

This requirement has been fully met although has not been used to its full extent. Any required datasets can be stored on Seafile and accessed by those that need the data.

Validation database

- This database represents a single source of observed impact data for validation purposes.
- The database would include currently used sources of historical impact information such as the Nexus Lexus and Flood Forecasting Centre resources as well as data collected from additional tools developed in EVER-EST (e.g. twitter scraping, web scraping, other social media sites)
- The database could be made publically available as a single source of impact data to be used for further validation and research.

The facility to store and provide access to a range of datasets is available on Seafile. Data can also be stored as Data-type Research Objects, providing the option to store data with associated metadata. The facility has been used to store original input data for workflow-type ROs, but its use has not been widespread. There is currently no immediate need to provide wider access to some of the larger datasets held by partner organisations such as the National Population Database. An additional feature has been provided – access has now been provided for discovery of Sentinel 1 data. This was not identified as a requirement at the start of the project but has since been added as the project has progressed.

Data sharing

- To be provided by live data services, if available.
- Ready access to static saved data e.g. case study based data. This would allow us to more effectively build models using readily available 'sample data'.
- Easier access to large datasets (e.g. National Population Database, Met Office data) to be provided online
- Alternative could be provision of access portals to database servers at partner organisations.
- Access platforms to geospatial data need to follow common standards – one access point to many disparate data sets.
- Data access method would need to allow extraction and interaction.
- Multiple data formats.

There is currently no facility to share results with NHP end users (i.e. those who receive natural hazard advice from the NHP) since the work of the NHP is still in development phase. With time, it will be useful to explore this as an option.

End users

- Could the VRE enable partners to share results or real-time data with end users (in trial mode only) – e.g. via WMS

The following requirement was not considered a priority, however all the information listed below could be stored in Research Objects with appropriate names and accompanying metadata.

Knowledge Base

- This would provide a central source of information which might include: journal article library of relevant science/technology papers; available skill sets held in each organisation (providing access to additional information/help); CVs of all EVE-REST collaborators/EU partners providing information on relevant expertise and sources of help and possibly data, metadata and glossary of terms.

<u>E-Learning services – Services providing access to learning materials</u>

The following tools have been made available, however the NHP has not had an opportunity to test and use these tools and therefore they cannot be validated at this time.

Knowledge transfer

- User-created tutorials. Could include walk-throughs of visualisation tools and references for stored code

Exploratory analysis/data playground

- Tools and services that are not currently used but might be available to be explored for integration in the VRE. This could include: analysis of social media, use of large-scale data processing tools (e.g. Hadoop, Apache Spark), and other open source software. This could utilise learning packages from the www (e.g. pluralsight).

2.1.3 Natural Hazard VRC Use Case Selection Criteria

Originally, the Hazard Impact Model (HIM) group was set up to:

1. Coordinate partner-led research projects to develop Hazard Impact Models capable of consistent quantification of multiple impacts from a range of hazardous natural elements to allow better understanding of the impact risks of single and multiple hazard events on critical assets and services.
2. Coordinate partner-led projects to develop an operational delivery system, the Hazard Impact Production System, an integrated, community of open platforms whose collective outputs are delivered to stakeholders such as category 1 (fire, police) and 2 (local government, utilities) responders.
3. Coordinate improved access to spatially and temporally explicit information about the occurrence and magnitude of natural hazards from media reports and various administrative records maintained by the emergency services and critical infrastructure providers via a Hazard Impact Database.

Hazard Impact Models (1.) are being developed for three hazards in the first instance: Surface Water Flooding, Land Instability and High Winds, via 3 distinct work NHP packages. Further work packages cover other HIM group activities, and include: a Vulnerability and Impact work package to develop a standard approach to modelling of the impact of these hazards; a Hazard Impact Productions System (2.) work package to develop and test the operational system and delivery of the impact model outputs; and a Hazard Impact Database (3.) work package to develop and apply data collection methods for model validation and testing. Each work package is led by one of the NHP partner organisations but requires significant collaboration by one or more of the other partners.

As the work is currently in research phase, the HIM group decided that the most advanced piece of work, the Surface Water Flooding work package (a), would provide the best opportunity to to use and validate the VRE. Following use of the VRE during the Everest project, a new use case opportunity arose – the creation of bibliographic ROs for storing and archiving documents used for decision making by the NHP (b).

(a) Surface Water Flooding work package – led by the NERC Centre for Ecology & Hydrology (NERC CEH)

The principal objective of this work package is, through joint working within the NHP, to develop and trial improved capabilities for forecasting surface water flooding (SWF) risk and its impact for emergency response and possibly alerting purposes. The work package involves input from hydro-meteorologists at the Flood Forecasting Centre (FFC) for England & Wales and impact/vulnerability specialists at the Health and Safety Executive (HSE). An end-to-end offline trial system has been demonstrated using historical case studies. Ongoing work is developing an end-to-end real-time trial system and gaining evidence from further offline case studies, sensitivity analyses and model validation (supported by King's College London). Offline fine-scale inundation

modelling by JBA is testing and possibly improving the link between CEH's hazard and HSE's impact assessments of SWF risk.

Collaboration

Our first priority was to test how the VRE might enhance collaboration amongst the Hazard Impact Model group partners. Currently, collaboration relies on the following methods of interaction.

- **Physical meetings**. These are held 4 times per year at a different partner organisation location across the UK. Work updates and plans are provided and any key issues requiring decisions are discussed.
- **Teleconferences**. These are held on an ad hoc basis as and when issues arise or specific discussions are required.
- **Physical work package meetings**. These meetings are held more regularly and are used to discuss technical developments, share expertise and discuss data requirements.
- **Ad Hoc physical meetings**. These are occasionally set up to discuss other overarching issues relevant to the Hazard Impact Model group. For example, a workshop was held to discuss the overarching Hazard Impact Framework, which is being designed to help ensure impact modelling, is developed by each work package in a consistent manner.

Hazard impact modelling requires participating partners to collaborate and share large datasets of different types. These include socio-economic datasets as well as scientific datasets. The Surface Water Flooding work package provided an ideal test case for testing interactive hazard modelling, impact analysis and verification between partners. This work package provided a useful starting point for testing how the VRE might support sharing of common methodologies and ontologies across different Hazard Impact Models.

As the user requirements were more tightly defined, the VRC decided to concentrate on the development of the surface water flooding models as this work package is the most advanced and provided the best opportunity to test the functionality of the VRE. Other work packages highlighted in the original user requirement were not considered as high priority as it was felt the surface water flooding experience, if successful, could be applied to a range of other hazard impact models once the research and development was more mature.

(b) Document storage and archiving

An additional requirement not highlighted in the original user requirement was the ability to store multiple Daily Hazard Assessments (DHA) and associated documentation. The DHA is a bulletin that is released once a day (and more frequently as required) and sent to all registered users – commonly the responder community, local government, and national agencies such as the Environment Agency. The DHA is a summary document detailing hazards forecast for the following 24 hours. Hazards include surface water flooding, high winds, landslides, wildfires etc. These summaries are compiled based on contributing evidence such as the Flood Guidance Statement (FGS) issued by the Flood Forecasting Centre, the National Severe Weather Warning Service (NSWWS) issued by the Met Office and the Daily Landslide Hazard Assessment (DLHA) issued by the British Geological Survey. Whilst the DHA is issued via the Met Office's Hazard Manager online service, there is no one location where all pieces of contributing evidence were stored. On discovering the power of Research Objects as a way of collecting and storing related pieces of information and the existence of the Bibliographic Research Object, the VRC decided to add this as a new user requirement.

2.1.4 Natural Hazard VRC selected use case descriptions

Use Case 1 - Surface Water Flooding The SWF HIM is a well-developed Hazard Impact Model approaching operational deployment with on-going work focussed on validation of impacts through chosen case studies, See figure 1.

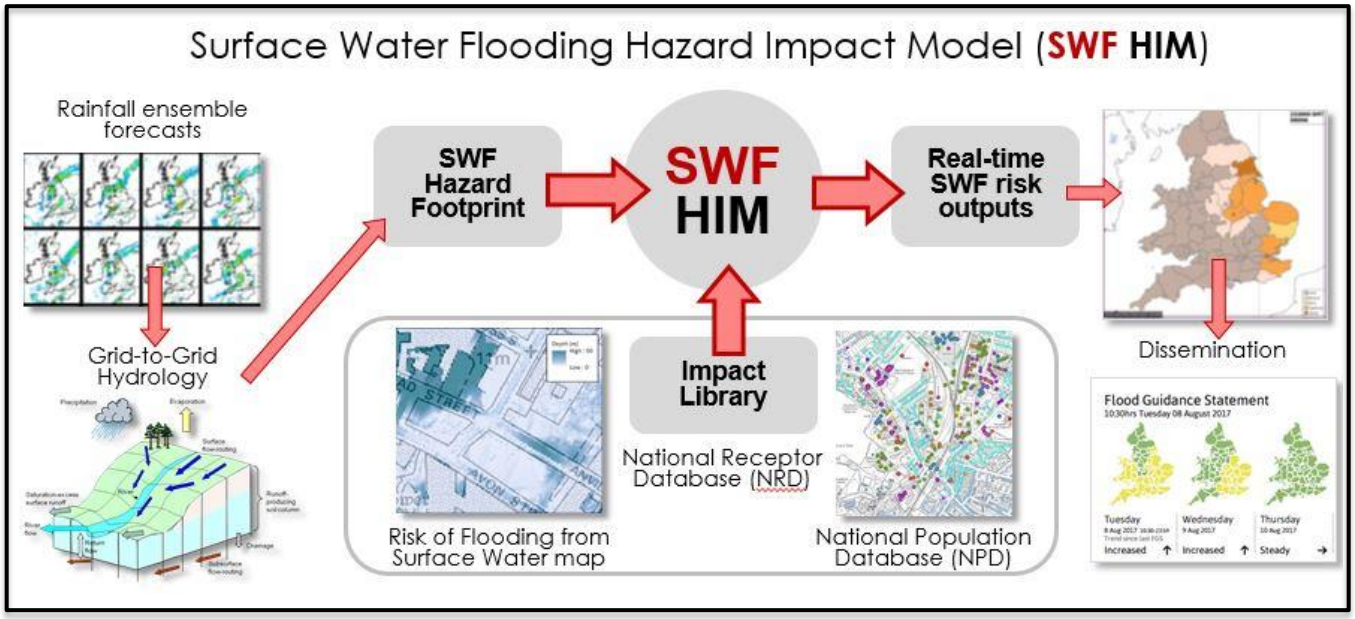

Figure 1 – Surface Water Flooding hazard Impact Model (SWF HIM)

This involves running a countrywide (1km grid, 15 min time-step) Grid-to-Grid (G2G) (https://www.ceh.ac.uk/services/flood-modelling) hydrological runoff and routing model (CEH) using rainfall inputs (Met Office), and linking its surface runoffs to potential impacts (HSE) and verifying these against observed impacts (King's College London, KCL). Currently this is achieved through accessing Met Office rainfall data (observed, forecast deterministic & ensemble) via a live data feed to CEH and retrieving archive data from the Met Office Managed Archive Storage System (MASS) for past events. A Big Data Analysis Platform, JASMIN, supports the latter with its MASS-on-JASMIN facility available to the NERC community, Met Office and environmental academia across Europe. FTP is used to share G2G model data from case study event runs between CEH and HSE, whilst email suffices for exchange of impact validation data. A live trial operational workflow producing and visualising SWF impact products was developed in 2016.

Following considerations of detailed user actions, the VRC concentrated on testing a smaller part of this workflow – the implementation of the impact modelling process once the hydrological modelling was completed in G2G, see figure 2.

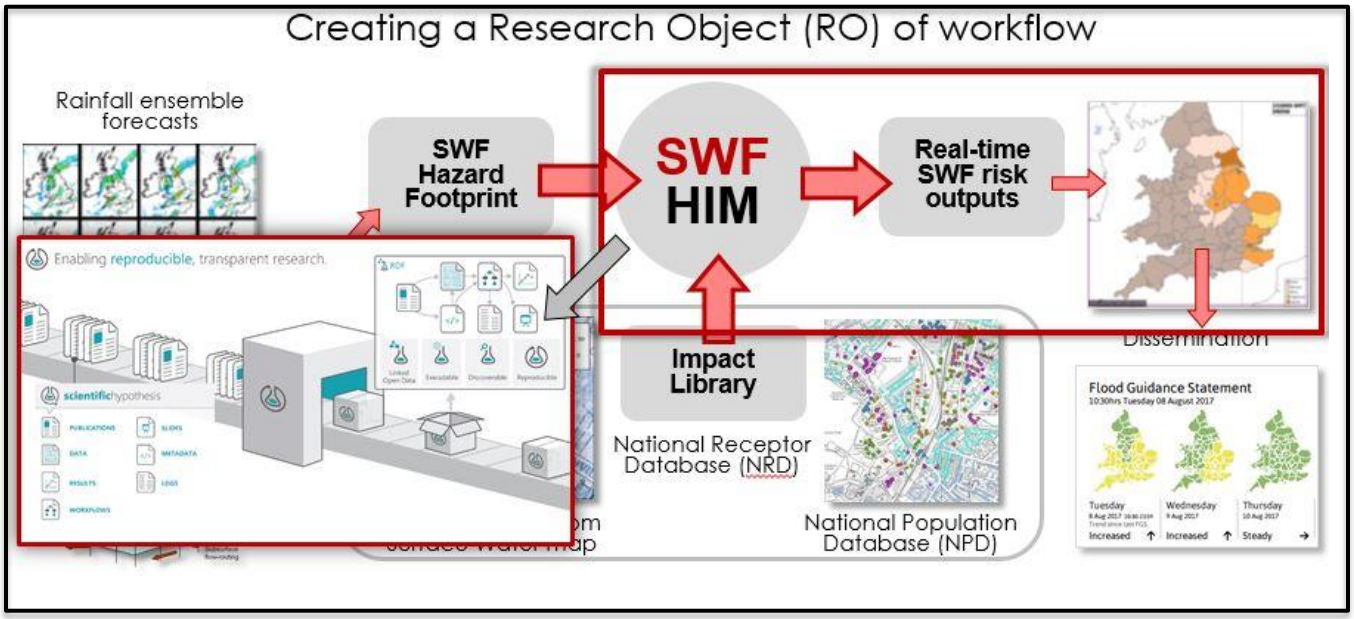

Figure 2 – Part of SWF HIM used to test creation of workflow Research Objects

The workflow utilises R and was successfully implemented on the VRE using the workflow management software Taverna. Each part of the workflow was re-written so it could run online in a virtual environment with all input data stored in the cloud (Seafile), and using Taverna Server and Rserve applications hosted on the VRE, instead of desktop software. Further, the workflow was parameterised so that users working in the VRE could specify key input variables without amending code. This enabled the process to be automated giving the potential for other VRC members to run the workflow. The completed automated workflow is stored as a workflow Research Object. This workflow is accessible to all VRC members that have access to the correct Seafile libraries. This workflow can easily be searched using the RO Search option and run multiple times. Workflow results, usually png and shape files, can be inspected within the VRE using the Virtual Globe or the integrated reader. Furthermore, they can be downloaded, stored in the current Research Object or in a new one.

Use Case 2 - Daily Hazard Assessment (DHA) The DHA is a summary of forecasted hazards released on a daily basis to the responder community, local government and national agencies. It is based on information provided by various partner organisations including the FGS, the NSWWS and the DLHA. Each contributing piece of evidence is submitted to the Met Office for inclusion in the DHA. Each piece of evidence is linked by date, however if any of the evidence is updated due to a change in the hazard forecast, then an updated piece of evidence is submitted for inclusion in the DHA. In addition, the NSWWS is not issued every day therefore there is not always a simple linear relationship between the DHA and its contributing evidence. Whilst all pieces of evidence are collected by the Met Office for inclusion in the DHA, there is no single repository for all this information – each original document is stored by its originating organisation i.e. the FFC stores all FGSs.

The VRC decided to test the storage of each DHA and its contributing evidence in a bibliographic Research Object. The RO contains all documents in a series of logically named folders to aid easy access. The RO also contains a sketch showing the user what is contained in the RO. 96 of these types of ROs have been automatically generated to date and are accessible by the user via the date on which the DHA was released. To fully test the usefulness of these types of ROs, one RO per day needs to be generated automatically.

2.1.5 Natural Hazard VRC use case data needs

The majority of datasets that are used are UK focussed, but there may be a future requirement to widen this scope. When the Icelandic Volcano eruption occurred in 2011, there was a need to share data across a wider area than just the UK.

Different types of data and information that were considered a requirement by the VRC included:

ESRI Shapefiles, MapInfo Files, Web Map Services (WMS), Web Feature Services (WFS), Met Office data (NIMROD or ASCII format), GRID Files, ESRI File geodatabases. Some datasets are static (only exist once), whilst others have a temporal aspect so will be updated, or at least new files uploaded, at regular intervals.

Specific examples included:

1. Surface Water Flooding (SWF) hazard footprints produced by G2G. These are ASCII grids either based on historical case study data or weather forecasts.
2. The SWF Impact Library used to estimate impacts in the SWF HIM. This is a set of 1km ASCII grids.
3. County areas – these are reporting areas used by the Flood Forecasting Centre and are used to generate county-level summaries of impacts. These data are held as ESRI shapefiles (vector).
4. Met Office Data - Raster data (NIMROD or 2km ASCII grid) e.g. Temperature, Relative humidity data, Precipitation Forecast
5. National Population Database (owned by Health and Safety Executive) - Is a vector dataset which contains details of population types, their locations, scale (how many people) and vulnerability ((e.g. how they might be susceptible to a hazard)
6. National Receptors Database (owned by Environment Agency) - this contains information on assets and infrastructure including roads, railways, buildings etc
7. BGS GeoSure Data – Is a vector dataset, containing information about different geological hazards in the UK. Of particular interest is the Slope Instability layer which is held as ESRI shapefiles.
8. Landslide hazard footprint produced by the British Geological Survey. These are held as ESRI shapefiles.
9. Landslide domains layer – this is a regional characterisation of landslides across the UK based on dominant failure type, style and geomorphology. These are held as ESRI shapefiles.
10. The National landslide database provides a record of landslide events and associated information for the UK (excl. Northern Ireland, Isle of Man and Channel Islands). It is an Oracle database linked to ESRI point shape files and excel spreadsheets.
11. Satellite imagery – Sentinel 1 and 2.
12. Various validation datasets including Nexus Lexus database (managed by KCL); Met Office WoW (Weather Observation Website) data; Other crowd sourced weather/impact data; Data that can be harvested from Social media.

Whilst all the above datasets are important to the work of the hazard impact modelling group, only the datasets 1-3 were used to validate the VRE. All 3 datasets are required for the generation of impact estimates from the SWF HIM. The success of the workflow run tests and subsequent model outputs, demonstrate that the required data access has been provided. In addition, access to Sentinel data has also been provided in anticipation of a future need. There is scope for storing all datasets listed in Seafile therefore providing the potential for widening the use of the VRE to include hazard footprint generation and subsequent impact modelling in a virtual space.

2.1.6 Natural Hazard VRC use case scenario for research object modelling

Use Case 1 – Hazard Impact Modelling

The VRC recognised that each Hazard Impact Model could be described as a Research Object. The Research Objects would be defined by configurations of common components including:

- Standards/Definitions – to support the consistent development of individual Hazard Impact Models to defined standards with consistent terminology;
- Approaches/Methods – based on case studies, development of re-usable impact modelling methods for individual hazards, guidance notes etc.;
- Impact descriptions and severities – standard terms and ontologies;
- Scenarios and narratives used to inform and verify new hazard impact models and scenario testing;
- Data formats;
- Data management standards;
- Code libraries for re-use by partners.

Components of the HIM are developed in a modular fashion as standalone components, each of which could also be defined as a Research Object. For example:

- The G2G component of the Surface Water Flooding HIM and its associated method descriptions and guidance notes etc.;

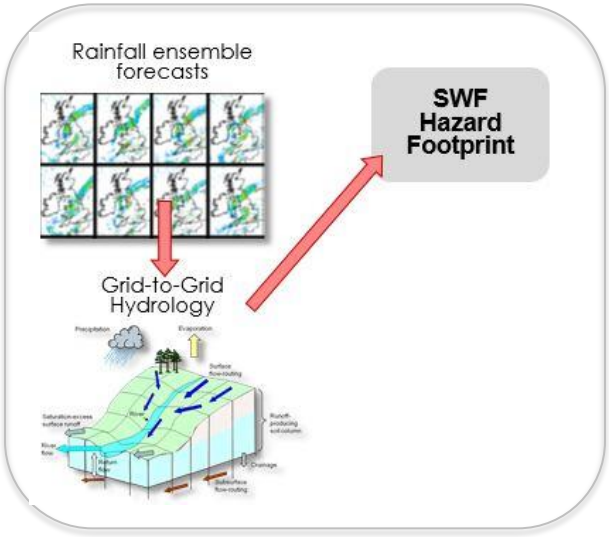

Figure 3 – Grid-to-Grid component of the Surface Water Flooding HIM

- The Impact Library set up as a repository of pre-calculated impact data for each HIM. The data sources and methods used are HIM-specific and developed and improved based on research, evidence and stakeholder input;

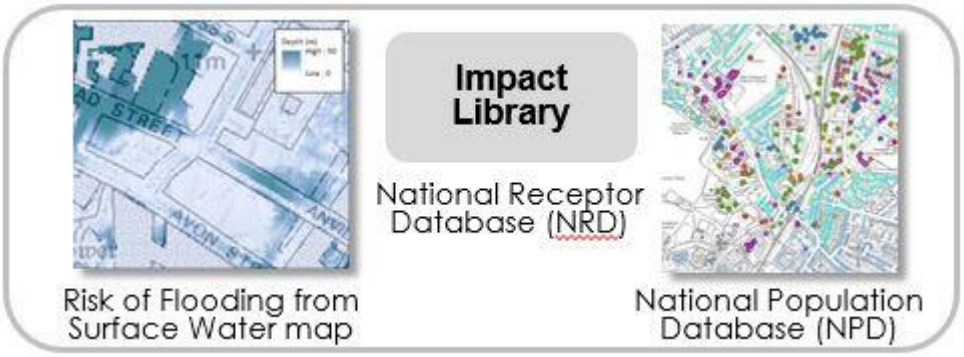

Figure 4 – The impact library component of the Surface Water Flooding HIM

- The risk/summary/visualisation component. The raw outputs of the HIM include impact results that require summary and presentation to end users, including an interpretation of the risk.

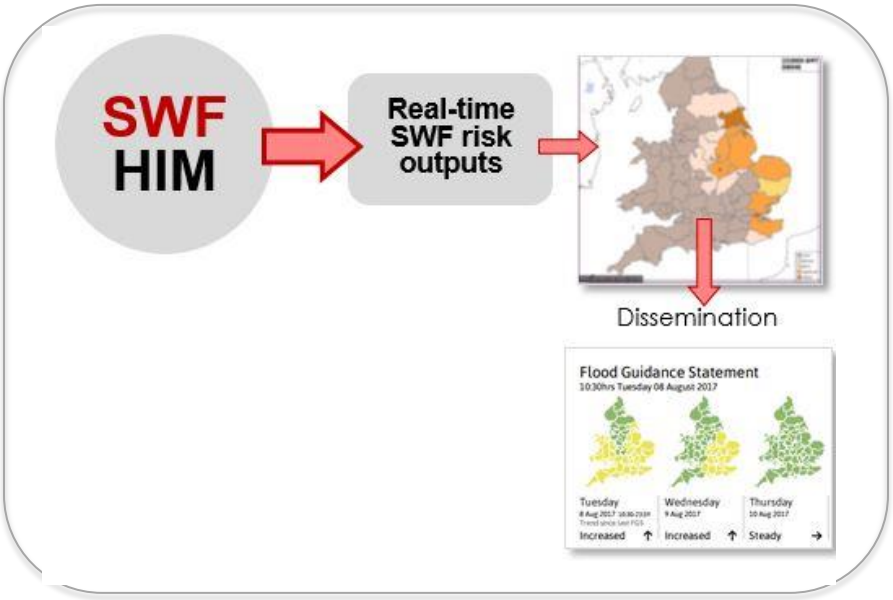

Figure 5 – The risk/summary/visualisation component of the Surface Water Flooding HIM

As described in 2.1.4, the VRC decided to concentrate on one part of the SWF HIM workflow and has created a workflow Research Object http://sandbox.rohub.org/rodl/ROs/swfhim_v2_0/ .This RO contains all the necessary components to apply and test the part of the SWF hazard impact modelling process depicted in figure 2, and save the results as Research Objects. These components include:

- R scripts encapsulated in a Taverna workflow. Previously, parts of the hazard impact modelling workflow were held in separate R scripts locally on a researcher's machine. Including all these components in a Taverna workflow has enabled the whole process to be automated by linking together each component R script. This also means that the workflow can now be stored on the VRE and access shared libraries on Seafile rather than relying on all data and files being stored locally on the researcher's machine. Encapsulation of this workflow on Taverna has also ensured its standardisation so that other VRC members can access the same workflow and run it.
- Configuration files. These are text files that contain parameters used to specify individual (or batches of) case studies, or that are used to configure the methods used (and being tested) in the SWF HIM. The latter parameters may represent different impact thresholds (e.g. how many buildings need to be modelled as flooded to flag a cell as high impact, or how large the extent of the flood impacts needs to be to flag a certain level of risk/response). A case study workflow can be run through multiple times with different configuration files to see what effect it has on the results. This allows workflows to be tested in order to optimise the configuration across all case studies.

The workflow RO accesses the SWF hazard footprint and Impact Library (see 2.1.5) to generate a hazard impact model output. This output can then be added to the RO and visualised on the VRE. This was tested by 2 additional VRC members. They found that they could search for the workflow RO and associated configuration files, run the workflow and inspect the results within the VRE.

Use Case 2 – Daily Hazard Assessment

As mentioned in 2.1.4, the Daily Hazard Assessment (DHA) is issued every day at 14:00 to category 1 and 2 responders (to support response activities), and local and national government departments (to support planning and resilience). A number of component parts of the DHA are produced by NHP partner organisations

(FFC, Met Office, BGS) and sent to the Met Office's hazard centre for collation and release. There is currently no one single place where all the documentation is stored therefore no easy access to all the pieces of information. An example bibliographic RO (http://sandbox.rohub.org/rodl/ROs/DHA_22_12_2016/) has therefore been created that holds all the component parts of the DHA that is now searchable from within RO-hub. These DHA ROs typically contain the following which are all linked by date of release:

- National Severe Weather Warning (Met Office) – a PDF document
- Daily Landslide Hazard Assessment (British Geological Survey) – an email
- Flood Guidance Statement (Flood Forecasting Centre) – PDF document
- A sketch of the RO content – a JPEG file
- The Daily Hazard Assessment – a PDF document

In addition to creating a golden example of the DHA, automatic creation of a series of DHAs was tested. 96 DHAs were successfully created as ROs and are searchable by date and are publically accessible on RO-hub.

Use Case 3 – Input data

Once the primary use cases (above) had been tested, the VRC decided to test the creation of a data RO. This data RO (http://sandbox.rohub.org/rodl/ROs/SWF_primary_data_for_specified_date/) contains all data and files necessary to run the workflow in the SWF HIM RO and has been assigned a DOI (10.5072/ro-id.XVAUMCCIKW). The data RO contains (also referred to in 2.1.5):

- The Flood Forecasting Centre reporting area boundaries (county boundaries)
- The SWF HIM Impact Library
- Estimates of surface runoff for a case study event (G2G output)
- The configuration files used to run the SWF HIM RO workflow (see Use Case 1 above)

This data RO is searchable and can be referenced using the assigned DOI.

2.1.7 Natural Hazard Use Case Validation Success Criteria

The original success criteria is as follows:

The EVER-EST e-infrastructure shall provide access to generated datasets and research objects as specified in the use case and detailed in the EVER-EST Data Management Plan. A major challenge for EVER-EST e-research services integration is to pursue the following performance and operational requirements success criteria:

1. Efficiency: there should be immediate, clear advantages for the scientists, either in terms of time required to carry out work, or as actual tasks, which could not be performed outside of the environment. This is the most important requirement, since the VRE is aiming to improve the efficiency of the research work, and scientists will not adopt an environment which is perceived as decreasing such efficiency.
2. User-friendliness: all operations should be intuitive and reflect the researcher's way of working. A steep learning curve may result in refusal to adopt the environment after a few attempts.
3. Flexibility: science normally implies the investigation of different lines of thought, sometimes very different. Flexibility in the VRE and an avoidance of rigid structures will allow the testing of different scientific hypotheses.
4. Easy integration of new software tools: the environment should provide interactive access to data processing tools (also from COTS), flexibility and simplicity to implement further tools (to maintain the state of the art).

What follows is an assessment of how well the VRE meets these success criteria as they pertain to the NHP use case.

1. The VRE increases the efficiency of the work of the NHP in the following ways:

a. The automation of the hazard impact modelling workflow within Taverna and stored within an RO means that any member of the VRC can run case studies and create ROs of the results. Not only this, but they can be assured that the ROs will be created to a standard set of rules with relevant configuration files containing tested parameters. This removes the need to test workflows locally and then share results either via emails and face-to-face meetings.
b. Sharing of large datasets held on Seafile allows VRC members instant access to these data in shared libraries whereas previously they were reliant on emailing, FTP and other file sharing applications. The fact that the data and any workflows are stored in the same place decreases the amount of time needed to search for all associated information
c. Instant visualisation of results means that researchers can view results of workflows in real time and share thoughts on these results easily. Previously, researchers had to email results to each other or rely on face-to-face meetings.
d. Access to a powerful online processing environment reduces the requirement to find and use local processing facilities that have sufficient computing power and storage capabilities.

2. The user interface is relatively intuitive although it does need some initial training to ensure VRC members know which functions are relevant to their work. Running of hazard impact models has until now been confined to 1-2 researchers therefore the process is new to other VRC members. For this reason, use of the VRE will be the first time many VRC members have been able to create new models therefore having to learn how to use the VRE to perform this function is not unexpected. Some of the terminology could be made clearer, however, with familiarity, such terminology becomes the norm.
3. The VRE allows the NHP researcher to access a range of different file types, data and workflows. This provides flexibility in the way workflows can be run and what outputs they create. However, by its very nature, the NHP use case is quite restrictive. It is unclear as yet how flexible the VRE would be for new uses within the VRC portal beyond running hazard impact models, creating ROs and searching for data.
4. Access to pre-defined software tools during the user requirement phase are all provided for. However, it is unclear how easy it will be to access software not previously defined in the user requirement. For example, the NHP VRC did not initially request a Virtual Machine for accessing additional software and processing. Subsequently the VRC has discovered that this would be a useful addition for enabling researchers to write and update code live on the VRE, therefore providing a true research and development environment. This point currently requires interaction with the Everest technical team.

3 Earth Science VRE User Requirements Validation

The following requirements applicable to the Natural Hazard VRC have been validated during the project since the deployment of the VRE during the validation and population phase. VRE infrastructure un-availabilities and SPR were raised to WP6 for operational support, bug fixing and work-around according to the procedure in place for anomaly handling and operation support in WP6.

3.1 EVER-EST Infrastructure Key Objectives

Text highlighted in red provides evidence of validation method and results.

EVER- GEN -010	EVER-EST GENERAL REQUIREMENTS	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall allow Earth Science data users to: • Discover, access and process both existing and new heterogeneous Earth Science datasets including the associated information and preserved knowledge held by distributed data centres; • Share data, models, algorithms, scientific results (including traceability of workflows and processes that would facilitate reproducibility of modelling and simulations) and their own experiences within a community or across communities (including those in other domains beyond Earth Science); • Capture, annotate and store the workflows, processes and results from their research activities; • Work together in a real-time environment that facilitates the sharing of expertise, information and data resources overcoming traditional working practices limitations related e.g. to physical meetings or the transfer of large datasets between users; • Ensure the long-term sustainability and preservation of data, models, workflows, tools and services developed by existing communities of practice that can potentially be re-used in the future by other users either for validation of existing research or for new applications.		Met Met Met Met Met at least for the short-term – will be maintained for at least a year
Source	Proposal, SMART OBJECTIVE s		
Verification Method	Verification Methods: Tested using platform to share data, knowledge, models etc.		

EVER-GEN-020	EVEREST GENERAL REQUIREMENTS	Mandatory	OUTCOME
Requirement	EVER-EST e-infrastructure shall enable and facilitate Earth Science Researchers by:		
	• Supporting the processes of conducting research and the creation and maintenance of collaborations across domains, institutions and countries;		Only partially tested – needs engagement from wider community which is ongoing
	• Be designed to meet user requirements and address usability and accessibility, with appropriate evaluation mechanisms and benchmarks for new service and tool development;		Meets most user requirements. Is usable and accessible.
	• Be secure and trustworthy – the EVER-EST components should interoperate with federated cross-institutional authentication and authorisation mechanisms;		Met
	• Be accountable, by providing adequate logging including supporting queries about provenance;		Met but not fully tested as option
	• Be compatible with other widely used and deployed systems, including at least: web, e-mail, instant messaging, SMS, wikis and video-conferencing tools i.e. lightweight desktop applications;		Not met
	• Allow easy and efficient Earth Science data browsing, query and access from multiple sources, adapting to different standards and services;		Partially tested via access to EO data. Other access provided by Seafile but this requires data to be shared by VRC members
	• Support the creation, sharing and curation of digital content, through ease of authoring, publishing, discovery and access. This implies adoption of appropriate metadata schemas and support for automatic generation of metadata. Resources to be described will include data, publications, computation, experimental or observational facilities and human researchers;		Met via ROs, checklists and sharing of data and experiments with colleagues
	• Maximise the processing of the data at EVEREST VRE.		Partially met. Data processing workflows trested on EVEREST VRE but not maximised
Source	Proposal		

Verification Method	Verification Methods: Tested through use of system and creation of ROs	

EVER-FUN –030	EVEREST GENERAL REQUIREMENTS	Mandatory	OUTCOME
Requirement	EVER-EST e-infrastructure shall enable and facilitate Earth Science Researchers by • Be based, as far as possible, on loosely coupled, distributed, interoperable services and tools, rather than monolithic applications; • Be extensible with enhanced or new tools (possibly domain-specific) from any developer, through the use of published standards and software development kits, software libraries, etc. It should be as easy as possible to make existing software and services (e.g. e-print repositories, portals and proprietary software) interoperate with the VRE; • Be open source and standards-compliant wherever possible. The licensing of the software should encourage and support improvements to the tools and development of new tools by the community. Intellectual property rights issues need to be investigated and addressed using easy ways to provide attribution identifiers and product licenses; • Support tailoring of the environment by individuals or groups to reflect their domain interests and personal preferences; • Support the delegation of routine tasks to intelligent personal agents where the means to realise these exists, e.g. by incorporation into workflow processes;		Met – Taverna workflows help users exploit interoperable tools and services. Processing via virtual machines may provide further application but not tested fully by NHP VRC. Not tested Partially met - tools developed using open source software but IPR-controlled publishing issues not addressed. Not met, not tested Met but only partially tested via sharing of ROs with colleagues
Source	Proposal		
Verification Method	Verification Methods: tested through use of system and creation of ROs		

3.2 EVER-EST Services User Requirements

3.2.1 EVER-EST e – Common Services Requirements

EVER-COM –010	User Authentication and Session Management	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall allow Earth Science data users for: • Authentication and Authorisation: have a reliable user authentication and authorization accessing the VRE; • Federation: enable federated identity management; • User management • Session management and accounting of resources. • Definition of VRE sub-communities (for authorization to access project restricted resources.		Met. VRC administrators can authorise users Met – single sign-in for systems used by EVEREST (VRE, ROHUB, Seafile etc.) Not tested Not tested Met – users can be assigned different levels of access
Source	Proposal, VRCs (all)		
Verification Method	Verification Methods: tested through use of system and creation of ROs		

3.2.2 EVER-EST e – Collaboration Services Requirements

EVER-COL-010	Community Building and Peer Reviews	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall allow Earth Science data users to foster community building by : • Integration, open access and sharing of knowledge, free licensed software tools, open data and scientific products; • enable dissemination of research objects and products in accordance to specific rights/authorization; • have an appropriate dissemination interface towards the general public/end-users, including a mechanism to provide feedback and share results that could be released on a staged approach for peer review and evaluation purpose; • enable VRCs community directory to be linked to external social networking sites (e.g. LinkedIn, Research-Gate) supporting exchange of information.		Met Met Met Not met
Source	Proposals, VRE (All),		
Verification Method	Verification Methods: tested through use of system and sharing		

EVER-COL-020	EVER_EST e- connect tools	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall foster collaboration and sharing of scientific results allowing to centralize tools for collaboration management either integrating open sources collaborations tools or supporting the integration of COTS owned and installed by the same VRC to: • connect the scientific and the institutional communities allowing exchange of information; • support audio-visual webinar/conferencing with audio/video recording functionality and other asynchronous services as instant messenger/chat service and asynchronous mail and message board. • showing who is online and where and enable simple click-and-connect functions allowing improved instant communications that could be useful for arranging last minute meetings, clarifications, or other collaborative discussions; • conduct and present demonstrations using shared data and codes;		Met Partially met – chat. Mailing and message board available Partially met via chat service Partially met. Results and data are easily shared but users can not show what they

	• support space/forum for communities members to post and describe progressions and research status and to support peer-review, separated into themes and/projects to be used by different sub-communities; • capture and archiving of discussions including keyword tagging to relate conversations to specific topics which are searchable.	are doing live on the system To a very small degree met via mail box and task list. Not properly tested Partially met via mail box but not option for searchable keyword tagging
Source	Proposals, VRE Use Case (All),	
Verification Method	Verification Methods: tested through use of system	

EVER-COL-030	Content Management Tools	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall integrate a content management tool : • for Documentation repository with collaboration editing capability for collective document/article writing, including document versioning and provenance metadata capture; • Assuring secure storage and knowledge management with different level of user accessibility rights • Supporting automatic metadata insertion according to the VRE metadata model for document, records, cataloguing and provenance for long term preservation. • Providing a central source of information and knowledge management which might include: journal article library of relevant science/technology papers; available skill sets held in each organisation (providing access to additional information/help); CVs of all EVER-EST collaborators/EU partners providing information on relevant expertise and sources of help and possibly data, metadata and glossary of terms.		Met via ROs Met via ROs and Seafile. Security managed at user status level, not at RO level – so access issues partially met Met via automatic annotation Met via ROs
Source	Proposal, VRE (All)		

Verification Method	Verification Methods: Creation and sharing of ROs. Accessibility of SeaFile.	

3.2.3 EVER-EST e – Research Services Requirements

EVER-RE-010	Data and Resources Discovery	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall allow the Earth Science community to: • retrieve in near real time, and through web services a wide range of satellite and in situ data for the Earth Science community; • access EO mission datasets in a rapid and efficient way; • access data using catalogue metadata harvesting tools as well as automatic tools for browsing high volumes of data catalogue; • support the use of semantics and ontologies developed by the VRCs; • increase the availability and use of open data and open sources tools		Met Not tested Not tested Not tested Not tested
Source	VRE Use Cases (All),		
Verification Method	Verification Methods: tested through use of system		

EVER-RE-020	Data Access and Storage	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall support the following data access, storage and needs/ functionalities: • Data must be retrievable in near real time through web services; • Local storage of data must be possible, else web services shall be provided to the VRC;		Met – Seafile Met – Seafile Met – Seafile Met – Seafile

	• Simple, file system based local data archiving is preferred; • Store data in local environments, exporting them to the community using web services; • Provide research products as web services to the community; • Archive data and products using simple, file system based archiving; • Associate metadata to all data and products generated by the community within the VRE. Metadata standards must be used; • Use tools and procedures for license and PID/DOI attribution to data, research products or documents; • Use tools for referencing algorithms, products, and documents.	Partially met via VRE for those who have used the VRE Met – via ROs Met – via ROs Met – via data-type ROs Met – via ROs
Source	VRE Use Case (SM, SS)	
Verification Method	Verification Methods: tested through use of system and creation of ROs	

EVER-RE-030	**Validation Data Base**	**Highly Desirable**	**OUTCOME**
Requirement	The EVER-EST e-infrastructure shall support the integration of validation data base: • validation database represents a single source of observed impact data for validation purposes. • The database would include currently used sources of historical impact information such as the Nexis Lexus and Flood Forecasting Centre resources as well as data collected from additional tools developed in EVER-EST (e.g. twitter scraping, web scraping, other social media sites) • The database could be made publically available as a single source of impact data to be used for further validation and research.		Can be shared via Seafile. Partially tested Not met but outside immediate user requirements once system was being used Could be made public via data RO but not tested
Source	VRE Use Case (NH)		
Verification Method	Verification Methods: tested through		

EVER-RE-040	Data Sharing	Desirable	OUTCOME
Requirement	The EVER-EST e-infrastructure shall support data sharing: • to be provided by live data services, if available. • ready access to static saved data e.g. case-study based data to support more effectively building models using readily available 'sample data'. • easier access to large in-situ datasets (e.g. National Population Database, Met Office data) • access portals to database servers at partner organisations. • access platforms to geospatial data following common standards (one access point to many diverse data sets).		Not tested as not available to be tested Met and available via Seafile Not tested but can be met via Seafile Not met but not current requirement Not met but not current requirement
Source	VRE Use Case (NH)		
Verification Method	Verification Methods: tested through use of system and sharing of data		

EVER-RE-050	Research Data Generation and Preservation	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall support preservation and dissemination of Earth Science users generated data set enabling the following functionalities: • Annotation of the data set; • Archiving of the data set; • Cataloguing of the data set ; • Cross search of the data set; • Data and test mining; • Implementation of dictionaries and ontologies defined by the VRCs; • Information resources discovery; • Packaging; • Data Set Provenance metadata capturing		Met Met Met Met Not tested Not tested Not tested Not tested Met via RO checklist
Source	VRE Use Case (SM, All)		
Verification Method	Verification Methods: tested via RO		

EVER-RE-060	Application Management	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall support the following application management functionalities: • Executable building; • Executable management; • Process building; • Application Validation and verification; • Workflow management.		Partially met. Additional requirement for Virtual Machine should mean this canbe fully met in the future Not tested Not tested Not tested Met via RO
Source	VRE Use Case (SM, All)		
Verification Method	Verification Methods: tested through use of system and creation of ROs		

EVER-RE-70	Workflow Management	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall allow Earth Science data users to: • Link Research Objects and work-flows/thought-flows together; • Enable documentation of model development e.g. model history; • Facilitate capturing ideas at each stage of model development e.g. ideas/thoughts added to Research Object.		Met Met Met but not fully tested
Source	VRE Use Case (NH)		
Verification Method	Verification Methods: tested via RO generation		

EVER-RE-090	Visualisation and Modelling Tool	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall allow Earth Science data users to: • Deploy and use innovative, interactive and flexible visualisation tools for processing outputs and data; • Browse, query, and analyse processing outputs (normally images or maps in raster/vector format) for interpretation/validation purposes, through a web GIS interface; • Validate processing outputs using the web GIS interface and embedded analysis tools; • Integrated modeling tools as defined by the VRCs use cases.		Met Partially met – browsing and styling of outputs possible but deeper analysis isn't Partially met Not met
Source	Proposals, VRE Use Case (All)		
Verification Method	Verification Methods: tested through use of system and generation of ROs with visual outputs		

EVER-RE-110	Research Object Development Environment	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall provide Earth Science data users: • Collaborative development on Research Objects; • Case studies re-run with updated datasets, configurations and parameters; • Staged approach to development releases to the operational environment; • Enable partners to interact regardless of their pre-selected operating system platform of use (e.g. Linux and Windows).		Met Met Not tested as still in research phase Not tested
Source	VRE Use Case (NH, All)		
Verification Method	Verification Methods: tested through use of system and creation of ROs		

EVER-RE-120	Collaborative Computational Environment	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall allow Earth Science data users to: • Place workable scripts (R, Python etc) for the VRCs and others to use, develop, copy and amend as necessary; • Provision of modules of generic model building tools; • provide access to modules at different levels of complexity (from detailed process models to simple I/O structures); • Modules should be user generated and accessible through a visual library of tools that is easy for other users to pick from; • This facility should include proper code versioning and software archiving where necessary; • support multiple scripting languages;		Available via e-learning module Available via e-learning module Available via e-learning module Available via e-learning module Not tested Not tested
Source	VRE Use Case (LM, NH)		
Verification Method	Verification Methods: tested through use of system		

EVER-RE-140	Big Data Analytics Tool	Desirable	OUTCOME
Requirement	The EVER-EST e-infrastructure should support exploiting and integrating of Big Data analytics COTS and open source tools and environments, such as: • Analysis of social media; • Large-scale bid data processing environment (e.g. Hadoop, Apache Spark); • Other open source software like learning packages from the www (e.g. pluralsight).		Not met or required currently Not met and not required currently Available via e-learning module
Source	VRE Use Case (NH)		
Verification Method	Verification Methods: not tested		

EVER-RE-150	Preservation Requirements	Mandatory	OUTCOME

Requirement	EVEREST infrastructure shall provide the means for Earth Science data sources (either external data providers or VRCs as data generator) to: • enhance their preservation strategies by facilitating their systematic monitoring of end-user experiences and their requests for system enhancements. • enhance the data provider data records and relevant knowledge packaging and ingestion in the VRE in accordance with long term preservation best practices defined in D5.1; • manage the research object life cycle preservation process as defined in D.4.1. • allow preservation of scientific products and RO generated by the community for the duration of the project and in accordance with EVER-EST sustainability plan expected timeframe • allow preservation of scientific codes generated by the community for at least for the duration of the project and in accordance with EVER-EST sustainability plan expected timeframe.	Met Addressed by Data Provider Impact Assessment Met Met Met
Source	Proposal, D.4.1, D. 5.1 Preservation Requirements	
Verification Method	Verification Methods: tested through use of system and creation of ROs	

EVER-RE-160	**Research Objects Dissemination**	**Mandatory**	**OUTCOME**
Requirement	The EVER-EST e-infrastructure shall support the following data dissemination needs/functionalities: • dissemination of research object as specified in the use case and detailed in EVER-EST Data Management Plan. • Research objects should be disseminated as web services to the community • Metadata should be associated to all data and products generated by the community within the VRC and in-capsulated in the dissemination package.		Met Not tested Met
Source	Data Management Plan, D 4.1, D 5.1. Preservation Requirements		
Verification Method	Verification Methods: tested through use of system and creation of ROs		

3.2.4 EVER-EST e – Learning Services Requirements

EVER-EL-010	E-Learning	Mandatory	OUTCOME
Requirement	The EVER-EST e-infrastructure shall support: • The collaboration in the community to foster capacity building of less experienced scientists • Tools supporting the organization of online courses (e.g. Massive Open On Line Course) • Training on the platform usage during dedicated events; • User-driven tutorials; • Walk-throughs of visualisation tools and references for stored code.		Not tested but available via e-learning module Not tested Met via workshops held by VRC Not tested but met via e-learning module Not tested but met via e-learning module
Source	Proposal, VRE Use Case (NH)		
Verification Method	Verification Methods: tested via workshops held by VRC		

4 EVER-EST Performance and Operational Requirements success criteria

4.1 Success Criteria

A major challenge for EVER-EST e-research services integration is to pursue the following performance and operational requirements success criteria:

Success Criteria	Definition	Final Evaluation
Efficiency	There should be immediate, clear advantages for the scientists, either in terms of time required to carry out work, or as actual tasks, which could not be performed outside of the environment. This is the most important requirement, since the VRE is aiming to improve the efficiency of the research work, and an environment, which is perceived as decreasing such efficiency, will not be adopted by scientists	This requirement has been met. We have shown that the work of the NHP and specifically the HIM group could become much more efficient through use of the VRE for specific tasks such as workflow runs.
User friendliness	All operations should be intuitive and reflect the researcher's way of working. A steep learning curve may imply refusal to adopt the environment after a few attempts.	Whilst the UI is attractive, not all functions are intuitive. Particularly the process for adding resources to ROs. However, buttons display prompts which do aid the user when searching ROs, discovering EO data. Any less intuitive functions will simply require some initial simple training which would be expected in any new software system. This is not seen as a big issue to overcome.
Flexibility	Science normally implies the investigation of different lines of thought, sometimes very different. This is why working environments with rigid structures and procedures should be avoided.	Partially met. Flexibility is available but requires interaction with the technical team who are usually very responsive.
Easy integration of new software tools	The environment should provide interactive access to data processing tools (also from COTS), flexibility and simplicity to implement further tools (to maintain the state of the art).	Partially met. Any new tools require requests to the technical team who are usually very responsive.

4.2 Key Performance Indicators

The NHP is a relatively new initiative in the UK and has taken time to establish itself as the source of trusted hazard information for government and the responder communities. The science undertaken by the Hazard Impact Modelling (HIM) group is also in its infancy with current work focussing on refining the methods for developing hazard impact models for surface water flooding, wind and landslides, all with varying degrees of maturity and success. For surface water flooding, work is ongoing on how to turn this research trial phase into an operational phase with collaborators at the Flood Forecasting Centre (FFC), Centre for Ecology and Hydrology (CEH) and Health and Safety Executive (HSE). Due to the immaturity of the work, VRC members that have an immediate need to engage with use of the VRE to run workflows, as described under Use Case 1, is limited therefore the numbers of registered and active users on the Natural Hazards portal is naturally small. The number of ROs generated as part of this use case has also been low; the work has involved model testing and output as ROs which have not necessarily been saved. However, Use Case 2 recognised the value of creating 100s of bibliographic ROs to serve as an archive for the Daily Hazard Assessments (DHA) and provide an audit trail to historic decision making. It is anticipated that once ROs have been created for all DHAs to date (approximately 1 per day for the past 6 years), general NHP users of the system will start to rise. In addition, many of the NHP partner organisations are scientific institutions. There is therefore value in encouraging more general use of the VRE, and particularly the concept of ROs, as a way of encapsulating scientific results and enabling early publication of scientific results ROs with DOIs.