



# ever-est

## D3.1

# Use Cases Description and User Needs Document

<b>Workpackage</b>	<b>WP 3</b>	<b>VRE User Cases</b>
<b>Task</b>	3.1	VRE Detailed Definition of Use Cases
<b>Author (s)</b>	WP3	ESA, NERC, INGV, ISMAR, SatCen
<b>Reviewer (s)</b>	Cristiano Silvagni Hazel Napier Stefano Salvi	ESA NERC INGV
<b>Approver (s)</b>	Rosemarie Leone Cristiano Silvagni	ESA ESA
<b>Authorizer</b>	Mirko Albani	ESA
<b>Document ID</b>	EVER-EST DEL WP3-D3.1	
<b>Dissemination Level</b>	Public	
<b>Status</b>	Approved by the EC	
<b>Version</b>	1.2	
<b>Date of issue</b>	14 December 2018	



## Document Log

Date	Author	Changes	Version	Status
29/03/2016	ESA, NERC, INGV, ISMAR, SatCen		1.0	Draft to be approved by the EC
14/10/2016	ESA, NERC, INGV, ISMAR, SatCen	Update of document initiated after interim review	1.1	Draft
18/10/2016	ESA, NERC, INGV, ISMAR, SatCen	Update of document after formal internal review	1.1	Draft to be approved by the EC
14/12/2018	Cristiano Silvagni	Document status	1.2	Approved by the EC



## Table of contents

<b>1</b>	<b>Introduction.....</b>	<b>9</b>
1.1	Document scope.....	12
1.2	Relations with other EVER-EST work packages.....	12
1.3	Use Cases, user scenarios and user requirements specification .....	13
1.3.1	User scenarios specification .....	13
1.3.2	User requirements specification.....	14
1.4	Earth Science virtual research communities and users .....	14
<b>2</b>	<b>Sea Monitoring VRC .....</b>	<b>16</b>
2.1	Overview of the marine strategy framework directive .....	16
2.2	MSFD needs and challenges .....	18
2.3	Sea monitoring VRC user operational scenarios .....	19
2.4	Sea monitoring VRC service requirements .....	19
2.5	SEA monitoring VRC use case selection criteria .....	21
2.6	Sea monitoring VRC use cases description .....	22
2.7	Sea monitoring VRC use cases data needs .....	23
2.8	Sea monitoring use case scenario for research objects .....	23
2.9	Sea monitoring VRC user actions specification .....	24
<b>3</b>	<b>Geohazard Supersites VRC.....</b>	<b>25</b>
3.1	Geohazard Supersites VRC user requirements .....	26
3.2	Geohazard Supersites VRC service requirements .....	27
3.3	Geohazard Supersites VRC use cases .....	29
3.4	Geohazard Supersites VRC Context framework description .....	29
3.5	Geohazard Supersites VRC use case selection criteria and methodology.....	32
3.6	Geohazard Supersites VRC use case data needs .....	33
3.7	Geohazard Supersites VRC use case scenario for research object modeling .....	33
3.8	Geohazard Supersites VRC selected use case description .....	35
3.8.1	Volcano deformation mapping use case .....	35
3.8.2	Volcano source modeling Use Case .....	37
3.8.3	Retrieval of ash and SO <sub>2</sub> contained in volcanic clouds use case .....	39
3.8.4	Water vapour content use case .....	40
3.9	Geohazard Supersites VRC user actions Specification .....	41
<b>4</b>	<b>Land Monitoring VRC .....</b>	<b>42</b>
4.1	Land monitoring VRC user requirements .....	42
4.2	Land monitoring VRC context framework description .....	43
4.3	Land monitoring VRC use case selection criteria and methodology.....	43
4.4	Land monitoring VRC selected use cases description.....	45



4.5	Land monitoring VRC data needs.....	47
4.6	Land monitoring VRC use case scenario for research object modeling .....	47
4.7	Land monitoring VRC use case validation methodology and success criteria .....	48
4.8	Land monitoring VRC user actions specification .....	49
5	Natural Hazard Virtual Research Communities .....	50
5.1	Natural Hazard VRC user requirements .....	50
5.2	Natural Hazard VRC context framework description .....	53
5.3	Natural Hazard VRC use case selection criteria and methodology.....	55
5.4	Natural Hazard VRC selected use case descriptions .....	56
5.5	Natural Hazard VRC use case data needs .....	56
5.6	Natural Hazard VRC use case scenario for research object modelling.....	58
5.7	Natural Hazard VRC use case scenario general description .....	58
5.8	Natural Hazard VRC user actions specification .....	59
6	Earth Science VRE User Requirements.....	60
6.1	EVER-EST infrastructure key objectives.....	60
6.2	EVER-EST services user requirements .....	62
6.2.1	EVER-EST common services requirements .....	63
6.2.2	EVER-EST e – collaboration services requirements .....	64
6.2.3	EVER-EST e – research services requirements.....	66
6.2.4	EVER-EST e – learning services requirements .....	73
7	Compliance to the Smart Objectives and Key Performance Indicators.....	75
	Appendix A – Sea Monitoring VRC Use Cases Scenario .....	79
	Appendix B – Supersites VRC Use Case Scenario.....	87
	Appendix C – Land Monitoring VRC Use Case Scenario .....	122
	Appendix D – Natural Hazards VRC Use Cases Scenario .....	130

## List of Figures

FIGURE 1 EARTH SCIENCE INFORMATION LIFECYCLE .....	11
FIGURE 2: WP DEPENDENCIES .....	12
FIGURE 3 MSFD IMPLEMENTATION SCHEDULE .....	17
FIGURE 4 MSFD IMPLEMENTATION SCHEDULE .....	18
FIGURE 5 LIST OF CRITERIA AND INDICATORS FOR THE DESCRIPTORS D1, D2 AND D6.....	22
FIGURE 6 SCHEMATIC DIAGRAMS ILLUSTRATING THE PRINCIPLE OF GROUND DEFORMATION ANALYSES ON ACTIVE VOLCANOES. ....	31
FIGURE 7 CHANGE DETECTION (IN RED, RIGHT IMAGE) ON TWO IMAGES ACQUIRED IN 2006 AND 2010 (SOURCE: LISA) .....	44
FIGURE 8 QUALITATIVE ANALYSIS IDENTIFYING THE CONSTRUCTION OF NEW BUILDINGS (SOURCE: LISA) .....	45
FIGURE 9: USE CASE WORKFLOW .....	46



## List of Tables

TABLE 1 USER SCENARIO TABLE TEMPLATE .....	13
TABLE 2 REQUIREMENT TABLE TEMPLATE .....	14
TABLE 3 PERMANENT SUPERSITES .....	26
TABLE 4 USE CASE SCENARIOS FOR RESEARCH OBJECT MODELING .....	35
TABLE 5 VOLCANO DEFORMATION MAPPING USE CASE.....	37
TABLE 6 VOLCANO SOURCE MODELING USE CASE .....	38
TABLE 7 RETRIEVAL OF ASH AND SO2 CONTAINED IN VOLCANIC CLOUDS USE CASE.....	40
TABLE 8 WATER VAPOUR CONTENT USE CASE .....	41
TABLE 9 SEA MONITORING VRC SMART OBJECTIVES .....	75
TABLE 10 GEOHAZARD SUPERSITES VRC SMART OBJECTIVES .....	76
TABLE 11 LAND MONITORING VRC SMART OBJECTIVES .....	76
TABLE 12 NATURAL HAZARD VRC SMART OBJECTIVES.....	77
TABLE 13 KEY PERFORMANCE INDICATORS.....	77



## Definitions and Acronyms

Acronym	Description
AFQM	Administrative Financial and Quality Manager
AIS	Automatic Identification System
ARIA	Advanced Rapid Imaging and Analysis
CB	EVER-EST Collaboration Board
CEOS	Committee for Earth Observation Satellites
CFSP	Common Foreign and Security Policy
COTS	Components Off-the-Shelf
DHA	Daily Hazard Assessment
DOI	Digital Object Identifier
DoW	Description of Work
EAB	EVER-EST External Advisory Board
EC	European Commission
EEA	European Economic Area
EU	European Union
FAIR	Findable, Accessible, Interoperable, Usable Data
FFC	Flood Forecasting Centre
GEP	Geohazard Exploitation Platform
GES	Good Environmental Status
GPS	Global Positioning Systems
GRD	Ground Range Detected Images
GSNL	Geohazard Supersites and Natural Laboratories Initiative
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
JRA	Joint Research Activities
MSFD	Marine Strategy Framework Directive
MoM	Minutes of Meeting
NA	Networking Activities
NHP	Natural Hazards Partnership
OCG	Open Geospatial Consortium
OLCI	Ocean and Land Colour Instrument



<b>PaaS</b>	Product as a Services
<b>PC</b>	Project Coordinator
<b>PID</b>	Persistent Identifier
<b>PM</b>	Project Month
<b>PQMP</b>	Project Quality and Management Plan
<b>RTDI</b>	Research, Technology Development and Innovation
<b>R&amp;D</b>	Research & Development
<b>SA</b>	Service Activities
<b>SAR</b>	Synthetic Aperture Radar
<b>SatCen</b>	European Union Satellite Centre
<b>SLC</b>	Single Look Complex
<b>SWF</b>	Surface Water Flooding
<b>TBC</b>	To Be Confirmed
<b>TechComm</b>	EVER-EST Technical Committee
<b>TC</b>	Technical Coordinator
<b>VOT</b>	Vehicle Overturning
<b>VRC</b>	Virtual Research Community
<b>VRE</b>	Virtual Research Environment
<b>WFS</b>	Web Feature Services
<b>WMS</b>	Web Map Services
<b>WP</b>	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, <a href="http://ec.europa.eu/research/participants/portal/desktop/en/home.html">http://ec.europa.eu/research/participants/portal/desktop/en/home.html</a>
[5]	EVER-EST Alfresco Platform, <a href="https://EVER-EST.eo-infrastructures.eu">https://EVER-EST.eo-infrastructures.eu</a>
[6]	EVER-EST Project Web Site, <a href="http://www.ever-est.eu">www.ever-est.eu</a>
[7]	D 4.1. Deliverable: Research Objects in Earth Science version 1.1



[8]	D. 5.1 Deliverable: EVER-EST VRE Infrastructure and Services Design version 1.1
[9]	D.6.4 Deliverable: EVER-EST VRE Data Management Plan

### Abstract

EVER-EST is a research and development project that will offer a framework based on advanced services to support each phase of the Earth Science Research and Information Lifecycle. The project follows a user-centric approach leveraging on the results of several FP7 projects which have produced a wealth of innovative and state-of-the-art technologies, systems and tools for e-collaboration, e-learning, e-research and long term data preservation. It will provide innovative services to Earth Science user communities for the generation, communication, cross-validation and sharing of knowledge and science outputs. This document describes EVER-EST pre-selected communities operational framework, their use cases, user scenarios, service requirements needs and challenges to enable innovation, interdisciplinary cross-fertilization and foster Findable, Accessible, Interoperable, Usable data management policies and best practices. User Requirements elicited from the latter communities and gathered by assessing various interdisciplinary user profiles in Earth Science research and operational services communities, will be translated into specifications for the services supported by EVER-EST Virtual Research Environment to ensure that the technical developments are well aligned with the needs of the VRE specific use cases, generic Earth Science user requirement and overall EVER-EST scope.





## 1 Introduction

Over recent decades Earth Science research and operational services communities, especially those involved in the global and regional environmental monitoring domains, have received significant stimulus thanks to the advancements in measuring techniques and equipment. Huge amounts of data about our Planet are now available and, if these were easily discoverable, accessible and properly exploited, preserved and shared, they would potentially represent a wealth of information for a whole spectrum of stakeholders: from scientists and researchers to the highest level of decision and policy makers. Within the Earth Science domain there are prime examples of research areas where the solving of issues around the exchange of data/information between communities are of critical importance. For example, climate change monitoring requires both synergy between the different Earth Science disciplines, and the strengthening of international cooperation with relevant initiatives and programs (e.g. ESA Climate Change Initiative - CCI or Copernicus).

On the one hand Earth Scientists need to easily discover, access and exchange reliable data, have access to suitable processing power and visualization and analytics tools, and then share with peer scientists and communities, for validation and reuse, their methods and approaches, observations, results and most of all the lessons they have learnt. On the other hand, Earth Science Data Providers must offer innovative data services and tools to release the potential and creativity of the teams of scientists actually working with the data. This is presently addressed by government-endorsed initiatives as EPOS, aiming to implement a coordinated infrastructure for Earth Science data and service provision (<https://www.epos-ip.org/>). It is expected that European infrastructures as EPOS will improve the Earth Science data access to the research communities but alone are not sufficient to exploit the full potential of the European Research Area. Further developments are needed to improve the efficiency, productivity, and impact of European researchers, e.g. by promotion of knowledge sharing (scientific methods and software), transnational collaboration, synergy with sectors of the civil society, public institutions and private sector. This cooperation should cover from dissemination activities and outreach of scientific results to the wider public, to the enhancement of services entailing any exchange of knowledge or technology between Earth Science communities and potential adopters, like civil protection institutes.

In recent years Virtual Research Environments (VRE) have emerged as a concrete technical solution to these needs. VRE e-infrastructures can facilitate collaboration amongst researchers and research teams by providing them with a more effective means of collaboratively collecting, manipulating and managing data, as well as working together for knowledge creation.

The EVER-EST project aims to develop a generic VRE, which will be tailored to the needs of, and validated by, the Earth Science community. To achieve this EVER-EST will provide earth scientists with the means to seamlessly manage both the data involved in their computationally intensive disciplines and the scientific methods applied in their observations and modelling, which lead to the specific results that need to be attributable, validated and shared within the community e.g. in the form of scholarly communications.

EVER-EST will be built on the results of several FP7 projects, which have produced a wealth of innovative, and state-of-the-art technologies, systems and tools in key related areas. EVER-EST will leverage previous work on Earth Science data catalogue federation, access, search and retrieval, data processing and long term data preservation. EVER-EST will also exploit the EPOS data services whenever possible, and will test solutions (as



the Research Objects) which could eventually be adopted by the EPOS infrastructure and Earth Science data and services provision infrastructure at later times. The data management capabilities will be augmented with the models, techniques and tools necessary for the preservation of scientific methods and their implementation in computational forms such as scientific workflows, which are increasingly used in the Earth Science domain. Central to this approach is the concept of Research Objects (ROs), semantically rich aggregations of resources that bring together data, methods and people in scientific investigations.

Research Objects enable the creation of digital artefacts that can encapsulate scientific knowledge and provide a mechanism for sharing and discovering assets of reusable research. The EVER-EST VRE shall implement features for sharing information between stakeholders, create and preserve complex processing workflows, where needed by the VRC.

The natural evolution of these technologies will lead to the design, implementation and deployment of the EVER-EST VRE as an innovative framework which enhances the ability of the Earth Science communities and, observational scientific disciplines in general, to interoperate and share knowledge and expertise. The proposed technology brings together, at the researchers' fingertips, key features and functionalities including:

- Mechanisms to preserve and share scientific workflows and link them to the data, products and models;
- Services for improving the communication among scientists;
- Services for the validation and reuse of scientific knowledge, methods, and best practices;
- Services for data exploration, visualization and exploitation;
- Access to underlying computational resources and processing infrastructure;
- Access to all the associated data related preserved knowledge.

The project will follow a user-centric approach. Four real Earth Science Virtual Research Communities (VRC) covering different research fields and applications will drive the implementation of the proposed system's services and capabilities. The EVER-EST project user community will include bio-marine researchers (Sea Monitoring VRC), the Common Foreign and Security Policy (CFSP) institutions and stakeholders (Land Monitoring VRC), the natural hazards forecasting system community (Natural Hazards VRC), and the seismic and volcanic hazard and Disaster Management communities (GEO Supersites VRC). The resulting requirements for data management (e.g. accessibility of heterogeneous data sources), preservation of data and data related workflows and user experiences, data exploitation and other important e-research services for communication, cross-validation, sharing of science products, will be addressed through the proposed solution. Each identified VRC represents a different research area and will be constantly involved in the project for the purposes of validating the EVER-EST infrastructure.

This user-centric approach will allow an assessment to be made of the capability of the proposed solution to satisfy the heterogeneous needs of a variety of Earth Science communities for more effective collaboration, and higher efficiency and creativity in research. Each of the identified VRCs involves different categories of users both within and outside the Earth Science domain (e.g. general public, policy and decision makers). This will allow the piloting and testing of the infrastructure in real-life scenarios, which may significantly differ from each other. Data providers and value adders will be constantly involved in this process during the project to maximize its scientific and societal return.

The "Earth Science Research and Information Lifecycle" can be defined as the continuous, iterative and on-going process used by scientists for conducting, validating and disseminating scientific knowledge. It can

undergo an unlimited number of iterations, which lead to the development of new and innovative ideas, concepts, techniques and technologies that ultimately benefit both science and society.



**Figure 1 Earth Science Information Lifecycle**

The lifecycle can be briefly summarized into four main phases that include different categories of stakeholders at each stage as represented in Figure 1:

1. Scientists access information (e.g. raw data or added value products generated by colleagues) and share results; this is reliant on researchers and data providers giving access to the data and related knowledge;
2. Shared results and information are analysed, interpretative models are generated and discussed with other colleagues (within the team and/or the wider community, which can include external stakeholders), and require the use of visualisation tools and data analytics;
3. Discussion leads to novel ideas and concepts which might need validation through further experimentation or data acquisition; requires access to additional data sets held by other data providers;
4. New results are validated and shared (together with the workflow and processes used to generate them) for further discussion; including dissemination to external stakeholders (e.g. general public, policy makers). This provides stimulus to new research bringing the process back to step 1.

EVER-EST will offer a framework based on advanced services, which will be delivered both at the e-infrastructure and domain-specific level, with the objective of supporting each described phase of the Earth Science Research and Information Lifecycle. It will provide innovative e-research services to Earth Science user communities for communication, cross-validation and the sharing of knowledge and science outputs.

## 1.1 Document scope

This document describes the EVER-EST pre-selected communities, their use cases, service requirements needs and challenges to enable innovation, interdisciplinary cross-fertilization and foster common FAIR data management policies and best practices. User Requirements elicited from the latter communities and gathered by assessing various interdisciplinary user profiles in Earth Science research and operational communities, will be translated into specifications for the services supported by EVER-EST Virtual Research Environment to ensure that the technical developments are well aligned with the needs of the VRC specific use cases, generic Earth Science user requirement and overall EVER-EST scope.

## 1.2 Relations with other EVER-EST work packages

This deliverable has direct relation to deliverable of WP2, WP4 WP5, WP6 respectively (refer to the links between the corresponding work packages in Figure 2: WP dependencies).

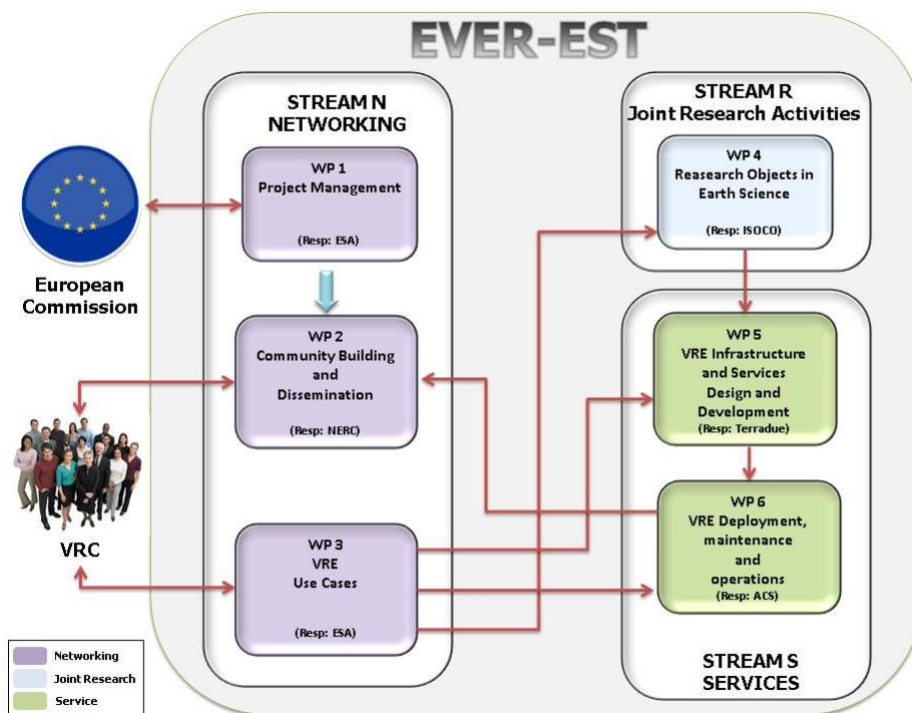


Figure 2: WP dependencies

A strong synergy between this deliverable and the other WPs ones is kept through the entire project lifecycle, as described in the Project Management Plan.

D3.1 analyses the needs, challenges and opportunities of the pre-selected user communities to boost innovative working practices for interdisciplinary cross-fertilization and synergy in the various fields of Earth Science and to enable common Findable, Accessible, Interoperable, Usable data management policies and guidelines.

From these use cases, user stories and learning-by-doing activities like the research object hackathons [AD7], a set of consolidated use cases, user stories and user services requirements, FAIR data management, quality preservation and curation requirements are gathered and further improved as addressed within WP4 deliverables. User requirements and detailed user actions, are, in turn, taken into account by WP5, as they are the drivers of the design and integration of the overall VRE platform components and its verification and validation.

At this stage of the project, VRC requirements and use cases are being deeply integrated in the architecture proposed by WP5. Up the end of the development phase, further requirements and user stories refinement will be discussed and addressed to ensure that the VRC's needs are fully addressed.

### 1.3 Use Cases, user scenarios and user requirements specification

EVER-EST Earth Science communities operational framework, use cases, user scenarios, needs and challenges are described in the following chapters for each of the pre-selected VRCs. To envision innovative way of working doable by improving the EVER-EST infrastructure components, continuous training-by-doing activities are being performed since the start of the project with the preselected VRCs acting as target audience and piloting the training approach development and communication activities for new users within the communities or new communities, described in WP2 deliverable. Attendance to conference and workshop and new communities responses to the EVER-EST VRE published results is, in turn, feeding the process of eliciting challenges, opportunities and requirements common to a broader spectrum of potential new Earth Science communities and driving EVER-EST VRE values propositions to new users and sustainability model.

#### 1.3.1 User scenarios specification

The following definition applies: A scenario is an end-to-end story describing a specific action/workflow performed within the VRC<sup>1</sup>. Each of the user actions description has the following template layout

Action ID	Action description	Dependencies	Challenge
US_VRC_01			

**Table 1 User scenario table template**

#### Legenda

- **Action ID:** progressive number to list all actions (e.g. US\_01, US\_02)
- **Action description:** A <user> performs an <action> to <achieve a goal >
- **Dependencies:** a pointer to all the possible dependencies and links related to an action. These include links to data, software, other infrastructure.
- **Challenge n°:** points to the proper challenge described in the use case which is addressed by an action.
- **Personas:** archetypical instance of an actor, at least one persona for each group of user profiles within the communities shall be addressed by the user stories.
- **Control Points:** human intervention specification, where and how the set of actions has a mandatory need for a human intervention and can't be automatized in any way. What is the action that the person is going to perform (e.g. validation), what are the options (e.g. Approved/Not Approved) and what are the actions that derive (e.g. proceed or go back to User Action N°).

---

<sup>1</sup> **Note:** The sum of all scenarios further refined during the entire project will reflect the entire VRC use cases in D.3.1 to be validated in accordance to the validation plan defined in D.3.2.

- **GUI Requirement:** functional behaviour requested by the user interface related to specific behaviours needed in relation to a specific action e.g. a specific popup window (e.g.)

VRCs user scenarios and user action specification, further developed from use cases, at this stage of the project, are detailed in the relevant Appendixes.

### 1.3.2 User requirements specification

User requirements have been elicited from use cases and user scenarios according to the main e-services categories preliminarily identified in AD[1]:

- Common Services;
- E-collaboration Services;
- E-research;
- Digital Information Services;
- E-learning Services ;
- Preservation and Curation Services.

Requirements are defined as illustrated in **Table 2** and are rated Mandatory (M – *shall* sentence), Highly Desirable (HD-*should* sentence) or Desirable (D- *may* sentence). The following user requirements verification methods will apply:

- Review of Design: identify an element of the design which satisfies a requirement;
- Inspection: verification performed by visually inspecting an element (e.g. source code);
- Demonstration: verification performed by executing a function of the service or software;
- Test: similar to demonstration, but requiring the definition of input data, test conditions and expected outputs.

EVER-XX-YYY	Requirement Title A	Rate
Requirement	Requirement Text	
	For readability, several sub requirements might be grouped in the same box. Sub requirements are identified by letters a) b) c)	
	All sub-requirements shall be verified individually	
Source	Traceability versus Proposals, VRE Use Case.	
Verification Method	Verification Methods: Inspection, Demonstration, Test or Review of Design	

**Table 2 Requirement table template**

## 1.4 Earth Science virtual research communities and users

The community of potential users of the Earth Science EVER-EST VRE is wide and heterogeneous and includes:

- Earth Science governmental panels, institutions, authorities and decision-makers;
- The scientific community at national and international level;
- The in situ data services providers;
- The EO data services providers (international space agencies and commercial ones);
- The end-users;
- The general public.



The scientific community involves multi-disciplinary scientists in all Earth Science disciplines and policy impact areas. Policy makers are responsible for defining the main Earth health indicators (e.g. Essential Climate Variable), disaster risk management actions and investments.

EVER-EST is an R&D project that follows a user-centric approach driven by four pre-selected communities:

- Sea Monitoring (SM VRC);
- Geo Hazard Supersites (SS VRC);
- Land Monitoring (LM VRC);
- Natural Hazards (NH VRC).

The pre-selected user communities represent a variety of interdisciplinary environments as well as of data and software and VRE service requirements will pilot the EVER-EST infrastructure. Hand-on training is provided to the VRCs since the start of the project to support and promote collaborative working across and between the individual communities and domains.





## 2 Sea Monitoring VRC

The community of potential users is wide and heterogeneous, including both scientists and national/international agencies and authorities (e.g. MPAs directors, technicians from regional agencies like ARPA in Italy, the technicians working for the Ministry of the Environment) dealing with the adoption of a better way of measuring the quality of the environment.

The scientific community involves multi-disciplinary scientists such as biologists, geologists, oceanographers and GIS experts. The scientific community has the main role of assessing the best criteria and indicators for defining the Good Environmental Status (GES) in their own sub regions, and implementing methods, protocols and tools for monitoring the GES descriptors.

The national agencies are in charge of leading the monitoring activities and they have the responsibility of applying the methods and protocols agreed by the scientific community.

The Ministry of Environment has the role of controlling, reporting and monitoring the phases and answer the questions raised by the European Commission related to the Marine Strategy Framework Directive (MSFD).

### 2.1 Overview of the marine strategy framework directive

The aim of the European Union's ambitious Marine Strategy Framework Directive is to protect more effectively the marine environment across Europe achieving the "Good Environmental Status (GES) of the EU's marine waters by 2020 and to protecting the resource base upon which marine-related economic and social activities depend. It is the first EU legislative instrument related to the protection of marine biodiversity, as it contains the explicit regulatory objective that biodiversity is maintained by 2020, as the cornerstone for achieving GES"

According to the MSFD, the environmental status is defined by 11 descriptors, and forms a proposed set of 29 associated criteria and 56 indicators that include biological, physic-chemical indicators as well as pressure indicators including hazardous substance, hydrological alterations, litter and noise and biological disturbance such as introduction of non-indigenous species (Cardoso et al., 2010; European Commission, 2010). Following the implementation schedule of the MSFD (**Figure 3**), EU Member States started to assess the environmental status of their marine waters in 2012. This assessment should be carried out in an integrative way, including measurement of many ecosystem components together with physicochemical parameters and elements of pollution.





**Figure 3 MSFD implementation schedule**

Implementation schedule of the MSFD including: a) The initial assessment of the current environmental status and the determination of GES, targets and associated indicators for national marine waters environmental to achieve GES by 2020 (by 15 July 2012); b) The establishment of a monitoring programme for the on going assessment and the regular update of targets (by 15 July 2014); c) The development of a programme of measures designed to achieve or maintain GES by 2020 (by 2015).

To help Member States interpret what GES means in practice, the Directive sets out, in Annex I, 11 qualitative descriptors which describe what the environment will look like when GES has been achieved. As these descriptors cover broad topics, the European Commission produced in 2010 a set of detailed criteria and indicators to help Member States determine what each descriptor means in practice and measure progress.

- **Descriptor 1.** Biodiversity is maintained;
- **Descriptor 2.** Non-indigenous species do not adversely alter the ecosystem;
- **Descriptor 3.** The population of commercial fish species is healthy;
- **Descriptor 4.** Elements of food webs ensure long-term abundance and reproduction;
- **Descriptor 5.** Eutrophication is minimised;
- **Descriptor 6.** The sea floor integrity ensures functioning of the ecosystem;
- **Descriptor 7.** Permanent alteration of hydrographical conditions does not adversely affect the ecosystem;
- **Descriptor 8.** Concentrations of contaminants give no effects;
- **Descriptor 9.** Contaminants in seafood are below safe levels;
- **Descriptor 10.** Marine litter does not cause harm;

- **Descriptor 11.** Introduction of energy (including underwater noise) does not adversely affect the ecosystem.

MSFD depends on measurable environmental targets and thus is linked to relevant indicators. Furthermore, the indicators and environmental targets will be related to the 11 descriptors and thus directly linked to the marine environmental pressures (Figure 3). Most importantly “to articulate quantitatively what GES looks like and/or set appropriate environmental targets it will be necessary to define for each of the criteria and, where appropriate, the indicators in COM Decision 2010/477/EU, environmental boundaries or thresholds above or below which GES is considered to have been met. . . . To that effect, a boundary between success and failure to achieve or maintain GES should be established. Thresholds/levels/limits in this sense represent that boundary between an acceptable and unacceptable status”.

Figure 4 illustrates an example of conceptual model showing how the criteria (and related indicators) for the 11 qualitative descriptors are linked, continuous lines indicate strong links, and dotted lines indicate weaker links. Dark green arrows indicate human pressures. Green boxes refer to pressure-related criteria, yellow boxes refer to state-related criteria, and orange boxes refer to impact-related criteria.

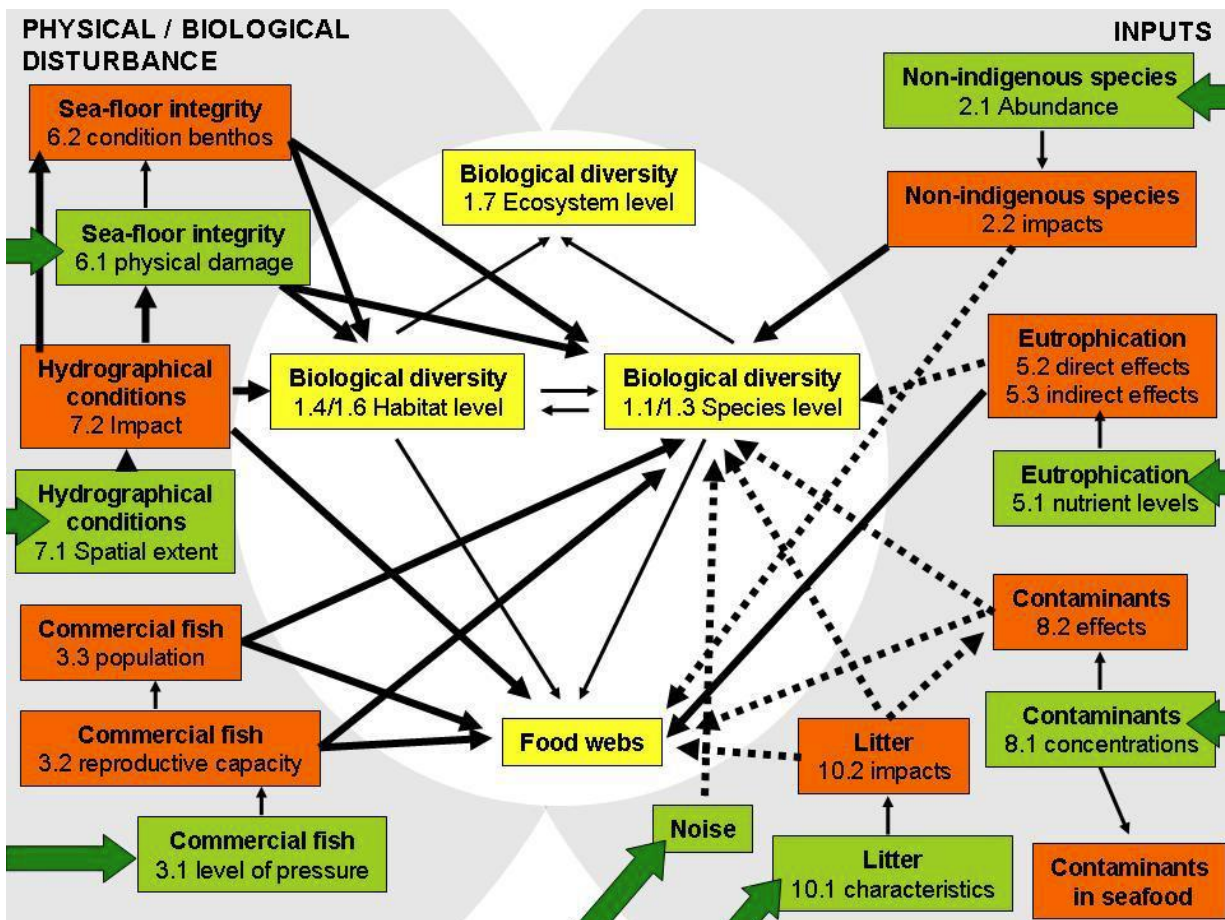


Figure 4 MSFD implementation schedule

## 2.2 MSFD needs and challenges

Within the MSFD rationalities, for similar issues there are different solutions and different Institutions involved (e.g. Universities, local entities, Research Centres)



In this perspective the MSFD needs to: 1) Strengthen regional cooperation and collaboration, governance (neighbour countries and Regional Sea Conventions) 2) Share Objectives (efficiency; integrated monitoring programme) and 3) increase knowledge exchange

In particular there is need for:

- New and/or applied knowledge to help define, assess and achieve GES and support MSFD implementation;
- Aggregation across geographic scales and comparison of results;
- Implement a modern and effective data and information sharing system between the EU (EEA) and RSCs ("WISE-Marine"), taking full advantage of the on going developments to improve accessibility and interoperability of marine data through the "Marine Knowledge 2020" initiative;
- Improve access to reliable, relevant, targeted and timely data & information.

In this perspective a VRE and Research Objects will give a great benefit to the community involved in the MSFD, including scientists, regional agencies dealing with monitoring programme and the Ministry of Environment managing at national level the MSFD legislation and attempts.

### 2.3 Sea monitoring VRC user operational scenarios

The users needs to fulfil the following general operational objectives:

- 1) Integration and homogenisation of available biodiversity data from different sources including literature:
  - Search for data related to MSFD descriptors and targets and provide links between platforms, easy access and discoveries;
  - Focus on specific key words for descriptor (Biodiversity) and MSFD targets for the discovery of spatial data.
- 2) Design and implementation of a repository for sharing analytical methods, protocols and algorithms for GES assessment and monitoring (focusing on Biodiversity D1 and Seafloor integrity D6);
- 3) Implementation of Research Objects including GES assessment, targets and monitoring protocols to be shared among the use case community.

### 2.4 Sea monitoring VRC service requirements

The following groups of services have been identified as needed by the Sea Monitoring VRC.

#### E-collaboration services

These services are important because the VRC community is heterogeneous and spread across different European countries. The community needs to communicate and share knowledge, ideas and protocols. There are great numbers of frontal meetings in different countries between different groups that could be reduced and optimized using these kind of services. Furthermore all monitoring campaigns could be better organized and managed. In this perspective the most useful services could be:

- Collaboration management;
- User registration;
- Asynchronous mail, message board;
- Resource sharing;



- Audio-visual conferencing and other synchronous services (e.g. chat).

#### E-research services

In general these services are needed to share data, protocols and workflows of monitoring activities. The community needs to compare and validate the results against previous data. Below the list of the most useful services at this stage:

- Application management;
- Executable building;
- Executable management;
- Knowledge discovery;
- Process building;
- Research resources discovery;
- Research resources management;
- Validation and verification;
- Visualization and interactive services;
- Workflow management.

#### E-learning services

The sea monitoring community normally is not involved in educational activities, however this services could be useful for increasing capacity building. Short courses could be organized concerning the monitoring programme and protocols to effectively share methods and workflow developed in the framework of MSFD between young scientists and less experienced technicians coming from regional agencies. It is not possible at this stage to define in details which services could be useful.

#### Digital Information services

This is one of the most important groups of services, as the Sea Monitoring community needs to effectively share data. The VRC needs to easily acquire and discover data, provide access to data, make available automatic tools for data processing and apply common semantic and ontologies. The VRC needs to make the data interoperable in order to ensure data access to the widest number of scientists and end users involved in the MSFD. Below a list of the useful services:

- Annotation;
- Archiving;
- Cataloguing;
- Content management;
- Cross search;
- Data and test mining;
- Data management;
- Dictionaries and ontologies;
- Information management;
- Information resources discovery;
- Packaging;
- Provenance.

#### Common Services

This group of services is needed by VRC given the heterogeneity of scientist involved to make the developed VRE efficient, user friendly, stable and interoperable. Below a list of useful services:

- Accounting;
- Alert and notification;



- Authentication;
- Authorisation;
- Digital Rights Management;
- Messaging;
- Metadata registry;
- Network management;
- Search;
- Security;
- Session management;
- User management.

## 2.5 SEA monitoring VRC use case selection criteria

Given the complexity of the MSFD focus will be given to the assessment of the GES and targets for 2 specific descriptors (D1 and D6) considering the indicators and criteria described in **Figure 5**.

These specific descriptors were selected because they encompass different aspects of biodiversity and the maximum number of possible indicators for the assessment of GES. Furthermore, assuming that the goal of the MSFD is to be an “instrument related to the protection of marine biodiversity” it was agreed to give the priority to D1 and D6 as first step for involving the maximum number of users within the VRE community. These descriptors are very interdisciplinary, and they require different kind of data from various and heterogeneous repositories and can be a good test for validating the infrastructure and for starting working with Research Objects. Focus will be given to 3 pilots for implementing a shared monitoring program trying to include key sites for GES assessment in terms of data requirements, tools and methodologies to be described.

DESCRIPTOR	CRITERIA	INDICATOR
1: Biological diversity	1.1 Species distribution	1.1.1 Distributional range
		1.1.2 Distributional pattern within the latter
		1.1.3 Area covered by the species (for sessile/benthic species)
	1.2 Population size	1.2.1 Population abundance and/or biomass
	1.3 Population condition	1.3.1 Population demographic characteristics
		1.3.2 Population genetic structure
	1.4 Habitat distribution	1.4.1 Distributional range
		1.4.2 Distributional pattern
	1.5 Habitat extent	1.5.1 Habitat area
		1.5.2 Habitat volume, where relevant
	1.6 Habitat condition	1.6.1 Condition of the typical species and communities
		1.6.2 Relative abundance and/or biomass, as appropriate
		1.6.3 Physical, hydrological and chemical conditions
	1.7 Ecosystem structure	1.7.1 Composition and relative proportions of ecosystem components (habitats, species)





2: Non-indigenous species	2.1 Abundance and state of non-indigenous species, in particular invasive species	2.1.1 Trends in abundance, temporal occurrence and spatial distribution of non-indigenous species
	2.2 Environmental impact of invasive non-indigenous sp.	2.2.1 Ratio between invasive non-indigenous species and native species
		2.2.2 Impacts of non-indigenous invasive species at the level of species, habitats and ecosystem
6: Seafloor integrity	6.1 Physical damage, having regard to substrate characteristics	6.1.1 Type, abundance, biomass and areal extent of relevant biogenic substrate
		6.1.2 Extent of the seabed significantly affected by human activities for the different substrate types
	6.2 Condition of benthic community	6.2.1 Presence of particularly sensitive and/or tolerant species
		6.2.2 Multi-metric indices assessing benthic community condition and functionality, such as species diversity and richness, proportion of opportunistic to sensitive species
		6.2.3 Proportion of biomass or number of individuals in the macrobenthos above specified length/size
		6.2.4 Parameters describing the characteristics of the size spectrum of the benthic community

**Figure 5 List of criteria and indicators for the descriptors D1, D2 and D6**

## 2.6 Sea monitoring VRC use cases description

Use Case 1. *Posidonia* meadows along Apulian coast: assessment of biological diversity focusing on habitat perspective considering the *Posidonia oceanica* meadows. All criteria and indicators from D1 and D6 will be considered. In this example the focus will be on the habitat approach producing thematic maps of sea grass communities and their regression from remote sensing data and historical map.

Use Case 2. Deep-sea corals along the Apulian continental slope: assessment of biological diversity focusing on habitat perspective. In particular the habitat extent will be considered and measured along with the distribution and condition found in specific sites of deep sea cold water corals considered by the MSFD as targets for monitoring programs. For this example focus will be on the habitat approach using habitat suitability model for assessing indicators and criteria.

Use Case 3. Jellyfish abundance in the Italian water: assessment of biological diversity focusing on the presence, outbreaks and distribution of indigenous and non-indigenous gelatinous organisms along the Italian seas, highlighting the matches with the environmental variables such as temperature, chlorophyll, nutrients, currents, wind, bathymetry and salinity. Moreover focus will be aimed at the impact of gelatinous blooms on fisheries. Criteria and indicators from D1 and D2 will be considered. For this example the Citizen Science data collected from 2009 by the University of Salento and the environmental data from the EMODNET and ESA will be used to realize: a) models of distribution and dispersions of gelatinous organisms and b) models assessing the economic impact on fisheries combining the information about the probability of jellyfish blooms with fishery landings and anthropogenic pressures (fish effort, fishing mortality, predator-prey interactions, migration fluxes and biological invasion).



## 2.7 Sea monitoring VRC use cases data needs

### Use Case 1. D1 and D6 for *Posidonia* meadows along the Apulian coast

- Seafloor bathymetry and derived environmental variables, habitat extent and biological related data available through COCONET repositories and EMODNET;
- Sea surface transparency, chlorophyll concentration, temperature and salinity available through ESA SAR imagery;
- Automatic Identification System (AIS) data;
- Chemical-physical data at bottom available through COCONET repositories and EMODNET.

### Use Case 2: D1 and D6 for cold water coral in deep water offshore Bari

- Seafloor bathymetry and reflectivity and derived environmental variables, habitat occurrences and biological related data available through COCONET repositories and EMODNET;
- Chemical-physical data at bottom available through COCONET repositories and EMODNET;
- Hydrodynamic condition at bottom, such as current velocities and direction available through COCONET, EMODNET and CNR-ISMAR local repository.

### Use Case 3: D1, D2 and D4 for jellyfish in the Med Sea

- Jelly fish occurrences data available through COCONET;
- Chlorophyll concentration, temperature and salinity, currents velocity and direction, Sea-surface height (SSH) available through ESA;
- Chemical-physical data (wind, nutrients and bathymetry, dissolved oxygen) at bottom available through COCONET repositories and EMODNET;
- Fish Effort available through Automatic Identification System (AIS);
- Fishing mortality, predator-prey interactions, migration fluxes and biological invasion.

## 2.8 Sea monitoring use case scenario for research objects

The Sea monitoring use cases will focus on the assessment of GES for D1, D2 and D6 for the following 3 cases:

### Use Case 1: Deep sea white corals: D1 Biological diversity- Criteria 1.4 Habitat extent – Indicator 1.5 Habitat area.

To derive the indicator 1.5 in deep-sea environment habitat suitability models needs to be implemented. The user needs to query the repositories storing multi beam bathymetry, seafloor reflectivity, hydrodynamic model and coral occurrence observations. The queries could be done through web portals within the VRE. The user needs raster data (ESRI GRID or .xyz) for seafloor bathymetry, backscatter and hydrodynamic models, vector data (.shp file) for the coral occurrences. The user will derive environmental predictor variables (EGVs = Eco geographic variables) such as slope, aspect, roughness using ArcGIS software or open sources GIS with dedicated algorithm. The contribution of each environmental variable together with the hydrodynamic component to the species distribution needs to be explained. To this aim an appropriate statistical methods for modelling needs to be chosen (such as ENFA, MAXENT or GLM) using scripts developed for the model. The final result is a continuous map of potential species or habitat useful to derive the habitat area.

### Use Case 2: D1 and D6 for *Posidonia* meadows along the Apulian coast: D1 Biological diversity- Criteria 1.6 Habitat condition – Indicator 1.6.1 Condition of typical species and communities.

To derive the indicator “1.6. Habitat condition”, an improved method needs to be implemented that is able to find out more about the distribution of *Posidonia oceanica* meadows in threatened areas and its spatial patterns in relation with the intensity of impacts and processes at various scales. The user needs to query the repositories storing multi beam bathymetry, seafloor reflectivity, historical map, SAR imagery and



environmental variables. The queries could be done through web portals within the VRE. The user needs raster data (ESRI GRID or .xyz) for seafloor bathymetry, backscatter, SAR imagery and vector data (.shp file) for the historical maps.

The user will derive “1.6. Habitat condition” through the integrative analysis of historical maps and SAR imagery, such as regression areas and classified cover density, using ArcGIS software or open sources GIS with dedicated algorithms (Diff). In addition the contribution of each environmental variable together with the human use must be explained, such as anchoring and fishing activity, in order to better define the distribution and condition of this habitat. To do this, appropriate statistical methods for modelling shall be chosen using scripts developed for different software like R. The final results will be a monitoring that should include procedures and systems, and which will provide regular and accurate information not only on habitat extent but also on the habitat condition and assessment.

### **Use Case 3: D1, D2 for Jellyfish in the Mediterranean Sea**

In order to derive the indicators of descriptors D1, D2, and D4, about the presence, outbreaks and distribution of Indigenous and non Indigenous gelatinous plankton along the Italian seas, a dedicated while limited Research Object shall be implemented, e.g. JF-RO) containing the following main items: the application capturing local info from citizens and researchers (JF-APP, to be evolved), the databases containing capture rough info (i.e. JFRDB to be maintained), the validated and geographically located rough data (i.e. JFVDB to be maintained) plus the relevant workflows aimed at capture and normalize data. Using this base the user needs raster data (ESRI GRID or .xyz) for seafloor bathymetry, sea surface temperature, sea surface salinity, chlorophyll, nutrients, geostrophic currents, and vector data (.shp file) of the jellyfish abundance (points) and track data to evaluate fishing effort, AIS data (lines). The users through the integrative analysis of environmental variables and jellyfish occurrence will be able to derive the distributional range and distributional patterns of different gelatinous plankton species along the Italian sectors of the Mediterranean Sea, and to derive the ratio between Non Indigenous species and Indigenous species. Finally, in few pilot areas, the users will be able to derive how environmental drivers affect the occurrence and abundance trend of functionally important selected species. In addition a model assesses the economic impact on fisheries combining the information about the probability of jellyfish blooms with fishery effort.

## **2.9 Sea monitoring VRC user actions specification**

Sea Monitoring user scenarios and user action specification, further developed from the above use cases, at this stage of the project, are detailed in the Appendix A.





### 3 **Geohazard Supersites VRC**

The [Geohazard Supersites and Natural Laboratories](#) initiative (GSNL) was launched in GEO in 2010 as a voluntary international partnership aiming to promote better scientific knowledge in the field of Geohazards, focusing on Seismic and Volcanic Hazards.

The GSNL goals are pursued providing open, easier and more complete access to a variety of space and ground-based data needed for geohazard assessment, focusing over selected, high risk areas of the world: the Supersites and Natural Laboratories.

On these focus areas a joint effort between the data providers and the scientific community is carried out. The space agencies comprised in the Committee for Earth Observations Satellites (CEOS) provide satellite imagery at no cost for scientific use, while national monitoring agencies commit to provide open access to ground based geohazard monitoring data. For each Supersite the partners also commit to facilitate the data access through standard web services and platform interoperability. Eventually, the global scientific community is enabled to exploit this continuously updated, large amount of data to generate new knowledge on the hazardous phenomena.

The benefits expected from the initiative are not only related to the production of new science, but also to the increase of knowledge transfer and capacity building, and to the promotion of a more efficient use of information resources (data, models, procedures, research products, etc.).

Moreover, as specified in the [Disaster SBA Task DI-01-C2](#), the goal of GSNL is also to provide a direct societal benefit: *"... The scientific information about geological disasters is the first element of the end-to-end approach to disaster management. This information will be openly available in timely manner to local governments for risk assessment ..."*

The general Geohazard Supersite concept is implemented using three instruments: the Permanent Supersites, which aim to promote new science to support seismic and volcanic hazard assessment for risk prevention, the Event Supersites, which have a limited duration and are dedicated to intensive scientific research on specific eruptions or earthquakes, and the Natural Laboratories which are instead larger regions of the world in which there is a high concentration of seismic and volcanic hazards (and other hazards as well), and where vulnerability, exposure and value are particularly high.

The following table lists the active seven Permanent Supersites, which were established between 2009 and 2015; new proposals are evaluated each year.

Permanent Supersite	Point of Contact	Point of Contact & Institution	Date Established
Hawaiian volcanoes	Michael Poland	U.S. Geological Survey, Hawaiian Volcano Observatory, USA	October 2012
Icelandic volcanoes	Freysteinn Sigmundsson	Institute of Earth Sciences, University of Iceland	November 2013
Mt.Etna volcano	Giuseppe Puglisi	Istituto Nazionale di Geofisica e Vulcanologia, Catania, Italy	April 2013
Campi Flegrei/Vesuvius	Sven Borgstrom	Istituto Nazionale di Geofisica e Vulcanologia,	April 2013



volcano		Naples, Italy	
Marmara Sea Region	Semih Ergintav	Kandilli Observatory and Earthquake Research Institute, Istanbul, Turkey	April 2013
Ecuadorian volcanoes	Patricia A. Mothes	Instituto Geofísico, Escuela Politécnica Nacional, Quito, Ecuador	October 2014
Taupo volcanic zone	Nico Fournier	GNS Science, New Zealand	October 2014

**Table 3 Permanent Supersites**

An initial team of scientists and institutions proposes each Supersite, including those research institutions, which maintain ground-monitoring networks, and academic institutions. After assessment and approval (this step involves the commitment by the space agencies to provide EO data) the Supersite is officially established, and a community of several scientists is built, who periodically need to access the in situ and EO data to carry out their studies.

The community is open, and new scientists can join it at any time. Involvement within the community implies the acceptance of the Supersite proposal objectives and data policy (the latter is specific for each Supersite). Belonging to a Supersite community allows access to the available data (EO and in situ) and to the scientific products generated by the community for that Supersite, but not for other Supersites. Other resources, as processing tools, are instead common to all Supersite communities.

Access to the data is regulated through specific data policies.

For EO data, there are two categories: data which are already fully open to all (as the ESA Sentinel data) for which the Supersite community in some cases can request additional specific acquisition plans, and limited access EO data, normally pertaining to missions which involve a commercial partner. For the latter (e.g. ASI COSMO-SkyMed, DLR TerraSAR X, CSA Radarsat 2) each scientist in the community must sign a license before being authorized to access the data.

For in situ data there is presently no license (there may be in the future) but the data policy may contain limits to the access, for example allowing access only to specific subsets of an entire catalogue, or to specific temporal periods, or to data produced before a certain limiting date.

Research products should be openly disseminated within the community, not just as publications, but also as digital files. This could be done in certain cases (e.g. during an emergency) even before publication. To provide a way to protect the Intellectual Property Rights of the data or scientific product owner, it is necessary to attribute a DOI to each item, and possibly to provide also a license (e.g. Creative Commons type), specifying the conditions under which the data or products can be used.

### 3.1 Geohazard Supersites VRC user requirements

The advantages which could be provided by a Virtual Research Environment to the overall Geohazard Supersite community are clear: in general terms these are the provision of services to improve collaboration and to make the scientist's work more effective. However, the actual success of the EVER-EST VRE will depend strongly on the capacity of the planned architecture and services to satisfy the following general requirements in order of importance:

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, an environment which is perceived as decreasing such efficiency, will not be adopted by scientists.

2. **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.
3. **Flexibility:** science normally implies the investigation of different lines of thought, sometimes very different. This is why working environments with rigid structures, procedures or interfaces should be avoided.
4. **Authorization and IPR protection:** the VRC needs an authorization mechanism to access data, products or tools, and to maintain IPRs of research results.
5. **Processing 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).
6. **Data access:** rapid and simple data discovery and access from a variety of sources. This is sometimes a very time consuming task for scientists, and efficiency of data access is considered important by most.

### 3.2 Geohazard Supersites VRC service requirements

In the following section the service requirements structured according to the categories listed in chapter 1.3 are described.

#### E-collaboration services

For the Supersite community these services are important, since typically the scientists are spread over different countries and may have different research capacities. Optimal service requirements are:

- Collective document/article writing, with document versioning;
- There should be the possibility to define specific VRE sub-environments within the general Supersite VRC (i.e. for authorization to access resources). In fact, the community is composed of sub-communities, e.g. there could be scientists interested to a Supersite but not to others, or scientists within the same Supersite involved in a funded project, in which they have specific IPR agreements which are not extended to all scientists in the same Supersite;
- Project planning and management tools are needed for the general Supersite network management, but may be used by sub communities of single Supersites;
- Community directory, linked to external social networking sites (LinkedIn, ResearchGate)
- The community will need synchronous and asynchronous communication tools (Chat, Teleconference, Announcements), with audio/video recording services for meetings;
- A dissemination interface towards the public/end-users is needed. Dissemination of Supersite products should be subject to specific authorization. A mechanism to provide feedbacks from the public on research results would be useful, also for evaluation purposes.

#### E-research services

The typical work procedure for a scientist in the Supersite community is to extract data from a repository, process them, validate the results against different data or past results, and share the results within the community. Depending on the objectives, the processing may be done using some COTS software (e.g. for satellite image processing), or in house software (more common for modelling and simulations). Since the in house software is normally very dynamically evolving, it is important to allow the scientists the capacity to easily test, run and modify their codes in the environment before publishing a stable update.

- InSAR data processing: there should be a baseline toolset mostly through COTS (SarScape is a probable option), with interactive GUI. This part is computationally intensive, and ad hoc technical specification will be given;
- Optical data processing: again basic toolset is based on COTS (ENVI is a probable option), plus scientific software, with interactive GUI;



- Modelling tools: the community will use or implement volcanic modelling software for ground deformation and volcanic clouds, based on EO and in situ data. This needs also interactive GUI. A finite element COTS may be needed (MARC or ABAQUS);
- Capacity to execute workflows is mostly useful for image data processing, in part for modelling, less so for validation tasks since these are less repetitive. Presently the community does not use any tool for the formal description of workflows (as for instance Taverna), however it is expected that there could be benefits in the adoption of these tools in the VRC;
- Most processing outputs are images or maps (raster/vector products) and for interpretation/validation purposes, they must be browsed, queried, and analysed through a web GIS interface;
- Specific validation tools are needed. They should be integrated into the web GIS interface.

### Digital information services

It is one of the most important groups of services for the Supersite VRE. The VRC should use data which is easily accessible through web services, although in some cases these services may not be available externally. Complete interoperability is mandatory, given the global scale of the VRC. The most relevant frameworks to which the data management in EVER-EST should be compliant or interoperable are: the EPOS framework at European level, and the [Common Framework on Earth-Observation Data](#)<sup>2</sup> from USGS and GEO-USA.

- Data must be retrievable in near real time through web services;
- Local storage of data must be possible, but in this case web services must be provided to the VRC;
- Research products should be also provided as web services to the community, either at VRE platform level or at local level;
- Simple, file system based local data archiving is preferred;
- Metadata should be associated to all data and products generated by the community within the VRE;
- Metadata standards will be provided;
- Data browsing, visualization, selection, should be done through friendly GIS interfaces;
- EO data needs are detailed in section 3.6;
- In situ data needs: GPS, levelling, epicentral locations, focal mechanisms, atmospheric radar, geochemical data, thermal imaging, etc;
- For Sentinel-1 data, access to the full archive is needed, to avoid the need of constant download from rolling archive;
- Tools and procedures for license and PID/DOI attribution to data, research products or documents are required;
- Full document management for each Supersite community and for the coordination of the Supersite network is needed;
- Tools for referencing algorithms, products, documents used in each Supersite, are needed;
- All data and products produced in the VRE should be managed in such a way to be searched and referenced by the GEOSS Geobrowser.

### E-learning services

The Supersite community does not normally carry out educational or training activities, however part of the collaboration in the community implies capacity building for less experienced scientists, or just brief courses regarding the use of codes or methods (e.g. the use of Research Objects, which are a new concept in Earth Science). Thus some e-learning services are required and will be specified later on.

---

<sup>2</sup> [Common Framework on Earth-Observation Data](#)



### 3.3 Geohazard Supersites VRC use cases

The main EVER-EST VRE user requirements are provided by the communities involved in the following volcano Supersites:

- Mt.Etna, Southern Italy;
- Campi Flegrei - Mt.Vesuvius, Central Italy.
- Iceland Supersite

The selected Supersites well represent the interests of the volcanological scientific community of potential VRE users. At the same time the role of INGV in the Supersite Scientific Committee will guarantee that also the seismological Supersite communities are able to verify the VRE requirements, once the main service identification is completed.

In this section use cases extracted from the various activities carried out by the scientific communities of the three mentioned Supersites are described. They are provided here as representative examples of such activities, and do not necessarily include all possible ways by which the scientists are expected to use the EVER-EST VRE.

Since the details of the use of the VRE for improving and making more effective the scientist's work and interactions will be discovered during the demonstration phase, it is expected that new use cases will be exploited to take advantage of the new capacities provided by the environment, as for instance the Research Object and workflow management.

### 3.4 Geohazard Supersites VRC Context framework description

The context framework for the Geohazard Supersite community has been described in 3.4. It is worth mentioning that given the dynamic character of the initiative, new Supersites are added every year, and different operational scenarios may come up with different Supersites. The various stakeholders involved (at different levels) in a Supersite are:

- The scientific community;
- The in situ data providers;
- The EO data providers;
- The end-users;
- The general public.

Within each category local (national or sub-national) and international or global stakeholders, with partially different interests or objectives can be distinguished. Two main categories of operations can be identified, those relevant to Event Supersites and those relevant to Permanent Supersites.

**Event Supersites** are areas of the world where a large earthquake or volcanic eruption occurred. Depending on magnitude of the phenomena, impact on human structures or environment, an Event Supersite is established by the GSNL community. The GSNL Chair then activates the data providers (in situ and satellite) to verify the willingness to provide monitoring data to the scientific community. Normally, while the CEOS space agencies are willing to provide image acquisition support to the Supersites, the in situ monitoring community is only providing, at best, a subset of the data from the local networks. This is due to the national security issues, which are inevitable for large events, and is often mandated by the civil protection or internal affairs agencies in charge of the emergency response.

The local scientists and monitoring agencies normally provide the satellite data requirements to the CEOS, sometime this task is carried out directly by the GSNL management to avoid delays, which could diminish the value of post-event satellite data. Pre-event archived data are also requested.

Normally the Event Supersite is active only for 3 to 12 months after the event, and the data and scientific products generated by the community are meant to support the detailed knowledge of the source and of the event effects, taking into account possible induced events during that period (e.g. large aftershocks).

For some of the scientific products (e.g. fault slip and location, magma source dimension and location, etc.) there are temporal constraints, since this information could lose immediate value for the responders after few days from the event. Some other products instead (e.g. ash plume monitoring, post-seismic ground deformation) should be constantly generated and updated during the event.

The scientific results generated during the emergency and used to support the response, should be provided to the end-users in an organised way. For most Supersites a main agency or institute is responsible for the interactions with the end-users. So, respecting the local procedures, the international scientific community involved in Event Supersite research, should coordinate with the local scientists to generate consensus reports providing a critical summary of all the research results. This procedure is necessary to avoid the diffusion of non-validated, incorrect information or even the misinterpretation or misuse of correct information by unprofessional media. To respect the intellectual property rights of the scientists who share their results even before publication, attribution of PID or DOI tags to the products must be guaranteed.

The use of a VRE is going to facilitate all activities carried out within the Event Supersites, improving the efficiency of data access, rapid product generation, easy collaboration for product validation and sharing, enabling a large community of stakeholders to work together to provide better and more timely response to disasters.

**Permanent Supersites** are areas of the world where seismic or volcanic hazard is high, and there are important scientific problems to address. Normally these areas are selected also on the basis of a high risk, implying the existence of considerable Exposure and Vulnerability. They are proposed by a wide community of scientists and data providers, and selected through a review process (see [Supersite selection and review procedures<sup>3</sup>](#)).

Contrary to Event Supersites, they are long-lasting initiatives, in which commitments are taken, by the CEOS and in situ data providers, to consistently provide data for many years.

One other goal of the Permanent Supersite is to provide a framework for the improvement of the data sharing infrastructures, or procedures. Supersites are also used to test innovative methods for the provision of open access to Earth Science data, for instance the four European Supersites are used during the development of the EPOS research infrastructure, to test its efficiency and suitability.

After a Permanent Supersite is officially established (by the GSNL Scientific Advisory Committee), the data start to be made available to the Supersite community. The community is open to scientists through registration.

### **Main activities in Permanent Supersites**

The main activities carried out in a Permanent Supersite are data production and scientific analysis of the phenomena and processes. These activities are routinely ongoing, sometimes (especially for volcanoes) triggered by some variations in the temporal evolution of some parameters. A typical case is for instance when geochemical measurements at the surface provide evidence of increased amounts or concentrations of gases (SO<sub>2</sub>, CO<sub>2</sub>, H<sub>2</sub>S) related to magma ascent in an open conduit. This is a sign of unrest of the volcano, which normally triggers further data analysis of, for instance, geodetic measurements from GPS and InSAR, or surface temperature mapping. The new data are then analysed and modelled to investigate parameters describing the evolution of the phenomena, as magma depth, size of magma chamber, strain accumulation at specific sites,

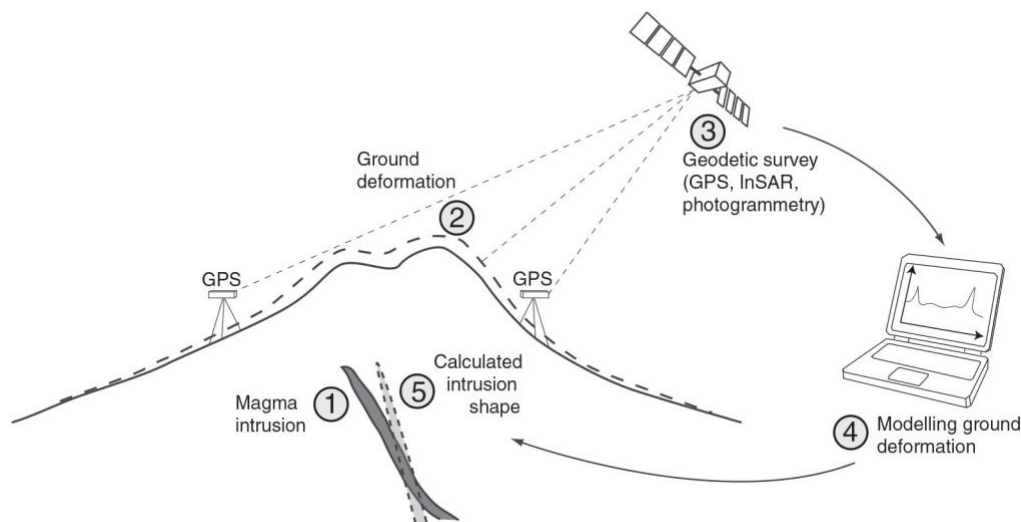
---

<sup>3</sup> [http://www.earthobservations.org/documents/gsnl/201510\\_Supersites\\_Selection\\_Review\\_Procedures.pdf](http://www.earthobservations.org/documents/gsnl/201510_Supersites_Selection_Review_Procedures.pdf)



etc., as shown in the scheme in **Figure 6** where the numbering gives the succession of the stages of the analyses. 1. Magma intrudes in volcano, feeding magma reservoir or forming a sheet intrusion (dark grey intrusion). 2. Magma intrusion triggers ground deformation at surface, leading to modified topography (dashed line). 3. Topography variation is measured by geodetic techniques (GPS, InSAR, Photogrammetry, etc.). 4. Geodetic data are compared with modelling of ground deformation due to various intrusion shapes. 5.

The best fit between models and ground deformation data provides a calculated intrusion shape (light grey dashed intrusion) responsible for the measured ground deformation. Nevertheless, the calculated intrusion shape is not a unique solution, and to validate the calculated model scientists need to compare their results with other datasets in a GIS interface.



**Figure 6 Schematic diagrams illustrating the principle of ground deformation analyses on active volcanoes.**

### Scientific collaboration in Permanent Supersites

At present there is little scientific coordination in some Permanent Supersites, and the scientists carry out their research independently. The main reasons are the lack of proper funding for collaboration, but also the lack of a practical way to collaborate and share results and tools, among scientists located in distant parts of the world. One other reason is the lack of established methods for early attribution of Intellectual Property (IP), which sometimes restrains the researchers from freely sharing their scientific results before publication. At present the IP attribution before publication is sometimes provided through the use of DOI or open licenses, as GNU or Creative Commons, but there is not established procedure or rule to guide the community through this task.

### Data and product management

Both data and scientific results are fundamental to estimate the hazard, but to reach this goal they should be discussed and evaluated by panels of scientists who should examine the different data, findings, models and uncertainties, and eventually reach consensus on some final pieces of information which could then be distributed to end-users. The consensus building and user dissemination tasks will certainly benefit from the collaboration tools provided by the EVER-EST VRE.



As previously mentioned, one of the main operational tasks carried out by the community is data browsing and processing. While satellite data is in most cases accessible through online catalogues, and sometimes also web services, only some in situ data types are similarly accessible (typically seismic, geodetic and geologic data). For the Supersites selected as pilot cases in EVER-EST, seismic web services are provided by INGV (<http://webservices.rm.ingv.it/>) for Italian earthquakes, but not yet for Iceland seismic data. For geodetic GPS data, web services are provided through a UNAVCO interface (<http://www.unavco.org/data/gps-gnss/data-access-methods>) but only for a small number of permanent stations.

Volcanological data for these Supersites are presently not provided as web services, but by mid-2016 the end of the three FP7 European Supersite projects should imply that at least some of these web services will be provided by the respective coordinating agencies.

### Processing services

Once data is accessed it is analysed by scientists through SW codes available to the single scientists, but not usable through common online platforms. Some of these codes are openly shared among scientists. Some other SWs are instead of commercial origin, typically those for satellite data analysis (ENVI, ERDAS, SarScape, GAMMA, etc.) or GIS (ESRI products).

There are some initiatives, as the Geohazard Exploitation Platform (GEP) by ESA, and the Advanced Rapid Imaging and Analysis (ARIA) platform by NASA/JPL, which are being established to provide processing services or processed data products to scientists. The GEP in particular is already used by part of the Supersite community to process SAR data.

The Supersite community processing needs normally fall into these categories: data processing tools (to extract geophysical parameters from EO and in situ data), modelling/analysis tools (used to model /analyse a given phenomena, as volcanic source, volcanic cloud ash density, etc.), validation tools (algorithms to be used during the validation of the processing results). While the validation tools implement simple statistical comparisons of numerical values, the modelling/analysis tools implement complicated algorithms and codes, which may also require particularly high computing power (e.g. for the Finite Element models). The same (to a lesser extent) occurs for the data processing tools.

Moreover, in the Supersite VRC the main focus is on the generation of new scientific results at the state of the art, which is a process involving constant evolution of data analysis methods, processing techniques, modeling methods and procedures. This implies that the scientific analysis modules during a timeframe of 1-2 years can evolve considerably, following the scientific advances in the field.

Since the EVER-EST VRE can be appealing to the community only if it provides state-of-the-art processing and analysis tools/services, it is extremely important that the environment architecture is designed in a way to allow the scientists to develop, implement, use, verify their models and algorithms in a continuous way.

More specifically, what is needed is in fact an environment very close to the researcher's workstation environment, in which the researcher can upload (possibly develop) his algorithms, carry out tests and experiments using the facilities of the platform, share the results with colleagues, validate the algorithms, and then when the codes are stable enough, publish the codes in a toolbox for the community to use.

Development of a too rigid infrastructure in which SW codes require important engineering efforts to be used by the community, is going to be a negative aspect if the EVER-EST VRE must be a lively, scientific-rich working environment able to attract many users.

## 3.5 Geohazard Supersites VRC use case selection criteria and methodology

Four typical use cases are presented, in order to exploit the whole spectrum of the VRE potential within the Supersite VRC. Use cases concerning the collaboration, IP management, data management, and in general those tasks, which are traversal among the four VRCs are not described. A list of the components and





respective subcomponents will be provided once a better idea of the architecture is provided. Moreover, the number of codes to implement and the progressive schedule of implementation depend on system design, and some iteration may be needed.

### 3.6 Geohazard Supersites VRC use case data needs

Information about the use cases data needs is reported in section 3.8 where each use case is treated as a whole.

### 3.7 Geohazard Supersites VRC use case scenario for research object modeling

The general Research Object content structure for the Geohazard Supersites VRC is described in the following table (see also Deliverable 4.1). Each of the four use cases described in the following sections can be described by a RO. This is expected to provide important benefits to the VRC.

Main Category Level	Sub Level 1	Sub Level 2	Notes
RO metadata			<p>Includes information about the version of the Research Object and changes made in relation to previous versions</p> <p>Includes information on how to reference and acknowledge this Research Object.</p> <p><b>Information</b> on authoring and scope of RO</p> <ul style="list-style-type: none"> <li>Authors</li> <li>Affiliations</li> <li>Origin</li> <li>Funding</li> <li>End-users</li> </ul> <p><b>Information</b> on acknowledgements</p> <p>Acknowledge contributions from: other ROs, workflows, people, projects, funding agencies, etc.</p> <p><b>Log book:</b></p> <p>One or more text files documenting the execution of the Use Case (actual log files from a specific SW execution, notes by the scientist, screenshots of specific WF steps, etc.)</p>
Workflows	Main WF	WF description	<p>A text file describing the WF. It should be including the assumptions, if any, related to the particular methods used in each WF.</p> <p>This reference could in some cases be provided by scientific publications (if they are detailed enough).</p>



		WF Graphical description	Simple WF flowchart in png or jpeg format
		WF digital files	May contain the actual processing configuration and parameters used in this specific WF run. E.g. a SarScape batch file (.sav).
	<b>Connected WFs (e.g. validation)</b>	WF description	As above
		WF Graphical description	As above
		WF digital files	As above
<b>Documentation</b>	<b>Scientific publications</b>	Used	Publications referencing the processing methods and/or the scientific reasoning and interpretation.  For example documents describing the methods, WFs, previous analysis of same data, previous investigations of same volcano or earthquake, etc.
		Produced	Publications of the specific WF results, as conference proc., journal articles, etc.
	<b>Manuals (SW, HW, DB)</b>	Only used	
<b>Data</b>	<b>Input</b>	Processed	Data processed by main WF.
		For validation	Data used for validation of main WF results
	<b>Output</b>	Intermediate, non-validated results	Includes authorization flags, IPR reference (licenses for reuse, DOI) of datasets, which should be shared.
		Final, validated	Includes authorization flags, IPR reference (licenses for reuse, DOI)
	<b>Comments on Use case data and results</b>		Document reporting, in a synthetic form, information, which is considered important on: data, results, conclusions, interpretation of results, and any other information useful for the reuse of the RO and WF.



			It may be similar to a publication abstract, and is mainly useful when the results of the particular scientific WF are not yet published.
<b>Processing components</b>	<b>Software</b>	Compilers	Normally this is commercial SW
		Scripts	Typically created by scientists (IPR management information needed)
		Data consumer SW	Normally this is commercial SW
	<b>Web processing services</b>		Used processing services offered by external resources.
	<b>Configuration parameters</b>		Configuration parameters relative to the SW used for the processing (e.g. allocated memory, working directories, etc.). Only in case the RO contains workflows.

**Table 4 Use Case Scenarios for Research Object Modeling**

### 3.8 Geohazard Supersites VRC selected use case description

The 4 examples of Supersite use cases are described in detail, in a step-by-step way, in the following sections.

1. Volcano Deformation Mapping: This Use Case deals with the process of investigating the ground deformation at an active volcano.
2. Volcano Source Modelling: This Use Case concerns the process of investigating the magmatic source of an active volcano.
3. Retrieval of ash and SO<sub>2</sub> contained in volcanic clouds: This Use Case concerns the process of assessing ash and SO<sub>2</sub> burden in the volcanic clouds generated during explosive eruptions from multispectral thermal infrared satellite data.
4. Water Vapour Content: This Use Case concerns the process of retrieving the atmospheric water vapour contained in a volcanic cloud over an active volcanic area.

#### 3.8.1 Volcano deformation mapping use case

<b>Number 1</b>	<b>Title:</b>
<b>Author:</b> S. Salvi, E. Trasatti	<b>Volcano Deformation Mapping</b>
<b>Description</b>	This Use Case concerns the process of investigating the ground deformation at an active volcano
<b>Triggers</b>	The generation of InSAR deformation maps is a routine procedure (monitoring) for an active volcanic area, and is performed every few months. In case of unrest (defined as a change in background seismicity rates, continuous GPS site velocities, geochemical parameters, etc.) the activity may be anticipated and deformation maps may be updated with higher frequency.



<p><b>Pre-Conditions</b></p>	<p><b>INPUT DATA:</b></p> <p>InSAR satellite data (SAR images periodically acquired using the same viewing geometry) of the volcanic area. They could be available through web services provided by the respective space agencies or data distributors (e.g. <a href="https://supersites.eoc.dlr.de/">https://supersites.eoc.dlr.de/</a>).</p> <p>Digital Elevation Model of the volcanic area is available, at a 90-30 m resolution (e.g. the ftp site of the SRTM mission, <a href="http://srtm.csi.cgiar.org">srtm.csi.cgiar.org</a>).</p> <p>(Optional) Local atmospheric parameters from ERA-Interim numerical weather model data produced by the European Centre for Medium-Range Weather Forecasts (ECMWF).</p> <p>User has access to continuous GPS site velocities and time series from an online repository (e.g. <a href="http://www.unavco.org/data/gps-gnss/data-access-methods/dai2/app/dai2.html#">http://www.unavco.org/data/gps-gnss/data-access-methods/dai2/app/dai2.html#</a>) in ascii/csv format.</p> <p>(Optional) User has access to vertical benchmark displacements from a leveling line. Same as above.</p> <p>User can display the 3D spatial patterns recent and NRT seismicity around the volcanic area (e.g., <a href="http://webservices.rm.ingv.it/">http://webservices.rm.ingv.it/</a>).</p> <p>User can display geological information pertaining the volcanic area (lithological map, structural map) from specific literature for the volcanic area, in graphical formats or ascii/xls data, or from online catalogues (<a href="http://geodb.ct.ingv.it:8080/geoexplorer/composer/">http://geodb.ct.ingv.it:8080/geoexplorer/composer/</a>).</p> <p>User can display other ancillary monitoring information (e.g. microgravity data, thermal data, geochemical data), same as above.</p> <p>Notes:</p> <ol style="list-style-type: none"> <li>1. Not all data could be available for each volcanic area or case;</li> </ol> <p><b>PROCESSING TOOLS:</b></p> <p>The SarScape and ENVI-IDL COTS are available for use by the operator through a graphic window.</p> <p>Validation tools in the EVER-EST map interface.</p>
<p><b>Post-Conditions</b></p>	<p>Existence of a repository for the processing results (ground deformation time series and maps).</p>
<p><b>Activities</b></p>	<p>The following procedure can be repeated for each InSAR data set pertaining to a specific satellite and orbit.</p> <ol style="list-style-type: none"> <li>1. User must select the appropriate InSAR and DEM data for the processing. To this aim the user needs to load and display on a web GIS interface the spatial and temporal distribution of the InSAR images, select them (temporally and spatially) and typically create a subset of a larger data set. The data subset must be in the format requested by SarScape.</li> <li>2. The user can also display in the same web GIS interface other data which are not input to the processing, as geological maps, seismicity patterns, GPS velocity vectors and values, previous InSAR deformation maps, etc. The user could also use data exploration tools, as profiling, time series graph generation, etc. to better constrain the ongoing deformation state and previous InSAR results.</li> <li>3. The user runs the SarScape software interface and displays one or more</li> </ol>



	<p>menus and graphic windows to select the input data, the processing method (e.g. PS or SBAS) and parameters.</p> <ol style="list-style-type: none"> <li>The user starts the processing.</li> <li>During the program execution a graphic window shows some inversion and model parameters. In case of problems the user can abort the execution.</li> <li>When the processing ends normally the user displays the results in the web GIS interface as vector shapefiles: velocity map, displacement time series (for single or groups of pixels), profiles, etc.</li> <li>The user must be able to carry out validation operations using validation tools and a web GIS interface. Typically, these operations include: best fitting of time series, selection of a reduced temporal interval, LoS component projection in cartesian coordinates, map and single point comparisons and statistics, profiling, etc.</li> <li>The user can save the processing results with a user-defined quality factor.</li> <li>If the user is not satisfied with the processing results can start the workflow all over again with different parameters.</li> </ol>
<b>Roles</b>	A geophysicist/volcanologist is in charge of all steps.

**Table 5 Volcano Deformation Mapping Use Case**

### 3.8.2 Volcano source modeling Use Case

<b>Number 2</b>	<b>Title:</b>
<b>Author</b> S. Salvi, E. Trasatti	<b>Volcano source modelling</b>
<b>Description</b>	This Use Case concerns the process of investigating the magmatic source of an active volcano
<b>Triggers</b>	A noticeable (i.e. above the measurement noise) recent ground deformation of the volcanic area (unrest).
<b>Pre-Conditions</b>	<p><b>INPUT DATA:</b></p> <p>An InSAR ground deformation map (or more than one, e.g. from different orbits or satellites) has already been calculated (as displacements or velocities).</p> <p>User has access to the InSAR deformation results, either as a single (or more) ground displacement map in cm, or as a ground velocity map in mm/yr. All maps have uncertainties and are available through an online catalogue or in the VRE database.</p> <p>User has access to GPS site displacements or velocities (same units of measure as above), with uncertainties. Same as above.</p> <p>User has access to vertical benchmark displacements from a levelling line. Same as above.</p> <p>User can display the 3D spatial patterns of recent and NRT seismicity around the volcanic area (e.g. obtained from <a href="http://www.iside.rm.ingv.it">www.iside.rm.ingv.it</a> in xls/klm/ascii formats).</p> <p>User can display geological information pertaining the volcanic area (lithological map, structural map) from specific literature for the volcanic area, in graphical formats or ascii/xls data, or from online catalogues (<a href="http://geodb.ct.ingv.it:8080/geoexplorer/composer/">http://geodb.ct.ingv.it:8080/geoexplorer/composer/</a>).</p> <p>User can display other ancillary monitoring information (e.g. microgravity data, thermal data, geochemical data), same as above.</p>



	<p>Notes:</p> <ol style="list-style-type: none"> <li>1. Not all data could be available for each volcanic area or case;</li> <li>2. The data are already treated to isolate the signals of volcanic origin.</li> </ol> <p><b>PROCESSING TOOLS:</b></p> <p>User has implemented his modelling software or has access to shared standard modelling tools as analytical forward models for volcanic sources (e.g. spherical, ellipsoidal, sill, and dyke sources).</p> <p>The ENVI-IDL, Fortran and Matlab COTS are available for use by the operator through a graphic window.</p>
<b>Post-Conditions</b>	Existence of a repository for the experiment results (volcanic sources).
<b>Activities</b>	<ol style="list-style-type: none"> <li>1. User must select the appropriate data (temporally and spatially) for the modelling experiment. To this aim the user needs to load and display on a web GIS interface the spatial distribution of the data values, select them spatially and typically create a subset of a larger data set. The data subset must be formatted as requested for the modelling tools.</li> <li>2. The user displays in a web GIS also other data which are not input to the modelling tools, as geological maps, seismicity patterns, previous models, etc. The user could also use data exploration tools, as profiling, time series graph generation, etc. to better constrain the possible source geometry, and define ranges for modelling parameters.</li> <li>3. The user runs the model interface and displays one or more windows to select the input data and their relative weights, then one window to define each of the modelling source types (more than one sources can be used) and parameter ranges.</li> <li>4. The user starts the modelling experiment.</li> <li>5. During the program execution a graphic window shows some inversion and model parameters. In case of problems the user can abort the execution.</li> <li>6. When the processing ends normally the user can display the model results as: table of best-fit model parameters with uncertainties, a graphical representation of the parameter statistics (1D e.g. frequency distribution, 2D e.g. trade off matrix); map display of the modelled data, and of the simulated and residual values; location and geometry of the source(s).</li> <li>7. The user must be able to carry out model validation operations using validation tools and a web GIS interface.</li> <li>8. The user can display and compare all data and parameters of this model with previous models.</li> <li>9. The user can save the model results with a user-defined quality factor.</li> <li>10. If the user is not satisfied with the model results can start the workflow all over again with different parameters.</li> </ol>
<b>Roles</b>	<p>A geophysicist/volcanologist is in charge of all steps.</p> <p>Input data preparation could be carried out by a different researcher.</p>

**Table 6 Volcano Source Modeling Use Case**



### 3.8.3 Retrieval of ash and SO<sub>2</sub> contained in volcanic clouds use case

Number 3	Title:
<b>Author:</b> L. Merucci, S. Corradini	<b>Retrieval of ash and SO<sub>2</sub> contained in volcanic clouds.</b>
<b>Description</b>	This Use Case concerns the process of assessing ash and SO <sub>2</sub> burden in the volcanic clouds generated during explosive eruptions from multispectral thermal infrared satellite data.
<b>Triggers</b>	On-going explosive eruptions that can release large amounts of ash and SO <sub>2</sub> in the atmosphere.
<b>Pre-Conditions</b>	<p><b>INPUT DATA:</b></p> <p>User has access to SEVIRI or MODIS multispectral satellite datasets in the thermal infrared range (TIR) collected during or immediately after explosive volcanic eruptions.</p> <p>SEVIRI and MODIS datasets can be acquired in real-time with dedicated antennas when available at user's premises or freely downloaded, also in near real-time, from specific websites (i.e. EUMETSAT <a href="https://eoportal.eumetsat.int/userMgmt/login.faces">https://eoportal.eumetsat.int/userMgmt/login.faces</a>, and NASA <a href="http://lance-modis.eosdis.nasa.gov/cgi-bin/imagery/realtime.cgi">http://lance-modis.eosdis.nasa.gov/cgi-bin/imagery/realtime.cgi</a> data portals). Main multispectral instruments datasets are level 1.5 calibrated radiances, while ancillary secondary higher level products can be downloaded when needed (i.e. cloud masks, cloud classification, cloud top height products, etc.). MODIS data are usually delivered in HDF format, while the user can choose between different data formats for SEVIRI datasets (i.e. native, H5, and others). The duration of an eruptive event capable of releasing volcanic clouds of ash and SO<sub>2</sub> in the atmosphere may vary from a few hours to many months and therefore the amount of data needed varies greatly for each specific eruption.</p> <p>User has access or has already produced ash detection maps based on the same SEVIRI or MODIS TIR datasets.</p> <p>User has access to reliable ash and SO<sub>2</sub> cloud masks generated from these detection maps.</p> <p>User has access or can estimate, either from the image or other independent sources, the volcanic ash cloud height.</p> <p>User can display the TIR images collected by SEVIRI or MODIS sensors and clearly identify by visual inspection the ash and SO<sub>2</sub> clouds released into the atmosphere by the volcanic eruption that is the object of the study.</p> <p>Note:</p> <ol style="list-style-type: none"> <li>1. Not all data could be available for each volcanic area or case.</li> </ol> <p><b>PROCESSING TOOLS:</b></p> <p>User has access to the Volcanic Plume Retrieval (VPR) procedure code implemented with coefficients previously calculated for specific ash type, sensor and volcanic area,</p>



	i.e. Mt. Etna (Sicily, Italy) or the Icelandic volcanoes.
<b>Post-Conditions</b>	Existence of a repository for the experiment results (volcanic sources).
<b>Activities</b>	<ol style="list-style-type: none"> <li>User must select the appropriate data (temporally and spatially) for the ash and SO<sub>2</sub> retrieval experiment. To this aim the user needs to select the multispectral SEVIRI or MODIS TIR images that actually contain the volcanic cloud that has to be investigated.</li> <li>User must select the ash and/or SO<sub>2</sub> cloud mask corresponding to each multispectral TIR satellite image.</li> <li>User must input to the procedure the estimated volcanic cloud height previously independently estimated.</li> <li>The user starts the ash and SO<sub>2</sub> retrieval experiment.</li> <li>The user runs the VPR procedure and obtains the maps of SO<sub>2</sub> and ash masses contained in the volcanic cloud and of the ash effective radius and aerosol optical depth at 550 nm (AOD).</li> <li>After the program execution the ash and SO<sub>2</sub> maps produced by the procedure can be displayed in dedicated graphic windows.</li> <li>When the processing ends normally the user can display the procedure results as: the aforementioned maps, a summary table of the main estimated parameters, i.e. ash and SO<sub>2</sub> total mass, mean ash effective radius, and mean AOD, and a table that summarise the VPR procedure settings used for the specific retrieval experiment, i.e. the specific VPR coefficients related to the ash type, the sensor and the volcanic area that were selected.</li> <li>The user must be able to carry out the validation operations using other approaches for the ash and SO<sub>2</sub> retrievals.</li> <li>The user can save the model results with a user-defined quality factor.</li> <li>If the user is not satisfied with the model results can start the workflow all over again with different parameters.</li> </ol> <p>OUTPUT DATA:</p> <p>Maps of SO<sub>2</sub> and ash masses contained in the volcanic cloud and of the ash effective radius and aerosol optical depth at 550 nm (AOD). Total amounts of SO<sub>2</sub> and ash contained in the volcanic cloud.</p>
<b>Roles</b>	<p>A geophysicist/volcanologist is in charge of all steps.</p> <p>Input data preparation could be carried out by a different researcher.</p>

**Table 7 Retrieval of Ash and SO<sub>2</sub> Contained in Volcanic Clouds Use Case**

### 3.8.4 Water vapour content use case

Number 4	Title:
<b>Author C. Spinetti, V. Romaniello</b>	<b>Water Vapor Content</b>
<b>Description</b>	This Use Case concerns the process of retrieving the atmospheric water vapour contained in a volcanic cloud over an active volcanic area
<b>Triggers</b>	The generation of water vapour content maps is based on a procedure using hyper and multispectral data in the NIR-SWIR spectral range. It is performed every day or depending on the satellite passes. The water vapour content of the volcanic cloud is





	retrieved in case of eruption.
<b>Pre-Conditions</b>	<p><b>INPUT DATA:</b></p> <p>EO satellite data: NASA EO-1 Hyperion, TERRA and AQUA MODIS data daily acquired of the volcanic area are available.</p> <p>Digital Elevation Model of the volcanic area available.</p> <p>Atmospheric parameters. They are generally available from Radio Sounding profiles close to the local volcanic area that need to be on line. Otherwise ERA-Interim numerical weather model data produced by the European Centre for Medium-Range Weather Forecasts (ECMWF).</p> <p>Notes:</p> <p>Not all data could be available for each volcanic area or case</p> <p><b>PROCESSING TOOLS:</b></p> <p>Validation tools in the EVER-EST map interface.</p>
<b>Post-Conditions</b>	Existence of a repository for the processing results (water vapour content maps).
<b>Activities</b>	<p>The processing of EO multi and hyperspectral data consist in the following consecutive steps:</p> <ol style="list-style-type: none"> <li>1. Pre-Processing. This step consists in DEM co-registration and extraction of SWIR calibrated bands. Input data as Level 1 data and DEM into the pre-processing step.</li> <li>2. Simulation. This step consists in simulating the radiance of the SWIR bands. Input of the RT model simulation is the output of pre-processing and pre-conditional data (eg., date and time of data acquisition, sun and view angles, atmospheric parameters). The model generates a defined simulations scenario.</li> <li>3. Inversion. This step consists in obtaining the resulting map. Input of the inversion algorithm is the output of the simulation step and pre-conditional data. The output of this step is the water vapour content map.</li> </ol>
<b>Roles</b>	A geophysicist/volcanologist is in charge of all steps.

**Table 8 Water Vapour Content Use Case**

### 3.9 Geohazard Supersites VRC user actions Specification

Geohazard Supersites user scenarios and user action specification, further developed from the above use cases, at this stage of the project, are detailed in the Appendix B.



## 4 **Land Monitoring VRC**

Land Monitoring can refer to the monitoring of urban, built-up and natural environments to identify certain features and anomalies or changes over areas of interest as well as the assessment of natural resources. Land Monitoring is a transversal issue common to a number of user communities; the main aim is to provide useful information to those entities that have to make informed decisions to address environmental, scientific, humanitarian, health, political and security issues as well as to adopt sustainable management practices. In this scenario, Earth Observation (EO) data play an important role and activities related to their usage in this context are currently coordinated by European programmes such as Copernicus.

The rapidly increasing amount and variety of data suitable for land monitoring coming from satellites and other sources imposes addressing new issues such as the management and exploitation of extremely large and complex datasets, the extraction of knowledge from very large volumes of data and the development of online collaborative systems; one of the main challenges is to improve the capacity to extract in a timely manner operational (i.e. useful and clear) information from a huge amount of heterogeneous data.

The development of a VRE for land monitoring scenarios aims to deal with these issues addressing different communities that potentially have different final goals but which are using the same space assets and similar services/techniques. The related workflows of relevant suitable scenarios will be inserted in a context of Research Objects to encapsulate essential information on experiments and investigations in order to facilitate their reusability and reproducibility as well as to allow a better understanding amongst the users of Earth Science data for land monitoring applications.

Following these considerations, the main objectives of the Land Monitoring use case in the framework of EVER-EST are:

- Fostering the use of Earth Science data for land monitoring;
- Building a VRE suitable for land monitoring purposes;
- Exploring the use of ROs for land monitoring applications.

The European Union Satellite Centre (SatCen) can represent, in the framework of EVER-EST and in line with the “Secure Societies” Horizon 2020 Societal Challenge, the stakeholders involved in the decision-making process of the EU in the field of the Common Foreign and Security Policy (CFSP). SatCen will be the player active in the domain collecting user requirements from the Land Monitoring community, implementing an appropriate use case related to Land Monitoring for Secure Societies and testing the functionalities and the infrastructure set-up in the project.

### 4.1 **Land monitoring VRC user requirements**

#### E-collaboration services

Main points are:

- To collect and share knowledge, tools and open data;
- To link the scientific and the institutional communities allowing exchange of information;
- To have a reliable user authentication and authorization accessing the VRE;
- To have an appropriate dissemination interface towards the end-users.

#### E-research services

Main points are:

- To guarantee a single access to properly archived data and tools;
- To have the possibility to create and upload data and tools;



- To guarantee provenance and traceability of data and tools;
- To have modularity allowing an easy modification and creation of single tools or of the entire workflow;
- To use the research environment as a PaaS;
- To provide easy and intuitive tools to support the assessment of methods and solutions;
- To have appropriate functionalities and graphical interfaces (e.g. GIS functions and visualization) to ingest, manage, analyse (visual analytics) and exploit geospatial data.

#### E-learning services

Main points are:

- To have tools supporting the organisation of webinar and online courses;
- To have training on the platform usage during dedicated events.

#### Digital information services

Main points are:

- To have easy acquisition and discovery of data;
- To create and access data/text mining tools as well as of automatic tools for processing high volumes of data;
- To have the possibility to use suitable semantics and ontologies;
- To increase the availability and use of open data and open sources tools;
- To rate the infrastructure and the related services.

#### Common services

The developed VRE has to be reliable, stable, efficient, time reducing, economically convenient, open source-based, user-friendly, multi-lingual, interoperable, improving current practices, customizable and extensible.

A key requirement for the Land Monitoring VRC is the security compliance. The VRE needs to comply with the existing data policies and standards for security (e.g. ISO27001) in terms of the information contained (classified data, secure data transfer and data encryption), property rights, data privacy, certified technical solutions and legal aspects (with respect to the output of mixing open and classified data).

### **4.2 Land monitoring VRC context framework description**

The development of a VRE suitable for Land Monitoring will aim at sharing applications, knowledge and expertise among different user communities. The use case will be structured with the ambition of: 1) addressing the requirements of the land monitoring community interested in security applications using mainly data coming from space assets; 2) addressing the requirements of other user communities that potentially have different final goals but which are using the same space assets and similar services/techniques for land monitoring (i.e. in the domains of climate, environment, transport, etc.), thus saving effort and money.

Relevant stakeholders will be asked to evaluate the VRE and to provide feedback through appropriate events regularly organized by the EU, SatCen and parallel research initiatives.

### **4.3 Land monitoring VRC use case selection criteria and methodology**

SatCen is currently building and consolidating a Space and Security community in the framework of its Research, Technology Development and Innovation (RTDI) activities. As a key element of this community, SatCen is working effectively on the development of technical and programmatic solutions needed for the

management and exploitation of data, always starting with the collection of user requirements from the community.

A number of discussions have taken place with SatCen staff belonging to different divisions and other relevant security stakeholders such as EU entities and international organizations; the topics were related to the current data management practises and how they could be improved.

In accordance with the general requirements described in section 4.1 it has been expressed the need of exploiting the increasing amount of data coming from space and other sources, where a major contribution of open data and tools is foreseen. Automatic tools for data management and processing are one of the key aspects, where the adopted solutions have to be developed and integrated in the whole data chain in order to reduce the human effort, to reduce the amount of needed economic resources and to efficiently deal with data analysis procedures.

As common requirements, the land monitoring community is interested in:

- Collecting efficiently data from different sources (e.g. from remote and social sensing);
- Detecting changes and anomalies on specific areas in defined time series analysis;
- Characterizing qualitatively the detected changes.

With respect to the use of open data, dedicated satellites under the Copernicus<sup>4</sup> programme can guarantee open and up-to-date information through a set of services dedicated to environmental and security issues.

The access and use of Copernicus Sentinel Data and Service Information is regulated under EU laws; the free, full and open data policy adopted for the Copernicus programme foresees access available to all users for the Sentinel data products, via a simple pre-registration on the Sentinel Scientific Data Hub<sup>5</sup>. Other data coming from social, commercial, open and other sources (such as video, photo, reports, social media information, GPS measurements, in situ measurements, etc.) could be a valuable support in addition to satellite images.

With respect to the detection of changes, the users expressed the interest in an automatic system raising an alert when a change is verified between two or more images on specific areas of interest. In this case, the user is then going to visualize, verify and validate the reasons why the alert has been triggered (e.g. with the support of a GIS interface).



**Figure 7 Change detection (in red, right image) on two images acquired in 2006 and 2010 (source: LISA)**

With respect to the characterization of changes, some users are interested in analysing qualitatively these changes investigating on their nature and thus increasing the complexity of the analysis. With the spatial resolution of the Sentinels' data (from 300 to 5 m), changes related (but not limited) to the following scenarios in built-up and natural environments were highlighted:

---

<sup>4</sup> <http://www.copernicus.eu/>

<sup>5</sup> [Sentinel Scientific Data Hub](#)



- Identification of main land cover surfaces (such as water, bare soil, man-made areas, vegetation) and assessment of their evolution (e.g. detection of a cleared area in a dense forest);
- Monitoring of critical infrastructures (e.g. large buildings, roads, airfields) and assessment of their condition (e.g. under construction, destroyed);
- Detection of large objects (e.g. aircrafts or ships).

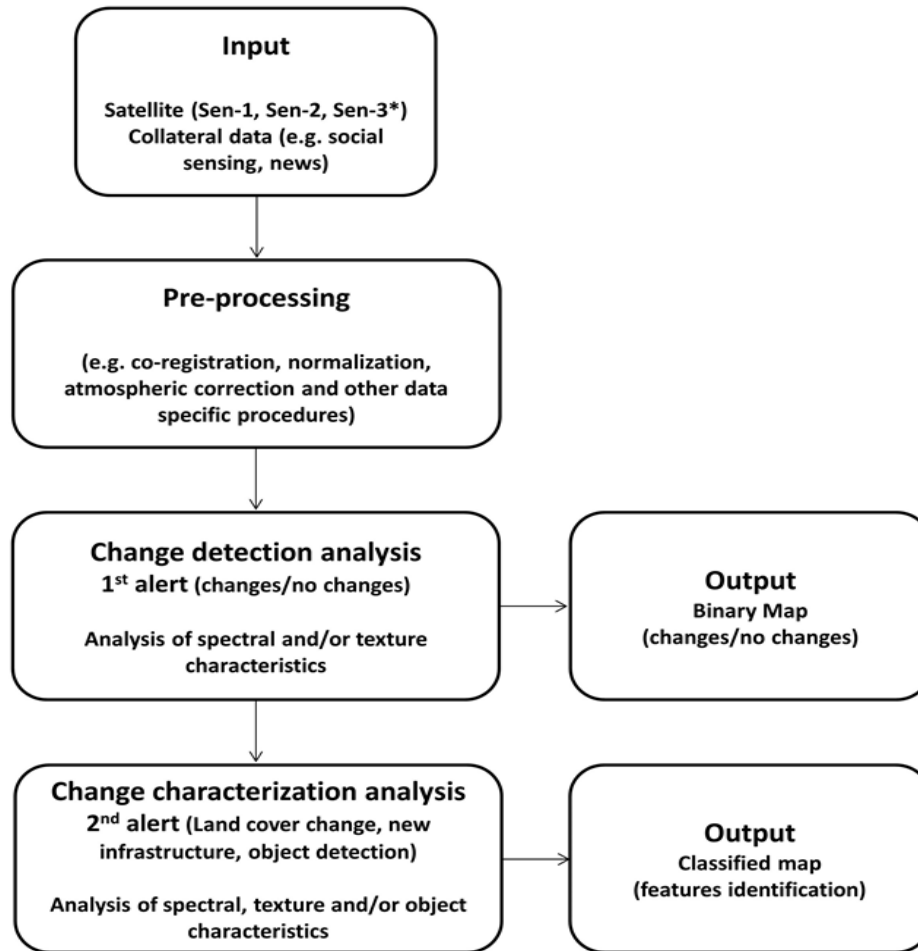
The qualitative analysis should provide an alert system indicating not only the simple detection of a change in a particular area of interest but also whether certain conditions of change have been reached.



Figure 8 Qualitative analysis identifying the construction of new buildings (source: LISA)

#### 4.4 Land monitoring VRC selected use cases description

Following the considerations expressed in the previous paragraphs, SatCen identified a use case based on a modular structure. The modularity characteristic is fundamental to allow the user to verify each single step of the chain, to perform only a subset of operations, to substitute a single module when needed and to link some modules to other modules. The following structure of the workflow in **Figure 9** has been built.



**Figure 9: Use case workflow**

This use case considers the ingestion of satellite images acquired on land areas from Sentinel 1, Sentinel 2 and Sentinel 3 (see section 4.5 for further details) with the support of other information from social sensing sources and other geotagged data. Appropriate pre-processing algorithms will be applied in order to create a normalized time series where change detection and characterization analysis could be performed. The first analysis is focused on the detection of changes, where an alarm will be sent to the user if changes occurred. The analysis will be based on spectral signatures (e.g. indexes extensively used or new indicators) and texture characteristics (e.g. granularity analysis). To this aim, there will be used processing tools already available for Sentinels data (such as tools provided by SNAP<sup>6</sup> or other existing sources); if needed in-house tools will be developed. A different access to the change detection processing chain should be granted to users with different levels of technical skills (users should be able to apply the default change detection parameters, to tune them or to plug-in other algorithms to perform the same processing). The visual interface should support the analysis of the 1<sup>st</sup> alarm raised with a map of changes/no changes by which the user will be able to verify and validate the output map and the input data, which brought to the alarm. Users should also be able to manage the confidence level of alarms they need (it is expected that a 95 % of confidence level will return only few areas where the changes are detected with respect to an 80 % confidence level).

<sup>6</sup><http://step.esa.int/main/toolboxes/snap/>





After raising the first alarm, the user will already have obtained a main result from the VRE; therefore the analysis can be stopped at this level or the results can be further explored with the change characterization analysis. The same approach used in the change detection step will be considered analysing spectral signatures (e.g. target detection for new infrastructures) and texture characteristics (e.g. homogeneity of a building) as well as adding object analysis (e.g. object characteristics such as elongation for roads, squared shape for buildings, etc.). A second alert is triggered if the change is identified as relevant with respect to a predefined list of changes (e.g. new building construction, road interruption, etc.) to be populated in accordance to user requirements. For this analysis a major contribution of in-house developed tools is expected with a different access level based on the user technical skills as in the change detection step. A visual interface should support again the analysis of the second alarm with visual analytics tools, validation and verification methods. Users should also be allowed to provide feedbacks in order to improve the alarm rise system.

The workflow can be connected to other workflows developed inside the EVER-EST project (e.g. by another VRC) or to other services/platforms (Sentinel level 1 data calibration, Thematic Exploitation Platforms, H2020 Big Data Europe platform, etc.).

#### 4.5 Land monitoring VRC data needs

Data to be used for the use case are described below:

- Sentinel 1 data - the required Sentinel data for this use case are available in Level-1 as Single Look Complex (SLC) and Ground Range Detected (GRD) images. The data acquisition can be configured with different modes: the default mode over land has a swath width of 250 km and a ground resolution of 5 x 20 m while the Stripmap mode provides a continuity of ERS and Envisat data offering a 5 x 5 m resolution over a narrow swath width of 80 km. Images have to be acquired in interferometric conditions in order to perform a proper co-registration. Images acquired in different time will be compared to perform the change detection;
- Sentinel 2 data - the required Sentinel data for this use case will be available in Level-1 (radiometrically and geometrically corrected) and ideally in Level-2 (atmospherically corrected). Images are composed by 13 spectral bands: 4 bands at 10 m (Visible), 6 bands at 20 m (Near Infrared) and 3 bands at 60 m (Shortwave Infrared), with a swath width of 290 km. Images acquired in different times will be compared to perform the change detection;
- Sentinel 3 data - the required Sentinel data for this use case will be derived essentially from the Ocean and Land Colour Instrument (OLCI) based on heritage from Envisat Medium Resolution Imaging Spectrometer (MERIS) with 21 bands at a resolution of 300 m. Data will be available in Level-1 (radiometrically and geometrically corrected) and ideally in Level 2 (atmospherically corrected). Algorithms deployed for vegetation assessment might be considered;
- High and Very High Resolution data (if available) - data with higher resolution than Sentinels could assist in the interpretation of the results and changes verification. In a second phase, the algorithms developed for the Sentinels might be adapted for these images;
- Other geotagged data, structured and unstructured, coming from social, commercial, open and other sources - these data can be available as video, photo, reports, social media information, in situ data, etc. and will be used as input to the analysis and to validate the output of the workflow.

#### 4.6 Land monitoring VRC use case scenario for research object modeling

Several ROs could be created for change detection and change qualitative analysis. The general structure can be a single RO encapsulating others ROs, where the specific modules can be exchanged.





With respect to the ROs available in the library, the RO “Extraction of Galaxy Samples from Images” could be adapted to the extraction of features (e.g. airfield with runways) from 2D satellite images. The land monitoring ROs can be furthermore associated to other ROs (e.g. calibration of input satellite data).

- Input data could be the ones described in the workflow and in a second phase other satellite data could be considered. The data selection should be present at the platform (e.g. interface with Sentinel Hub) or the user should be able to upload its own data. Documentation on the data (e.g. Sentinels guidebooks or data provenance) should be available;
- All the processing algorithms (e.g. co-registration, change detection, feature extraction) should be properly described with reference papers with a different access level granted to users with different technical skills;
- Output can be visualized or downloaded by the user as image or report;
- Other geotagged data upload procedures should be considered, where these data can be stored on the infrastructure. If results are validated in the VRE (extended GIS and raster processing capabilities are required), validation reports should be available in the RO.

#### **4.7 Land monitoring VRC use case validation methodology and success criteria**

The use case will be validated and evaluated by the user community to provide its view and feedback on the VRE in accordance with the different phases of the EVER-EST project, e.g. ranging from the definition of user needs and requirements and the continuous evaluation of the VRE capabilities to the final assessment of the VRE added value in addressing real use case and operational user scenarios. The collection of the user requirements methodologies has been discussed in section 4.1

The evaluation of the VRE capabilities will be mainly based on criteria related on qualitative and quantitative indicators.

- The qualitative method will be mainly supported by visual interface, where the users will be able to display the output to assess the quality of the final product. For instance, in the case of the change detection, a simple comparison of the output with the input images will be a preliminary evaluation of the results. As indicated in data section 4.5 other images (such as Very High Resolution images) and geotagged data could support the interpretation of the results;
- The quantitative method will be based on different criteria related to the platform characteristics (which might be common to other use cases) and the Land Monitoring use case. The first category might include the computational response, the easiness of the platform use and the platform extensibility. The second category might include specific criteria related to the use case which could be associated to the extraction of ad-hoc indicators (e.g. range of alarm values due to the entities of changes detected) or different accuracy values (e.g. if ground truth and in-situ data are available, the latter in particular for the qualitative analysis), according to the available services;

The success of the Land Monitoring use case will be measured by the feedback provided by the majority of users reached within the Land Monitoring VRC members. SatCen will organize appropriate events to collect user feedbacks and will report about these activities.

The Land Monitoring use case aims at providing a concrete case of cross-disciplinary interaction between Earth Scientists and Institutional entities to transfer knowledge, best practices and tools. The user scenarios, the workflows and the land monitoring infrastructure services will be constantly updated based on the user requirements coming from the Land Monitoring community, demonstrating the VRE compliance with the community needs. It is expected that the users will contribute to the workflow testing the existing modules and populating the Research Objects with new features. The Land Monitoring VRC should receive proper training on the VRE functionalities during the different phases of the project. Ideally, when the methodologies will be tested and validated, the workflow would be potentially tested in an operational scenario.



#### **4.8 Land monitoring VRC user actions specification**

Land Monitoring VRC user scenarios and user action specification, further developed from the above use cases, at this stage of the project, are detailed in the Appendix C.



## 5 **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. By establishing a consortium of relevant public bodies the NHP aims to:

- Establish a forum for the exchange of knowledge, ideas, expertise, intelligence and best practice in relation to natural hazards;
- Provide a timely, common and consistent source of advice to government and emergency responders for civil contingencies and disaster response;
- Create an environment for the development of new services to assist in disaster response.

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 HIM group partners are: the Met Office, the Centre for Ecology & Hydrology, the British Geological Survey, the Health and Safety Laboratory and Kings College London. Initially they are concentrating on modelling the impact of 3 key hazards – surface water flooding, land instability and high winds – on people, their communities and key assets such as road, rail and utility networks. The partners share scientific expertise, data and knowledge on hydrological modelling, meteorology, engineering geology, GIS and data delivery and modelling of socio-economic impacts.

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.

### 5.1 **Natural Hazard VRC user requirements**

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

#### **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.

### **Document Storage and Sharing**

- A document repository with editing capability.
- Needs to be accessible by different partner organisations with different internal security issues.

### **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

### **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.

### **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

### E-Research services – Services which support research processes and tools

#### **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.

#### **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.
- The VRE could provide access to modules at different levels of complexity (from detailed process models to simple I/O structures).
- 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.
- 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.
- Needs to allow for multiple scripting languages.
- Library of tools could be provided by a smart user interface providing easy drag and drop facility.

### **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.
- 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.

### **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)

### **Visualisation**

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

### Digital Information services

#### **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.
-



### 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.

### 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

### 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.

### E-Learning services – Services providing access to learning materials

#### 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).

## 5.2 Natural Hazard VRC context framework description

The Hazard Impact Model group was set up to:

- 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 to customer specific visualisation systems.
- 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.
- 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.

Hazard Impact Models are being developed for three hazards in the first instance: Surface Water Flooding, Land Instability and High Winds, via 3 distinct work packages. A standard approach to modelling of the impact of these hazards is being developed as part of a 4<sup>th</sup> Vulnerability work package. The operational delivery of the impact model outputs is being developed and tested as part of the 5<sup>th</sup> Hazard Impact Productions System work package. Validation and testing of the impact model outputs is provided by the Hazard Impact Database work package. Each work package (detailed below) is led by one of the NHP partner organisations but requires significant collaboration by one or more of the other partners.

[1] 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 Laboratory (HSL). An end-to-end offline trial system has been demonstrated using historical case studies. On going work is developing an end-to-end real-time trial system and gaining evidence from further offline case studies, sensitivity analyses and verification (supported by Kings College London). Offline fine-scale inundation modelling by JBA is testing and possibly improving the link between CEH's hazard and HSL's impact assessments of SWF risk.

[2] Land Instability work package – led by NERC's British Geological Survey (NERC BGS)

The aim of this work package is to ensure that a rainfall-induced land instability hazard impact model is developed and delivered for the NHP using the best available science. In order to develop a HIM for land instability, the science used to determine a real-time hazard footprint in the UK is currently the main focus of this WP. This WP relies on Met Office rainfall data and collaboration with HSL to determine potential impacts of hazardous events and with CEH on gridded estimates of soil water and runoff.

[3] Wind work package – led by the UK Meteorological Office (Met Office)

The aim of the Wind work package is to develop a comprehensive set of modules within the Hazard Impact Model (HIM) to forecast the impact of high winds on assets, infrastructure and human activities in the UK. A prototype probabilistic version of the Vehicle Overturning (VOT) model has been developed which is now running live and forecasts the impact of the road network of overturning vehicles during high wind events. This model uses the Met Office convective-scale MOGREPS-UK weather forecast ensemble as input for the hazard forecast. This WP relies on collaboration with HSL in reporting the likely impact of events in a consistent manner.

[4] Vulnerability work package – led by Health and Safety Laboratory (HSL)

This work package is developing proof of concept HIM prototypes/interfaces and gathering additional impact and vulnerability information and science to deliver the overarching aims of the HIM. It involves close liaison with the other work packages to identify and deliver appropriate vulnerability/impact information. HSL has experience in demographic modelling, population vulnerability and spatial impact analysis. Much experience has been gained on impact modelling over the last 2 years through the proof-of-concept SWF work that has



been developed together with the FFC. HSL is also leading the development of the Hazard Impact Framework which is an overarching framework put in place to ensure impact modelling is done in a consistent and repeatable manner.

#### [5] Hazard Impact Production System (HIPS) work package – led by the Met Office

The key concept of the HIPS is that diverse models (termed Modules within the system) will deliver outputs via Open Geospatial Consortium (OGC) Web Services. Use of these open standards will allow a modular approach to be taken to the production system: data clients and production servers can be added, removed or upgraded without requiring the redesign of other parts of the system. The Hazard Impact Framework (HIF) mentioned above provides the guidelines for all the HIM Modules to follow. The pilot implementation of a HIM Module dealing with VOT (see Wind WP above) has been developed in collaboration with Met Office scientists. OGC Web and Feature Services are generated by the prototype and have been demonstrated with a range of client software. Emphasis has now switched to the Surface Water Flooding Module (described above) in collaboration with CEH, HSL and the FFC.

### 5.3 Natural Hazard VRC use case selection criteria and methodology

#### Collaboration

Our first priority would be 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.

A priority for the Hazard Impact Modelling group would be to test how the VRE might enhance this collaboration by providing an easy to use and accessible online meeting place (see user requirements for specifics).

#### Large data sharing and model development

Each work package described in 5.2 requires participating partners to share large datasets of different types. These include socio-economic datasets as well as scientific datasets. The Land Instability work package provides an ideal test of what the VRE could provide in terms of sharing of large disparate datasets.

#### Model development, workflow and code sharing



Each work package is developing its own Hazard Impact Model via collaboration with other partners. However each work package is at a different stage of maturity. The Surface Water Flooding work package would provide an ideal test case for testing interactive hazard modelling, impact analysis and verification between partners.

### Methodology development

The Hazard Impact Framework activity would provide a useful starting point for testing how the VRE might support sharing of common methodologies and ontologies across different Hazard Impact Models.

## 5.4 Natural Hazard VRC selected use case descriptions

E-Collaboration could be valuable to all Hazard Impact Modelling activities described in 5.2 so no further detail is provided here.

**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. This involves running a countrywide (1km grid, 15 min time-step) Grid-to-Grid (G2G) hydrological runoff and routing model (CEH) using rainfall inputs (Met Office), and linking its surface runoffs to potential impacts (HSL) and verifying these against observed impacts (Kings 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 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 HSL, whilst email suffices for exchange of impact validation data. A live trial operational workflow producing and visualising SWF impact products will be in place in 2016.

**The Landslide Instability** Hazard Impact Model is at an earlier stage of development. There is planned work to explore the use of gridded time-series G2G outputs of surface runoff and soil moisture variables in the context of landslide hazard, involving partnership working between CEH (G2G runoff and soil water modelling), BGS (landslide modelling), HSL (impact assessment) and BGS/KCL (impact validation); again, meteorological data (e.g. rainfall) is sourced in part from the Met Office. Alternative model formulations will be trialled within a model development, calibration and verification framework involving collaborative working across partners working on separate Linux and Windows platforms.

**The Hazard Impact Framework** (HIF) aims to identify common methodologies and ontologies across HIMs for different hazards (e.g. wind hazard and vehicle overturning, SWF and landslides), including their interdependencies and joint occurrence. This will map vocabularies across hazard disciplines, lead to a better common understanding and bring benefits through exploiting generic features.

## 5.5 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 are shared, which include:

### **1) Spatial Data (i.e. with X,Y and Z values)**

Raster and Vector Data is shared. This could be in a number of different data formats e.g . ESRI Shapefiles, MapInfo Files, Web Map Services (WMS), Web Feature Services (WFS), NIMROD format, ASCII, 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 include:

1. National Population Database (owned by Health and Safety Laboratory) - Is a vector dataset which contains details of population types, their locations (and therefore exposure to a potential hazard) and vulnerability e.g. are they school, care homes, hospitals etc.
2. BGS GeoSure Data – Is a vector dataset, containing information about different geological hazards in the UK.
3. MET Office Data - Raster data (NIMROD or ASCII format) e.g. Temperature, Relative humidity data, Precipitation Forecast
4. Met Office Hourly Precipitation Data -
5. Soil Moisture Data – Raster Data
6. Satellite imagery?
7. Data stored in a database, but with XY attributions e.g Nexus Lexus database (managed by KCL)
8. National Receptors Database (owned by Environment Agency) - this contains information on assets and infrastructure including roads, railways, buildings etc

Other datasets would include but not limited to:

Validation Datasets

- 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.

### **2) Documents (e.g Word, PDF, Excel Files)**

- Historical precipitation records (for calculating antecedent rainfall) are often stored in Excel or txt files.
- Validation Datasets e.g Responder reports – fire, police and ambulance incident reports.
- Data that can be harvested from Social media.

### **3) Images (e.g .jpg, .tiff)**

### **4) Presentations (e.g Presi, Powerpoint)**

Many of these data would consist of both living (constantly updated) and final (published) files. An example of a living file would be our glossary of terms. Different users add to this document, defining new terminology or refining existing terms. Our glossary would benefit from a Wikipedia style-sharing environment.



## 5.6 Natural Hazard VRC use case scenario for research object modelling

These scenarios shall cover the information needed for modelling a digital RO with respect to submission, archival and dissemination stages.

Research objects - possible candidates for research objects

- 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 being developed in a modular fashion as standalone components, each of which could also be defined as Research Objects. For example:
  - The G2G component of the Surface Water Flooding HIM and its associated method descriptions and guidance notes etc.;
  - 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;
  - 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.

## 5.7 Natural Hazard VRC use case scenario general description

One example of a Use Case scenario is for the validation and testing of HIMs. The HIM components described above fit into a workflow that starts with observed data on rainfall and river flow conditions as input to the G2G model to establish the present hydrological state, and then ensemble forecast rainfalls input to G2G yields ensemble surface runoff forecasts. These are post-processed to assess SWF hazard and, using the Impact Library, impact and likelihood are calculated and this flood risk information summarised for end-users. For scenario runs or case studies, each component owner is currently required to be part of the process, creating a barrier to use based on resource and communication. If this process was set up as a fully automated workflow based on the research object model, end-users could run the process themselves. Further, other interested parties could access the workflow to do their own testing, or validation work. These parties might include universities (e.g. KCL) or other working groups (e.g. within the NERC-funded programme “Flooding From Intense Rainfall”), as well as other partners in the NHP.

This would also allow the organisations involved in the development work to explore how changes to their HIM components might affect the final outputs. For example, changes to the G2G model configuration, or thresholds used in the hazard and impact post-processor, can be tested and assessed against verification datasets.



Used in this way the NHP would adopt the EVER-EST platform as a development environment, providing access to working versions of the HIM for testing and comparison with the live operational environment. Versioned literature and documentation would be an important element of the RO, providing potential users with instructive material, recent publications and links to relevant material. Discussion forums could be used to facilitate dialogue about the HIM, with instant chat functions providing the means for ‘live’ collaborations.

## **5.8 Natural Hazard VRC user actions specification**

Natural Hazard VRC user scenarios and user action specification, further developed from the above use cases, at this stage of the project, are detailed in the Appendix D.



## 6 Earth Science VRE User Requirements

The EVER-EST VRCs differ substantially one from each other in terms of the data they use, their teams setup, the working environments and open data policy related to the data they use and generate, the workflows adopted and have heterogeneous needs. EVER-EST VRE platform architecture while not forcing any of its users to radically change the way they work shall support new, easy to use, innovative features for overcoming problems and transforming current challenges in new opportunities.

Particular effort has been put in applying the H2020 guidelines for implementing Findable, Accessible, Interoperable and Reusable (FAIR) data while taking into account the VRCs specific requirements in relation to openness and protection of scientific information, commercialisation and Intellectual Property Rights (IPR), privacy concerns, security and long term preservation needs. EVER-EST VRE projects activities and requirements gathering from the VRCs, aiming on harmonization of metadata and ontologies for the VRE are being addressed in WP4. To be quoted that data and research object produced and used during the project will be identifiable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers).

In the following chapters, user requirements elicited from the EVER-EST communities operational scenarios, use cases, needs and challenges described in Chapter 2, have been assessed, grouped for commonalities of needs, categorized and rated as described in Chapter 1.3.

User Requirements will be translated into specifications for the services supported by EVER-EST VRE re-usable building blocks components, open sources and VRCs owned COTS integrated in a SOA architecture, to ensure that the technical developments are well aligned with the common requirements as well as the needs of the individual use cases and fulfil the overall EVER-EST main objectives described in 6.1. To this scope detailed user requirements have been listed in chapter 6.2, grouped under the following main services categories:

- EVER-EST Common Services;
- EVER-EST e-Collaboration Services;
- EVER-EST e-Research services which include digital information and preservation;
- EVER-EST e-Learning services.

### 6.1 **EVER-EST infrastructure key objectives**

Key objectives of EVER-EST Earth Scientists user communities (pre-selected and potential new ones), is to easily discover, access and exchange reliable data, have access to suitable processing power, visualization and analytics tools shared across communities, maximise processing locally at the VRE, share for validation and reuse their methods and approaches, observations, results and the lessons they have learnt with peer scientists and external communities.

The following general requirements elicited from the EVER-EST proposal are mandatory for the EVER-EST VRE Infrastructure and shall be verified end-to-end according to the verification methods identified in the requirements.



EVER- GEN -010	EVER-EST General Requirement	Mandatory
Requirement	<p>The EVER-EST e-infrastructure shall allow Earth Science data users to:</p> <ul style="list-style-type: none"> <li>Discover, access and process both existing and new heterogeneous Earth Science datasets including the associated information and preserved knowledge held by distributed data centres;</li> <li>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);</li> <li>Capture, annotate and store the workflows, processes and results from their research activities;</li> <li>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;</li> <li>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.</li> </ul>	
Source	Proposal, SMART Objectives	
Verification Method	Verification Methods: Test, Review of Design	

EVER-GEN-020	EVER-EST General Requirement	Mandatory
Requirement	<p>EVER-EST e-infrastructure shall enable and facilitate Earth Science Researchers by:</p> <ul style="list-style-type: none"> <li>Supporting the processes of conducting research and the creation and maintenance of collaborations across domains, institutions and countries;</li> <li>Be designed to meet user requirements and address usability and accessibility, with appropriate evaluation mechanisms and benchmarks for new service and tool development;</li> <li>Be secure and trustworthy – the EVER-EST components should interoperate with federated cross-institutional authentication and authorisation mechanisms;</li> <li>Be accountable, by providing adequate logging including supporting queries about provenance;</li> <li>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;</li> <li>Allow easy and efficient Earth Science data browsing, query and access from multiple sources, adapting to different standards and services;</li> <li>Support the creation, sharing and curation of digital content, through ease of authoring, publishing, discovery and access. This implies</li> </ul>	





	<p>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;</p> <ul style="list-style-type: none"> <li>• Maximise the processing of the data at EVER-EST VRE.</li> </ul>
<b>Source</b>	Proposal
<b>Verification Method</b>	Verification Methods: Test, Review of Design

EVER-FUN –030	EVER-EST General Requirement	Mandatory
<b>Requirement</b>	<p>EVER-EST e-infrastructure shall enable and facilitate Earth Science Researchers by</p> <ul style="list-style-type: none"> <li>• Be based, as far as possible, on loosely coupled, distributed, interoperable services and tools, rather than monolithic applications;</li> <li>• 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;</li> <li>• 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;</li> <li>• Support tailoring of the environment by individuals or groups to reflect their domain interests and personal preferences;</li> <li>• Support the delegation of routine tasks to intelligent personal agents where the means to realise these exists, e.g. by incorporation into workflow processes;</li> </ul>	
<b>Source</b>	Proposal	
<b>Verification Method</b>	Review of Design	

## 6.2 EVER-EST services user requirements

For each services group requirement, a brief description will be provided summarising VRCs main use cases challenges and rational for the corresponding requirements defined for the VRE, along with the definition of the scope of the service group within the EVER-EST research and development project context.

The definition of the scope of each main service categories, rational for the requirements, rating, categorization and relevant sub-requirements grouping are explained in the following sub-chapters:

- EVER-EST Common Services;
- EVER-EST e-Collaboration Services;
- EVER-EST e-Research Services which include Digital Information and Preservation Services;
- EVER-EST e-Learning Services.

## 6.2.1 EVER-EST common services requirements

### Rationale for EVER-EST Common Services Requirements: Needs and Challenges

The definition of the scope of the EVER-EST common services takes into account the highly heterogeneous nature of the VRCs communities operational scenarios and research digital objects generated, which span from experimental data, computational results, enhanced publications to executable papers. Common to all VRCs are the challenging requirements for EVER-EST to foster scientific coordination and facilitate sharing of experiences by:

- Improving the adoption for collaboration environment among scientists located in distant parts of the world (addressed in 6.2.2);
- Complement the collaboration environment with means to share results and tools within projects sharing the same funding source by enabling the possibility to define specific VRE sub-environments (i.e. for authorization to access resources) as the community could be composed of sub-communities with restricted resources, e.g. there could be scientists belonging to one project within the community or scientists involved in a subset of projects within the communities for which specific agreements are in place that cannot be extended to all scientists of the community;
- Establish methods for early attribution of Intellectual Property (IP) and citation, which sometimes restrains the researchers from freely sharing their scientific results before publication. It is noted that currently the IP attribution before publication is sometimes provided through the use of DOI or open licenses, as GNU or Creative Commons, this should be complemented by practical way and established procedure to implement a staged released of the results.

### Definition of EVER-EST VRE Common Services

The following definition applies: EVER-EST VRE common services are defined as those services that form the basis of the operational VRE and shall ensure that it is efficient, user friendly, stable and interoperable and secure. The EVER-EST VRE needs to comply with best practice security standards, support open and restricted data policies and accommodate property rights and citation.

EVER-COM-010	User Authentication and Session Management	Mandatory
<b>Requirement</b>	The EVER-EST e-infrastructure shall provide Earth Science data users: <ul style="list-style-type: none"> <li>• Authentication and Authorisation: have a reliable user authentication and authorization accessing the VRE;</li> <li>• Federation: enable federated identity management;</li> <li>• User management;</li> <li>• Session management and accounting of resources;</li> <li>• Definition of VRE sub-communities (for authorization to access project restricted resources).</li> </ul>	
<b>Source</b>	Proposal, VRCs (all)	
<b>Verification Method</b>	Review of Design, Test	

EVER-COM-020	Research Object, Digital Management, IPR and Citation	Mandatory
<b>Requirement</b>	The EVER-EST e-infrastructure shall:	



	<ul style="list-style-type: none"> <li>• Allow scientists to define the terms in which their ROs can be used creating confidence on the RO and establishing etiquette for acknowledgement and citation;</li> <li>• Allow dissemination of products and research objects in accordance to specific rights/authorization;</li> <li>• Provide basic access control to the content of the RO, particularly by third parties accessing the RO;</li> <li>• Implement Authorization and IPR protection: need for an authorization mechanism to access data, products or tools, and to maintain IPRs of research results;</li> <li>• Support Citation of ROs vs. citation of papers.</li> <li>• Support tools and procedures for license and PID/DOI attribution to data, research products or documents;</li> <li>• Provide tools for referencing algorithms, products, documents used in each Supersite;</li> </ul>
<b>Source</b>	Proposals, VRC (All)
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Test

EVER-COM-030	Security Requirements	Mandatory
<b>Requirement</b>	The EVER-EST e-infrastructure shall: <ul style="list-style-type: none"> <li>• Comply with best practices and standards for security in terms of the information contained (secure data transfer and) supporting property rights and product data policy.</li> </ul>	
<b>Source</b>	Proposals, VRC Use Case (All)	
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Test	

### 6.2.2 EVER-EST e – collaboration services requirements

#### Rationale for EVER-EST Collaboration Services Requirements: Needs and Challenges

E- Collaboration services are of outmost relevance to foster communities capacity and knowledge building and facilitate sharing of experiences. The as-is situation is currently still relying mainly on physical co-located meeting and teleconference for the various stages and milestones of the cooperation projects. Main challenge for the communities is to assess how the VRE might enhance this collaboration by providing an easy to use and accessible virtual meeting place. Some VRCs are heterogeneous and spread across different countries, with different time-zones and different research capabilities and skills. As a common need, all the communities shall be able to communicate and share knowledge, ideas and protocols on line. Content management tools with collaborative editing functionalities and project management tools (e.g. to share Gantt chart) are deemed necessary to assist in increasing efficiencies within project and to facilitate knowledge capture and improve real-time communication across scientists. It is worth noting that data collection campaigns will also benefit by the use of collaboration services, for the organisation of the overall campaign data acquisition and for the harmonisation of the metadata capturing in line with the best metadata preservation guidelines.

#### Definition of EVER-EST VRE e- Collaboration Services

The following definition applies: EVER-EST VRE E-Collaboration Services support the adoption and sharing of common collaboration management tools by the user communities. These services will provide the



functionalities necessary to assist the users in collaborating within heterogeneous research groups. Synchronous and asynchronous services will enable "same time-different place" and "different time-different place" communication and collaboration mode.

EVER-COL-010	Community Building and Peer Reviews	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall allow Earth Science data users to foster community building by:</p> <ul style="list-style-type: none"> <li>• Integration, open access and sharing of knowledge, free licensed software tools, open data and scientific products;</li> <li>• Enable dissemination of research objects and products in accordance to specific rights/authorization;</li> <li>• Have an appropriate dissemination interface towards the general public/end-users, including a mechanism to provide feedbacks and share results that could be released on a staged approach for peer review and evaluation purpose;</li> <li>• Enable VRCs community directory to be linked to external social networking sites (e.g. LinkedIn, Research-Gate) supporting exchange of information.</li> </ul>	
<b>Source</b>	Proposals, VRE (All)	
<b>Verification Method</b>	Verification Methods: Demonstration, Inspection, Test	

EVER-COL-020	E-Connect Tools	Mandatory
<b>Requirement</b>	<p>E-infrastructure shall foster collaboration and sharing of scientific results through the adoption of collaboration management tools by either integrating open sources tools or supporting the integration of COTS owned and installed by the same VRC to:</p> <ul style="list-style-type: none"> <li>• Connect the scientific and the institutional communities allowing exchange of information;</li> <li>• 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.</li> <li>• 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;</li> <li>• Conduct and present demonstrations using shared data and codes;</li> <li>• 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;</li> <li>• Capture and archiving of discussions including keyword tagging to relate conversations to specific topics, which are searchable.</li> </ul>	
<b>Source</b>	Proposals, VRE Use Case (All)	
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Test	



EVER-COL-030	Content Management Tools	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall integrate a content management tool:</p> <ul style="list-style-type: none"> <li>• For documentation repository with collaboration editing capability for collective document/article writing, including document versioning and provenance metadata capture;</li> <li>• Assuring secure storage and knowledge management with different level of user accessibility rights</li> <li>• Supporting automatic metadata insertion according to the VRE metadata model for document, records, cataloguing and provenance for long-term preservation.</li> <li>• 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 EVE-REST collaborators/EU partners providing information on relevant expertise and sources of help and possibly data, metadata and glossary of terms.</li> </ul>	
<b>Source</b>	Proposal, VRE (All)	
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Test	

EVER-COL-040	Project Management Tools	Highly Desirable
<b>Requirement</b>	<p>EVER-EST e-infrastructure shall provide project planning and management tools to be used by the VRCs communities that includes presentation and editing of Gantt charts, reporting and recording of minutes and actions.</p>	
<b>Source</b>	VRE Use Cases (VM, SS)	
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Test	

### 6.2.3 EVER-EST e – research services requirements

#### Rationale for EVER-EST E-Research Services Requirements: Needs and Challenges

Earth Science scientists' main daily tasks to be supported by the EVER-EST VRE are mainly related to data discovery and data processing. Concerning data discovery, satellite data are mostly accessible through online catalogues and in some cases via web services. The same applies for in situ data, similarly accessible via catalogues, databases and for some VRCs use cases, via web services. Data discovery among various data providers and heterogeneous sources is today a very time consuming task and efficient data access is considered of outmost importance by scientists. The main common data access requirement for all VRCs is therefore to have reliable, automated, fast data discovery and metadata harvesting from a variety of heterogeneous data sources.

Once accessed, data is downloaded and analysed by scientists through SW codes currently installed on the scientist's local desktop, not usable through common online platforms. Some of these codes are openly shared among scientists. Meanwhile other SW is instead subject to commercial licensing schemes, typically those for satellite data analysis (ENVI, ERDAS, SarScope, GAMMA, etc.) or GIS (ESRI products).

After data discovery, the typical next steps work procedure for a scientist are: extract data from a repository, process the data, validate the results against different data or past results, and share the results within the community and/or publish the results for peer review. Depending on the objectives, the processing may be



done using COTS software (e.g. for satellite image processing), or in house software (more common for modelling and simulations). Since the in house software is normally very dynamically evolving, one major challenge to be supported by EVER-EST is to allow the scientists the capacity to easily test, run and modify their codes in the environment before publishing a stable update to be shared with the community.

For requirements grouping purpose, processing needs and challenges have been grouped in the following categories:

- Data processing tools (e.g. to extract geophysical parameters from EO and in situ data);
- Modelling/analysis tools (e.g. used to model /analyse a given phenomena, as volcanic source, volcanic cloud ash density, etc.);
- Validation tools (e.g. algorithms to be used during the validation of the processing results).

The generation of new scientific results at the state of the art is a process involving constant evolution of data analysis methods, processing techniques, modelling methods and procedures. Since the EVER-EST VRE can be appealing to the community only if it provides state-of-the-art processing and analysis tools/services, one major challenge is that the environment architecture shall be designed in a way to allow the scientists to develop, implement, use, verify their models and algorithms. What is needed is in fact an environment very close to the researcher's workstation environment, in which the researcher can upload (possibly develop) his algorithms, carry out tests and experiments using the facilities of the platform, share the results with colleagues, validate the algorithms, and then when the codes are stable enough, publish the codes in a toolbox for the community to use.

Some VRCs need not only to discover the data but also to effectively share their own generated data with their stakeholder or with the others VRCs. VRCs need to provide access to end-users to data generated by in-situ campaign and/or calculated, make available large data set for further model development and validation and web-services for data processing by different partners and apply common semantic and ontologies to facilitate data sharing and preservation among communities.

The VRC itself needs to make the data interoperable in order to ensure data access to the widest number of scientists and end users. Interoperability is a major challenge, given the global scale of the VRCs. The most relevant frameworks to which the data management in EVER-EST should be compliant or interoperable are: the EPOS framework at European level, and the [Common Framework on Earth-Observation Data](#)<sup>7</sup> from USGS and GEO-USA.

The EVER-EST e-infrastructure shall provide access to generated dataset and research object as specified in the use case and detailed in 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:

- 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.
- 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.
- 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.
- 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).

---

<sup>7</sup> <https://www.whitehouse.gov/blog/2015/12/09/improving-access-earth-observations>



### Definition of EVER-EST VRE E- Research Services

The following definition applies: E-Research Services are one of the core EVER-EST Services that enable the VRC users to accomplish their research activities. These services will provide suitable functionalities to support e-research daily activities, ranging from resources discovery to cloud infrastructure as a services, software lifecycle and workflow management, application and processing job management and research object dissemination. E-Research Services include services for Digital Information, Preservation and Research Object dissemination. Digital Information services span from cataloguing, data and text mining to data management and information. Preservation Services support data preservation best practices in Earth Observation according to the preservation stages defined in D4.1 and preservation requirements defined in D5.1. Research object dissemination services are described in detail in D.4.1.

EVER-RE-010	Data and Resources Discovery	Mandatory
<b>Requirement</b>	The EVER-EST e-infrastructure shall allow the Earth Science community to: <ul style="list-style-type: none"> <li>• Retrieve in near real time, and through web services a wide range of satellite and in situ data for the Earth Science community;</li> <li>• Access EO mission datasets in a rapid and efficient way;</li> <li>• Access data using catalogue metadata harvesting tools as well as automatic tools for browsing high volumes of data catalogue;</li> <li>• Support the use of semantics and ontologies developed by the VRCs;</li> <li>• Increase the availability and use of open data and open sources tools</li> </ul>	
<b>Source</b>	VRE Use Cases (All)	
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Test and Review of Design	

EVER-RE-020	Data Access and Storage	Mandatory
<b>Requirement</b>	The EVER-EST e-infrastructure shall support the following data access, storage and needs/ functionalities: <ul style="list-style-type: none"> <li>• Data must be retrievable in near real time through web services;</li> <li>• Local storage of data must be possible, else web services shall be provided to the VRC;</li> <li>• Simple, file system based local data archiving is preferred;</li> <li>• Store data in local environments, exporting them to the community using web services;</li> <li>• Provide research products as web services to the community,;</li> <li>• Archive data and products using simple, file system based archiving;</li> <li>• Associate metadata to all data and products generated by the community within the VRE. Metadata standards must be used;</li> <li>• Use tools and procedures for license and PID/DOI attribution to data, research products or documents;</li> <li>• Use tools for referencing algorithms, products, and documents.</li> </ul>	
<b>Source</b>	VRE Use Case (SM, SS)	
<b>Verification Method</b>	Verification Methods: Inspection, Test	

EVER-RE-030	Validation Data Base	Highly Desirable
-------------	----------------------	------------------





<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall support the integration of validation data base:</p> <ul style="list-style-type: none"> <li>• Validation database represents a single source of observed impact data for validation purposes;</li> <li>• 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);</li> <li>• The database could be made publically available as a single source of impact data to be used for further validation and research.</li> </ul>
<b>Source</b>	VRE Use Case (HM)
<b>Verification Method</b>	Verification Methods: Inspection, Test

EVER-RE-040	Data Sharing	Desirable
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall support data sharing that shall be:</p> <ul style="list-style-type: none"> <li>• To be provided by live data services, if available;</li> <li>• Ready access to static saved data e.g. case study based data to support more effectively building models using readily available 'sample data';</li> <li>• Easier access to large in-situ datasets (e.g. National Population Database, Met Office data);</li> <li>• Access portals to database servers at partner organisations;</li> <li>• Access platforms to geospatial data following common standards (one access point to many diverse data sets);</li> <li>• Provide the sharing of research products and data using web services</li> <li>• Provide the single VRCs with web service applications to be installed locally, to share data and products, which are security-sensitive.</li> </ul>	
<b>Source</b>	VRE Use Case (HM)	
<b>Verification Method</b>	Verification Methods: Inspection, Test	

EVER-RE-050	Research Data Generation and Preservation	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall support preservation and dissemination of Earth Science users generated data set enabling the following functionalities:</p> <ul style="list-style-type: none"> <li>• Annotation of the data set;</li> <li>• Archiving of the data set;</li> <li>• Cataloguing of the data set;</li> <li>• Cross search of the data set;</li> <li>• Data and test mining;</li> <li>• Implementation of dictionaries and ontologies defined by the VRCs;</li> <li>• Information resources discovery;</li> <li>• Packaging;</li> <li>• Data Set Provenance metadata capturing.</li> </ul>	



<b>Source</b>	VRE Use Case (SM, All)
<b>Verification Method</b>	Verification Methods: Inspection, Test, Review of Design

EVER-RE-060	Application Management	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall support the following application management functionalities:</p> <ul style="list-style-type: none"> <li>• Executable building;</li> <li>• Executable management;</li> <li>• Process building;</li> <li>• Application Validation and verification;</li> <li>• Workflow management.</li> </ul>	
<b>Source</b>	VRE Use Case (SM, All)	
<b>Verification Method</b>	Verification Methods: Inspection, Test, Review of Design	

EVER-RE-70	Workflow Management	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall allow Earth Science data users to:</p> <ul style="list-style-type: none"> <li>• Link Research Objects and work-flows/thought-flows together;</li> <li>• Enable documentation of model development e.g. model history;</li> <li>• Facilitate capturing ideas at each stage of model development e.g. ideas/thoughts added to Research Object.</li> </ul>	
<b>Source</b>	VRE Use Case (HM)	
<b>Verification Method</b>	Verification Methods: Inspection, Review of Design	

EVER-RE-080	Data Processing Tools	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall allow Earth Science data users to:</p> <ul style="list-style-type: none"> <li>• Extract data from a repository, process them, validate the results against different data or past results, and share the results within the community;</li> <li>• Process satellite and in situ data using COTS software or in house software;</li> <li>• Easily test, run and modify their codes in the environment before publishing a stable update;</li> <li>• Use a baseline toolset mostly through COTS, provided and configured by the VRCs with their own interactive GUIs;</li> <li>• Process satellite optical data using a basic toolset based on COTS or scientific software, with interactive GUI;</li> <li>• Analyze and model data using volcanic modeling software (including finite element COTS) for ground deformation and volcanic clouds, with interactive GUI, and accounting for computationally intensive calculus;</li> </ul>	



	<ul style="list-style-type: none"> <li>Execute workflows for image data processing and modeling;</li> <li>Process the data where the data resides.</li> </ul>
<b>Source</b>	VRE Use Case (SS, LM)
<b>Verification Method</b>	Verification Methods: Test, Review of Design

EVER-RE-090	Visualisation and Modelling Tool	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall allow Earth Science data users to:</p> <ul style="list-style-type: none"> <li>Deploy and use innovative, interactive and flexible visualisation tools for processing outputs and data;</li> <li>Browse, query, and analyze processing outputs (normally images or maps in raster/vector format) for interpretation/validation purposes, through a web GIS interface;</li> <li>Validate processing outputs using the web GIS interface and embedded analysis tools;</li> <li>Generate, discovery and manage Research Objects through a simple to use, graphical interface, also with geographic tagging capabilities;</li> <li>Integrated modeling tools as defined by the VRCs use cases.</li> </ul>	
<b>Source</b>	Proposals, VRE Use Case (All)	
<b>Verification Method</b>	Verification Methods: Inspection, Test, Review of Design	

EVER-RE-100	Research Environment as IaaS	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall allow Earth Science data users to:</p> <ul style="list-style-type: none"> <li>Customize the VRE research environment as an Infrastructure as a Service environment to be used by the communities to integrate tools and product for their specific needs as specified in the relevant use cases.</li> </ul>	
<b>Source</b>	VRE Use Case (LM, SS)	
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Review of Design	

EVER-RE-110	Research Object Development Environment	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall provide Earth Science data users:</p> <ul style="list-style-type: none"> <li>Collaborative development on Research Objects;</li> <li>Case studies re-run with updated datasets, configurations and parameters;</li> <li>Staged approach to development releases to the operational environment;</li> <li>Enable partners to interact regardless of their pre-selected operating system platform of use (e.g. Linux and Windows).</li> </ul>	



<b>Source</b>	VRE Use Case (HM, All)
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Test, Review of Design

EVER-RE-120	Collaborative Computational Environment	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall allow Earth Science data users to:</p> <ul style="list-style-type: none"> <li>• Place workable scripts (R, Python etc.) for the VRCs and others to use, develop, copy and amend as necessary;</li> <li>• Provision of modules of generic model building tools;</li> <li>• Provide access to modules at different levels of complexity (from detailed process models to simple I/O structures);</li> <li>• Modules should be user generated and accessible through a visual library of tools that is easy for other users to pick from;</li> <li>• This facility should include proper code versioning and software archiving where necessary;</li> <li>• Support multiple scripting languages.</li> </ul>	
<b>Source</b>	VRE Use Case (LM)	
<b>Verification Method</b>	Verification Methods: Inspection, Review of Design	

EVER-RE-130	Data Interoperability and Compatibility	Mandatory
<b>Requirement</b>	<p>Data management in EVER-EST should be interoperable with the EPOS framework at European level, and the <u>Common Framework on Earth-Observation Data</u> from USGS and GEO-USA. Compliancy with GEOSS should be provided.</p>	
<b>Source</b>	Proposals	
<b>Verification Method</b>	Verification Methods: Review of Design	

EVER-RE-140	Big Data Analytics Tool	Desirable
<b>Requirement</b>	<p>The EVER-EST e-infrastructure should support exploiting and integrating of Big Data analytics COTS and open source tools and environments, such as:</p> <ul style="list-style-type: none"> <li>• Analysis of social media;</li> <li>• Large-scale big data processing environment (e.g. Hadoop, Apache Spark);</li> <li>• Other open source software like learning packages from the www (e.g. pluralsight).</li> </ul>	
<b>Source</b>	VRE Use Case (HM)	
<b>Verification Method</b>	Verification Methods: Review of Design	

EVER-RE-150	Preservation Requirements	Mandatory
<b>Requirement</b>	<p>EVER-EST infrastructure shall provide the means for Earth Science data sources (either external data providers or VRCs as data generator) to:</p> <ul style="list-style-type: none"> <li>• Enhance their preservation strategies by facilitating their systematic</li> </ul>	



	<p>monitoring of end-user experiences and their requests for system enhancements.</p> <ul style="list-style-type: none"> <li>• 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;</li> <li>• Manage the research object life cycle preservation process as defined in D.4.1.</li> <li>• 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</li> <li>• 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.</li> </ul>
<b>Source</b>	Proposal, D.4.1, D. 5.1 Preservation Requirements
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Test and Review of Design

EVER-RE-160	Research Objects Dissemination	Mandatory
<b>Requirement</b>	<p>The EVER-EST e-infrastructure shall support the following data dissemination needs/ functionalities:</p> <ul style="list-style-type: none"> <li>• Dissemination of research object as specified in the use case and detailed in EVER-EST Data Management Plan.</li> <li>• Research objects should be disseminated as web services to the community</li> <li>• Metadata should be associated to all data and products generated by the community within the VRC and in capsulated in the dissemination package.</li> </ul>	
<b>Source</b>	Data Management Plan, D 4.1, D 5.1. Preservation Requirements	
<b>Verification Method</b>	Verification Methods: Inspection, Demonstration, Test or Review of Design	

#### 6.2.4 EVER-EST e – learning services requirements

##### Rationale for EVER-EST E- Learning Services Requirements: Needs and Challenges

VRCs community are partially involved in providing pure educational activities to public users, however this services might be exploited to foster community building. On line courses might be organized to effectively share methods and workflow developed by the VRCs for mentoring young scientists, trainees and less experienced technicians coming from regional agencies through user-driven tutorials which might include walk-throughs of visualisation tools and references for stored code.

##### Definition of EVER-EST VRE E- Learning Services

The following definition applies: EVER-EST E-Learning Services support the adoption of the VRE by the user communities and will be exploited to provide training on the full research object life cycle from data collection, data cleansing, analysis, and reproducibility to preservation and information package dissemination. These services will also be used to provide on line training to mentor future scientists to overcome the challenges of increasing data volumes and support their ability to validate, analyse, visualize, store, and curate the information.



EVER-EL-010	E-Learning	Mandatory
Requirement	<p>The EVER-EST e-infrastructure shall support:</p> <ul style="list-style-type: none"><li>• The collaboration in the community to foster capacity building of less experienced scientists;</li><li>• Tools supporting the organization of online courses (e.g. Massive Open On Line Course);</li><li>• Training on the platform usage during dedicated events;</li><li>• User-driven tutorials;</li><li>• Walk-throughs of visualisation tools and references for stored code.</li></ul>	
Source	Proposal, VRE Use Case (All)	
Verification Method	Verification Methods: Review of Design	

## 7 Compliance to the Smart Objectives and Key Performance Indicators

For each of the VRCs specific SMART Objectives defined by the tables below a set of use cases, user scenarios, personas and user actions have been derived to represents archetypical instance of an actor, at least one persona for each group of user profiles, within the Earth Science domains. To envision innovative way of working doable by improving the EVER-EST infrastructure components, continuous training-by-doing activities are being performed since the start of the project with the preselected VRCs acting as target audience and piloting the training approach development and dissemination activities for new users within the communities or potential new communities.

Attendance to conference and workshop and new users communities responses to EVER-EST VRE published results is, in turn, feeding the process of eliciting challenges and requirements common to a broader spectrum of potential new Earth Science communities and driving EVER-EST VRE values propositions to new users and sustainability model. Needs, challenges and opportunities of the pre-selected and potential new communities user profiles are continuously assessed to transform current research working practices limitations and problems into opportunities for boosting innovative working practices for interdisciplinary cross-fertilization and synergy in the various fields of Earth Science and to enable common Findable, Accessible, Interoperable, Usable data management policies and guidelines.

Since the start of the project each VRC has received hands-on training of the VRE components. From these use cases, user stories and learning-by-doing activities like the research object hackathons, a set of consolidated VRE services requirements and FAIR data management, quality preservation and curation requirements have been further improved. User requirements and detailed user actions, are, driving the design and integration of the overall VRE platform components and its verification and validation to ensure that the technical developments are well aligned with the needs of the preselected and potential new VRCs users.

At this stage of the project, VRC requirements and use cases are being deeply integrated in EVER-EST architecture. Up the end of the development phase, further requirements and user stories refinement will be discussed and addressed to ensure that the smart objectives below are fully addressed.

SM_OB#2.1 Validate and demonstrate the VRE functionalities within the Sea Monitoring VRC	
<b>Measured by</b>	Number of users reached (around 100 researchers); increased number of publications of web map layers of species and habitat distribution in space and time in key areas for biodiversity monitoring; increased number of publications on shared and clear protocols assessing a well-defined methodology for biodiversity monitoring.
<b>Achievable</b>	Sea Monitoring community can leverage on a well defined set of data catalogues which will be federated by the EVER-EST VRE and supported by tailored tools for data exploitation and exchange.
<b>Relevant</b>	This use case is an example of inter-disciplinary work in terms of disciplines involved, data and software used (biology, chemistry, physics, earth observation).
<b>Timely</b>	Each VRC will receive proper training on the VRE functionalities during the first phase of the project. The VRC will be able to use the VRE services starting from M18 and eventually request changes and adaptation.

**Table 9 Sea Monitoring VRC Smart Objectives**





<b>SM_OB#2.4 Validate and demonstrate the VRE functionalities within the GEO Supersites VRC</b>	
<b>Measured by</b>	Number of users reached (around 100 researchers) and positive feedback received; increased use of EO/in situ data and science products by users; assessment of VRE usability performances through ad hoc metrics.
<b>Achievable</b>	VRE validation using two INGV Supersites. VRE demonstration by entire GEO Supersite community.
<b>Relevant</b>	Globally distributed community and local Disaster Risk Reduction users will ensure a comprehensive test of research product generation and uptake by users.
<b>Timely</b>	Each VRC will receive proper training on the VRE functionalities during the first phase of the project. The VRC will be able to use the VRE services starting from M18 and eventually request changes and adaptation.

**Table 10 GeoHazard Supersites VRC Smart Objectives**

<b>SM_OB#2.3 Validate and demonstrate the VRE functionalities within the Land Monitoring VRC</b>	
<b>Measured by</b>	Positive usability feedback provided by the majority of users reached within the Land Monitoring (and possibly others) Virtual Research Community members (e.g. 50).
<b>Achievable</b>	Initial scenarios have been based on real Land Monitoring VRC operational scenarios. SatCen will organize appropriate events to collect user feedbacks and will report about these activities both in D3.6 and D2.5.
<b>Relevant</b>	The Land Monitoring use case provides a concrete case of cross-disciplinary interaction between Earth Scientists and Institutional entities.
<b>Timely</b>	Each VRC will receive proper training on the VRE functionalities during the first phase of the project. The VRC will be able to use the VRE services starting from M18 and eventually request changes and adaptation.

**Table 11 Land Monitoring VRC Smart Objectives**

<b>SM_OB#2.2 Validate and demonstrate the VRE functionalities within the Natural Hazard VRC</b>	
<b>Measured by</b>	Number of users reached (around 50 researchers); usability feedback provided by the Natural Hazards Partnership (NHP) members engaged with the development of the Hazard Impact Model (HIM) forecasting system; demonstration of the VRE implementation for the NHP to encourage adoption by the wider hazard forecasting community both in the UK and in Europe; documented examples of new opportunities for collaboration with other agencies engaged with natural hazard forecasting.
<b>Achievable</b>	NHP will organise several opportunities to demonstrate this use case to relevant UK stakeholders with the objective of obtaining feedback on this specific deployment and use of the EVER-EST VRE both from organisations concerned with the forecasting of natural hazard and those involved with disaster response planning and management.



<b>Relevant</b>	The VRE will be used to support the development of specific modules of the Hazard Impact Model forecasting system for the UK e.g. those for flooding, extreme weather events, ground instability (landslides). As part of this use case the NHP will also look for opportunities to use the VRE for knowledge exchange with the Supersites VRC.
<b>Timely</b>	Each VRC will receive proper training on the VRE functionalities during the first phase of the project. The VRC will be able to use the VRE services starting from M18 and eventually request changes and adaptation.

**Table 12 Natural Hazard VRC Smart Objectives**

The VRC's use cases are expected to evolve after the second release of this document and will be further refined in WP3 deliverables during the project life cycle to sustain and increase the project key performance indicator. Compliance to smart objectives key performance indicators will be demonstrate and measured during the validation activities.

KPI Number	Objective / Result	Indicator
<b>1</b>	Establish a VRE for Earth Science	Number of user requirements successfully addressed
<b>2</b>	Establish a VRE for Earth Science	Number of Earth Science Datasets managed by the system
<b>3</b>	Establish a VRE for Earth Science	Number of algorithms managed by the system
<b>4</b>	Establish a VRE for Earth Science	Number of workflows managed by the system
<b>5</b>	Establish a VRE for Earth Science	Facilitated sharing of experiences: number of positive feedbacks from the user community
<b>6</b>	Validate the VRE with four main Virtual Research Communities	Number of additional members of the VRC involved in the project
<b>7</b>	Implement and validate the use of "Research Objects" in Earth Science	Number of ROs implemented in Earth Science
<b>8</b>	Implement and validate the use of "Research Objects" in Earth Science	Number of experiments successfully run by a team different from the one who created the experiment
<b>9</b>	Impact on data providers	Number of additional or improved contacts with Earth Science stakeholders
<b>10</b>	Impact on data providers	Number of hosted data processors activated, thus avoiding input data download by the user
<b>11</b>	Impact on data providers	Number feedbacks regarding end-user experience

**Table 13 Key Performance Indicators**



This project is co-funded by  
the European Union

everest





## Appendix A – Sea Monitoring VRC Use Cases Scenario

### Personas

The following personas have been identified:

Personas	Representing group	Description
<b>Federica F.</b>	Scientist (CNR–ISMAR, Bologna)	Expert user. Creates and manage ROs, she works with Francesco and Valentina.
<b>Francesco D.</b>	Scientist (CNR–ISMAR, Bologna)	Expert user. He manages jellyfish App, validates citizen sightings, he elaborates workflow and data need to elaborate MSF descriptor concerning jellyfish. He works with Federica and Valentina.
<b>Ferdinando B.</b>	Scientist (University of Salento)	Expert user. He collaborates with Federica and Francesco to develop Jellyfish App content and evaluates dataset and results.
<b>Valentina G.</b>	Scientist (CNR–ISMAR, Bologna)	Expert user. Database manager. She works with Federica and Francesco.
<b>Fantina M.</b>	Scientist (CNR–ISMAR, Venezia)	Data provider and potential RO user.
<b>Lorenzo A.</b>	Scientist (CNR–ISMAR, Bologna)	Data provider and potential RO user.
<b>CNR</b>	Public agencies	Data provider and potential RO user.
<b>ARPA</b>	Public agencies	Data provider and potential RO user.
<b>EU</b>	Institutional user	Potential RO user. The UE checks the results in terms of MSF descriptors.
<b>Pippo</b>	Citizen	Pippo uses the Jellyfish App to inform about the presence of jellyfish and accesses the map of risk (meteo meduse).



## User actions

Scenarios	Free text description (max 50 words)
<b>1. Habitat suitability model for CWCs</b>	Federica foglini develops the habitat suitability model for the cold water corals (CWCs) in the Bari canyon.
<b>2. The Citizen science and jellyfish distribution</b>	A crowdsourcing app sponsored by Italian magazine and others different media provides scientific data to study jellyfish. CNR wants to fully exploit within the EVER-EST initiative the app DB potential to generate meaningful indicators in MSF perspective.
<b>3. Trend in the evolution of invasive jellyfish distribution</b>	Starting from Jellyfish sightings, we elaborate data to produce explicit geographical information concerning trend about the evolution and distribution of alien species according with MSF directive descriptors.
<b>4. Habitat suitability model for jellyfish species, study cases <i>Pelagia noctiluca</i> and <i>Velella velella</i></b>	Starting from Jellyfish sightings according with environmental variables we produce suitability model to evaluate habitat extent for <i>Pelagia noctiluca</i> and <i>Velella velella</i> .

### Scenario 1: Habitat suitability model for CWCs

Action ID	Action description	Dependencies	Challenge
<b>CNR_CWC_01</b>	Fantina or Lorenzo collect new data (bathymetry or CWC occurrences/absence data) and load them on the ISMAR ftp site	DS_CNR_01 DS_CNR_02 DS_CNR_03 Infra_CNR_01	
<b>CNR_CWC_02</b>	Valentina updates the CoCoNet database with the new data	GUI_CNR_01 DS_CNR_01 DS_CNR_02 Infra_CNR_02 DS_CNR_16 DS_CNR_17	
<b>CNR_CWC_03</b>	Federica downloads the bathymetry, the occurrences layer and the absence data layer	Infra_CNR_01 DS_CNR_16 DS_CNR_17	
<b>CNR_CWC_04</b>	Federica produces the EGVs	SW_CNR_01 Algo_CNR_01 Algo_CNR_02 Algo_CNR_03 Algo_CNR_04 Algo_CNR_05 Algo_CNR_06	



		DS_CNR_04	
<b>CNR_CWC_05</b>	Federica carries out a statistical analysis to evaluate which EGVs are relevant for the suitability model	SW_CNR_02 DS_CNR_04	
<b>CNR_CWC_06</b>	Federica choose the EVGs useful for the analysis	HI_CNR_01 DS_CNR_04	
<b>CNR_CWC_07</b>	Federica uses the selected EVGs and the occurrences layer to generate the habitat suitability model	SW_CNR_03 DS_CNR_16 DS_CNR_17 DS_CNR_04	
<b>CNR_CWC_08</b>	Federica checks the results	DS_CNR_05 HI_CNR_02	
<b>CNR_CWC_09</b>	Federica creates the new RO	DS_CNR_16 DS_CNR_17 DS_CNR_04 DS_CNR_05	
<b>CNR_CWC_10</b>	CNR and EU access the RO	GUI_CNR_01	

## Scenario 2: The Citizen science and jellyfish distribution

Action ID	Action description	Dependencies	Challenge
<b>CNR_JFD1.1_01</b>	Citizen provides a new sighting using the App about presence absence and outbreaks of jellyfish along the Italian coasts	Infra_CNR_05	
<b>CNR_JFD1.1_02</b>	Francesco needs to know when new sightings were sent out	GUI_CNR_02	
<b>CNR_JFD1.1_03</b>	Francesco checks and validates the new data	SW_CNR_01 DS_CNR_06 HI_CNR_03	
<b>CNR_JFD1.1_04</b>	Francesco processes the data to produce a daily map of distribution for each species sighted	SW_CNR_01 DS_CNR_06 DS_CNR_09	
<b>CNR_JFD1.1_05</b>	Francesco processes the data to produce map of distribution for each species to inform Institutional users (D1.1 MSF)	SW_CNR_01 DS_CNR_10	
<b>CNR_JFD1.1_06</b>	Francesco sends the new data to Valentina	DS_CNR_09 DS_CNR_10 DS_CNR_06 Infra_CNR_01	
<b>CNR_JFD1.1_07</b>	Valentina recives results and updates the CoCoNet	GUI_CNR_01	



	database with new validated sightings	Infra_CNR_01 DS_CNR_06 DS_CNR_09 DS_CNR_10 Infra_CNR_02 DS_CNR_07	
<b>CNR_JFD1.1_07</b>	Federica finalizes the results and updates it on Institutional users interface	DS_CNR_10 GUI_CNR_01	

### Scenario 3: Trend in the evolution of invasive jellyfish distribution

Action ID	Action description	Dependencies	Challenge
<b>CNR_JFD2.1_01</b>	Francesco needs to access jellyfish dataset on CoCoNet DB	DS_CNR_07 Infra_CNR_02	
<b>CNR_JFD2.1_02</b>	Francesco and Federica elaborate sightings to produce density raster map for each invasive species for each year	SW_CNR_01 DS_CNR_11	
<b>CNR_JFD2.1_03</b>	Francesco and Federica analyse density rasters and made a quantitative analysis to represent change in time distribution of each alien/invasive species	SW_CNR_01 DS_CNR_11 DS_CNR_12	
<b>CNR_JFD2.1_04</b>	Valentina recives results and produces the final map (D2.1)	GUI_CNR_01 Infra_CNR_01 DS_CNR_11 DS_CNR_12 Infra_CNR_02 SW_CNR_01 DS_CNR_13	
<b>CNR_JFD2.1_05</b>	Federica finalizes the results and updates it on the Institutional users interface	DS_CNR_13 GUI_CNR_01	

### Scenario 4: Habitat suitability model for jellyfish species, study case *Pelagia noctiluca* and *Velella velella*.

Action ID	Action description	Dependencies	Challenge
<b>CNR_JFD2.2_01</b>	Francesco need to access the CoCoNet DB to retrieve the jellifish occurrences	Infra_CNR_02 DS_CNR_07	
	Francesco selects the occurrences of <i>Pelagia noctiluca</i>	DS_CNR_14	





	and <i>Velella velella</i>		
	Francesco need to access the Copernicus portal to retrieve the EGVs	Infra_CNR_04 DS_CNR_08	
<b>CNR_JFD2.2_03</b>	Francesco loads the EGVs on the ftp site	DS_CNR_08 Infra_CNR_01	
<b>CNR_JFD2.2_04</b>	Francesco and Federica perform a statistical analysis using Maxent and R	SW_CNR_02 SW_CNR_03 DS_CNR_14 DS_CNR_08 DS_CNR_05	
<b>CNR_JFD2.2_05</b>	Ferdinando receives the results from Federica and checks it	DS_CNR_05 HI_CNR_02	
<b>CNR_JFD2.2_06</b>	Francesco and Federica receive results after Ferdinando check it and refine them	DS_CNR_05 SW_CNR_01	
	Francesco loads the results on the ftp site	DS_CNR_05 Infra_CNR_01	
<b>CNR_JFD2.2_07</b>	Valentina receives results and produces the final map (D2.2 in MSF)	GUI_CNR_01 Infra_CNR_01 DS_CNR_05 DS_CNR_15	
<b>CNR_JFD2.2_08</b>	Federica finalizes the results and updates it on the Institutional users interface	DS_CNR_15 GUI_CNR_01	

## Dependencies

### Datasets/ Products

N°	Name	Source	Volume	Format
<b>DS_CNR_01</b>	Bathymetry		80 MB	Ascii file
<b>DS_CNR_02</b>	CWCs Occurrences		200 KB	.shp, .csv
<b>DS_CNR_04</b>	EVGs for CWCs		250 MB	Rater, Ascii, netCDF
<b>DS_CNR_05</b>	Habitat suitability map		80 MB / 2MB	Ascii file / Image
<b>DS_CNR_06</b>	Jellyfish occurrences		200 KB	.shp, .csv
<b>DS_CNR_07</b>	CoCoNet jellyfish occurrences		200 KB	.shp, .csv
<b>DS_CNR_08</b>	EVGs for jellyfishes		10T	Rater, Ascii, netCDF



<b>DS_CNR_09</b>	Daily map of distribution for each species of jellyfish			
<b>DS_CNR_10</b>	Map of distribution for each species of jellyfish			
<b>DS_CNR_11</b>	Density raster map for each invasive species of jellyfish by year			
<b>DS_CNR_12</b>	Time distribution rasters			
<b>DS_CNR_13</b>	Final distribution maps			
<b>DS_CNR_14</b>	<i>Velella</i> and <i>Pelagia</i> CoCoNet occurrences			
<b>DS_CNR_15</b>	Final suitability map of jellyfishes			
<b>DS_CNR_16</b>	CoCoNet CWCs Occurrences			
<b>DS_CNR_17</b>	CoCoNet bathymetry			

### Infrastructures/catalogues

<b>Infrastructure/ Catalogue Name</b>	<b>Infrastructure/ Catalogue Description</b>	<b>Access policy</b>	<b>Com. Protocol</b>	<b>Contacts</b>
<b>Infra_CNR_01</b>	ISMAR ftp site	Password protected		Valentina G.
<b>Infra_CNR_02</b>	CoCoNet WebGIS	Data policy acceptance		Valentina G.
<b>Infra_CNR_03</b>	Emodnet	Open access		Francesco D. L., Federica F., Valentina G.
<b>Infra_CNR_04</b>	Copernicus Marine environment monitoring service online catalogue	Open access		Francesco D. L.
<b>Infra_CNR_05</b>	Jellyfish APP	Free		Francesco De Leo



## Software requirements

N°	Name	Description	Open source/ IPR	Contacts
SW_CNR_01	ArcGIS	GIS software	IPR	
SW_CNR_02	R		Open	
SW_CNR_03	Maxent	The software is based on the maximum-entropy approach for species habitat modeling. It takes as input a set of layers or environmental variables (such as elevation, precipitation, etc.), as well as a set of georeferenced occurrence locations, and produces a model of the range of the given species	Open	
SW_CNR_04	QGIS	GIS software	Free and open	
SW_CNR_05	ODV	Ocean Data View (ODV) is a desktop software for analysis and visualization of oceanographic, atmospheric and other geo-referenced profile or time-series data	Open	

## Algorithms

N°	Name	Description	Sw Used/Code type	Interaction
Algo_CNR_01	Terrain Ruggedness		QGIS	
Algo_CNR_02	Surface Area to Planar Area		ArcGIS	
Algo_CNR_03	Calculate Slope		QGIS	
Algo_CNR_04	Build Fine Scale BPI / Topographic position index		ArcGIS/ QGIS	
Algo_CNR_05	Curvature		QGIS	
Algo_CNR_06	Calculate Statistical Aspect		QGIS	



## User stories-actions GUI requirements

Specific GUI Requirement	Free text description/Image
GUI_CNR_01	Update alert
GUI_CNR_02	Sighting alert

## Control points

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
HI_CNR_01	Selection of the variables	Yes: the variable is necessary No: the variable is not necessary	Yes: the user inserts the variable in the model No: the user doesn't insert the variable in the model
HI_CNR_02	Validation of the results	Yes: the products are acceptable No: the products are not acceptable	Yes: the user loads the new products in the CoCoNet database No: the user rejects the products
HI_CNR_03	Validation of the data	Yes: the data are acceptable No: the data are not acceptable	Yes: the user loads the new data in the CoCoNet database No: the user rejects the data



## **Appendix B – Supersites VRC Use Case Scenario**

The following scenarios are related to the Supersites VRC Use Case:

1. The INGV-UniBo InSAR investigation (Use case 1)
2. The water vapour content investigation (Use case 4)
3. The ash cloud investigation (Use case 3)
4. The generic scientific data analysis investigation (Use case 2)
5. The Volcano Source Modelling investigation (Use case 2)
6. The generic reference search generation

Each of the user actions description has the following layout

<b>Action ID</b>	<b>Action description</b>	<b>Dependencies</b>	<b>Challenge</b>
<b>Sx_US_INGV_01</b>			

***Note to Action ID: the initial code Sx is used to distinguish between the 4 different scenarios. This was done to avoid confusion on the dependencies. Now each Scenario has a list of dependencies, which is in many cases different from the other Scenarios.***

### **Personas**

<b>Personas</b>	<b>Representing group</b>	<b>Description</b>
<b>Elisa</b>	INGV – Roma	Geophysicist. Creates RO's. Works on data with Cristiano
<b>Cristiano</b>	INGV – Roma	SAR processor. Works on data with Elisa
<b>Maurizio</b>	Univ. Bologna	Interprets data and publish papers with Elisa and Cristiano
<b>Claudia</b>	INGV – Roma	Expert user. Creates RO's. Works on data with Massimo and Vito
<b>Massimo</b>	INGV – Catania	Works on data with Claudia and Vito
<b>Vito</b>	INGV – Roma	Works on data with Claudia and Massimo
<b>Francesco</b>	INGV – Catania	Interprets data and uses the output to publish with Claudia, Vito and Massimo
<b>Simone</b>	INGV - Roma	Modeler. He writes data analysis/modeling algorithms, tests them on the VRE VM, then he creates an interface and publishes them for other users.
<b>Giuliana</b>	INGV-Roma	Contributes to Supersite community in various ways. She writes bibliographic ROs. Works on ROs with Stefano, Elisa and Vito.



## Detailed Use Case scenarios

Scenarios	Free text description (max 50 words)
<b>1. The Mt.Etna ground deformation investigation</b>	INGV and the University of Bologna work together on the data analysis and interpretation of the ongoing ground deformation of Mt. Etna. INGV processes data and returns to Univ. Bologna for data modelling/interpretation and final result for publication.
<b>2. The water vapour content investigation on Mt.Etna</b>	INGV-Roma and INGV-Catania work together on the Precipitable Water Vapour analysis of Mt. Etna. INGV-Roma processes data and returns the intermediate product to INGV-Catania for data modelling and final result validation and RO publication.
<b>3. The ash cloud investigation on Mt.Etna</b>	A researcher from INGV runs a code for calculating maps of quantities which characterise the volcanic cloud. The results are displayed in the map browser and validated against different data (in situ/EO). The resulting RO is shared within the community and/or with the end-users and/or the public.
<b>4. The generic scientific data analysis investigation</b>	Stefano is a researcher from the Supersite community. He uploads an algorithm in the VRE VM, he runs it with test or real data, he validates the results and the procedure, then he writes an interface to allow other users to run his algorithm in the VRE. Elisa then runs the code without the need to access the VM.
<b>7. The Campi Flegrei Volcanic Source Modelling</b>	Cristiano obtained a robust deformation map of the Campi Flegrei unrest during 2011-2013. Starting from these data, Elisa proceeds to infer the volcanic source responsible of the observed deformation using the Volcano Source Modeling (VSM) tool. The results of the processing are the source parameters (position, depth, potency), the modelled data and the residuals (observed minus modelled data).
<b>8. The generic reference search</b>	Giuliana <sup>8</sup> performs a reference search <sup>9</sup> on selected repositories <sup>10</sup> , on a given subject, e.g. a Supersite and acquisition technique or investigation method <sup>11</sup> . She stores the queries in a RO <sup>12</sup> made available via a link/button in the

<sup>8</sup> Giuliana or another contributor in the Supersite community.

<sup>9</sup> The search aims to make a collection of materials which includes not only publications but also other resources like reports and citations for example. Hence the reason for searching inside more than one repository.

<sup>10</sup> Repositories are open access and/or may require an account to access the papers.

<sup>11</sup> Example of reference search: search for “InSAR and Mount Etna” in Google Scholar. Geographic search like that of Pangea ([www.pangea.de](http://www.pangea.de), see Advanced search with geographic search coverage) could be interesting and eventually explored.

<sup>12</sup> In case of supervision needed, for example in the case of bibliographic search performed by a student, an early career researcher, or a generic collaborator supporting other researchers, the RO is first created “temporary” and, once validated by a Supervisor belonging to the Supersite community, it becomes “permanent”.



VRE's interface. Another researcher runs the search without the need to reprocess it.<sup>13</sup>

## Scenario 1: The Mt.Etna ground deformation investigation

User actions for Scenario 1

Action ID	Action description	Dependencies	Challenge
<b>S1_US_INGV_01</b>	Cristiano logs in the VRE - Supersite portal, and selects the Mt.Etna Supersite as his area of interest.		
<b>S1_US_INGV_02</b>	Cristiano discovers all the ERS1-2 passes over Mt. Etna, ascending and descending orbits, using the discovery service in the VRE portal. He selects the data of interest through spatial and temporal queries.	S1_Infra_INGV_01 S1_SW_INGV_01	
<b>S1_US_INGV_03</b>	Cristiano wants to place the dataset in a special area provided by the VRE, from where the SarScape COTS will read the data. The dataset S1_DS_INGV_01 is created.	S1_DS_INGV_01	
<b>S1_US_INGV_04</b>	Elisa and Maurizio (who are registered as members of the same working group of Cristiano) are automatically notified that Cristiano has created a new dataset of the Mt Etna volcano.		
<b>S1_US_INGV_05</b>	Cristiano discovers & downloads the DEM of Mt. Etna from a web catalogue. Data S1_DS_INGV_02 created in the same area as the ERS datasets.	S1_Infra_INGV_02 S1_Infra_INGV_03 S1_DS_INGV_02	
<b>S1_US_INGV_06</b>	Cristiano opens a VM to perform InSAR analysis with the COTS Sarscape. He needs a VM with limited access to his research group, including Elisa and Maurizio. He launches the SarScape SW and starts the analysis. The SarScape run begins with input data selection and parameter definition, and is composed of some steps after which Cristiano may want to display the intermediate results in the map browser or in a graphic window of the COTS (ENVI). All SarScape outputs are stored in the same VRE storage area where the input data are located, using a specific directory tree.	S1_SW_INGV_02 S1_DS_INGV_01 S1_DS_INGV_02 S1_DS_INGV_03 S1_GUI_INGV_01	

<sup>13</sup> In general terms, one added value consists in the fact that the stored query is “smart”, i.e. embeds the expertise of the researcher who has refined the query string after several validation cycles upon retrieved results.





	There is no need to create a RO for each of the intermediate files.		
<b>S1_US_INGV_07</b>	Cristiano opens an ENVI window to check the intermediate SarScape results (e.g. the several tens of single interferograms created for the time series analysis). Depending on his expert judgement he may want to modify the SarScape previous or following steps, e.g. he may delete some of the interferograms files before proceeding to the next step, or delete all the intermediate results and start all over again with the same data but new parameters, or just continue to the next step, in which he will input more parameters.	S1_SW_INGV_02 S1_DS_INGV_03	
<b>S1_US_INGV_08</b>	<p>Cristiano ends the SarScape run. The final result is several tens of files stored in the same area of the input data. Dataset S1_DS_INGV_03.</p> <p>A RO is automatically created including reference to all the input and output files. The SarScape log file is embedded in the RO. Possibly also the workflow.</p> <p>Within the RO the most important files (the final results) are a raster file of the mean ground velocities and a vector file of the ground displacement at each coherent pixel.</p> <p>Cristiano checks that the temporary RO has been correctly created with all the needed information.</p>	S1_DS_INGV_03 S1_RO_INGV_01       S1_HI_INGV_01	
<b>S1_US_INGV_09</b>	<p>Maurizio and Elisa (the working group) receive notification that a temporary RO has been created for the Mt Etna deformation processing. The notification has a link by which they open directly the results in the map browser or in the ArcMap or ENVI windows.</p> <p>They upload and display in the chosen window other data or results they need for the validation. The latter are discovered from other catalogues (UNAVCO SSARA, Supersite catalogues, the GEP catalogue, the EVER-EST catalogue, etc.).</p> <p>They carry out the validation of the final results in the RO, by comparing the results against different datasets or previous results (e.g. from past ROs), or by performing statistical analysis.</p>	S1_SW_INGV_03 S1_DS_INGV_04 S1_SW_INGV_04 S1_GUI_INGV_02 S1_GUI_INGV_03	
<b>S1_US_INGV_10</b>	If the validation process is satisfactory Maurizio and Elisa approve the final results of S1_RO_INGV_01, they convert its status into permanent. They may want to edit the RO and add other information into it.	S1_RO_INGV_02    S1_HI_INGV_02	



	S1_RO_INGV_02 is created and S1_RO_INGV_01 is deleted. The scenario is ended		
<b>S1_US_INGV_11</b>	If the validation process is not satisfactory (e.g. poor accuracy, high residuals, too much noise, etc.) Maurizio and Elisa reject the final results of S1_RO_INGV_01. They notify Cristiano and the scenario should restart from S1_US_INGV_06. S1_RO_INGV_01 can be deleted or kept in the database for reference.	S1_RO_INGV_01  S1_HI_INGV_02	

## Dependencies for Scenario 1

### Datasets/ Products for Scenario 1

N°	Name	Source	Volume	Format
<b>S1_DS_INGV_01</b>	SAR images	ERS 1-2 images downloaded from ESA catalogue (several tens of images for each processing instance)	5GB	Raw or SLC
<b>S1_DS_INGV_02</b>	DEM	Accessed and Downloaded from the Etna Supersite data infrastructure, or from SRTM <a href="http://srtm.csi.cgiar.org">srtm.csi.cgiar.org</a> , or from other web services or file systems.	300Mb	Raster
<b>S1_DS_INGV_03</b>	SarScape InSAR results (include tens of different data files also for intermediate results)	Result of the SarScape SW run.	5GB	Raster
<b>S1_DS_INGV_04</b>	Datasets for the validation of the processing results	Data and product catalogues (UNAVCO SSARA, Supersite catalogues, the GEP catalogue, the EVER-EST catalogue, etc.).	50-500 Mb	Raster and vector

### Research Objects for Scenario 1

N°	Name	Source	Description	Format
<b>S1_RO_INGV_01</b>	Non-validated ground deformation of	SarScape processing	Contains input data, and intermediate and output results of the Sarscape processing. Should contain the	Temporary



	Mt.Etna		Sarscape log file. Contains the workflow if possible. Contains some textual description explaining the knowledge used for the Sarscape processing.	
<b>S1_RO_INGV_02</b>	Validated ground deformation of Mt.Etna	SarScape processing + validation	Contains all information listed above plus the results of the validation procedure and some textual description explaining the knowledge used for the entire scientific process.	Permanent

### Infrastructures/catalogues for Scenario 1

Infrastructure/ Catalogue Name	Infrastructure/ Catalogue Description	Access policy	Com. Protocol	Contacts
<b>S1_Infra_INGV_1</b>	ESA Catalogue	Open access	Ask ESA	Cristiano T.
<b>S1_Infra_INGV_2</b>	srtm.csi.cgiar.org	Open access	Ask USGS	NA
<b>S1_Infra_INGV_3</b>	MEDSUV catalogue of in situ data available for the Etna Supersite.	Open access	Ask INGV Catania	Giuseppe P., INGV Catania

### Software requirements for Scenario 1

N°	Name	Description	Open source/ IPR	Contacts
<b>S1_SW_INGV_01</b>		Software to browse and download images from the ESA catalogue. Mass download is required for InSAR dataset creation.		ESA
<b>S1_SW_INGV_02</b>	Sarscape	Software to perform InSAR analysis. Requires ENVI/IDL.	Commercial	Cristiano T.
<b>S1_SW_INGV_03</b>	GIS	ArcMap and/or QGIS software for data visualization/analysis	Commercial	
<b>S1_SW_INGV_04</b>	Statistical tools	Software to perform some statistical computations on the obtained results (e.g. statistics of residuals, boolean operations, spatial profiling, spatial gradient analysis, time series spatial averaging, time series linear-non linear fitting, etc.).	Commercial/INGV	Cristiano T.



## Algorithms for Scenario 1

N°	Name	Description	Sw Used/Code type	Interaction
S1_Algo_INGV_01	SBAS	InSAR time-series processing algorithm, providing ground motion analysis of surface elements, based on Berardino et al. (2002)	Sarscape COTS	high
S1_Algo_INGV_02	PS	InSAR time series processing algorithm providing motion analysis on single scatterers, based on Ferretti et al. (2001)	Sarscape COTS	High

## User stories-actions GUI requirements for Scenario 1

Specific GUI Requirement	Free text description/Image
S1_GUI_INGV_01	<p>A graphic window should be available to inform the user where the input data have been placed (for the processing SW to use them) and at the end of the processing the same window should display the location and directory tree of the final results.</p> <p>The user must see and use the ENVI/Sarscape interface launched within a graphic linux/windows VM.</p> <p>A notification window appears in the VRE once the COTS processing is over. The user shall be able to interact (display/download/validate, etc.) with the results obtained by Sarscape, in the map browser or in ENVI or in Arcmap/QGis.</p>
S1_GUI_INGV_02	<p>The user wants to discover the temporary and permanent ROs in the VRE map browser. The GUI should allow the user to query (spatial, temporal, author, content, etc.) the RO database of the VRC, display the ROs as symbols on the map, select them and display their content, edit the RO, save it as a permanent RO, or delete it. If the RO is saved, a list of people (the working group) should be notified via email. Something that graphically notifies the creation of a new RO (temporary or permanent) should be visible in the map browser and in the information windows of the main GUI.</p>
S1_GUI_INGV_03	<p>The User will use the map browser or a GIS interface (in the VM) to validate the results, e.g., load a previously computed dataset, open other kind of data, perform some statistical analysis. Once the results are validated the temporary RO should be saved or discarded (see above).</p>

## Control points for Scenario 1

Only control points external to the Sarscape processing are listed.

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
S1_HI_INGV_01	User checks the	YES: User verifies	YES: Press OK button and and



	temporary RO content. Uses GUI_INGV_02	completeness of temporary RO. NO: The temporary RO was not correctly created and is incomplete.	publishes the temporary RO NO: manually edits the temporary RO and fills the gaps. Then GOTO YES
<b>S1_HI_INGV_02</b>	User validates the final results. Uses GUI_INGV_03	YES: The final velocity map and displacement time series are validated NO: The final velocity map and displacement time series are not validated	YES: Press OK button and converts the temporary RO into permanent (an editing of the RO content may be necessary using GUI_INGV_02). NO: The Sarscape processing is restarted from the beginning or from an intermediate step within action US_INGV_05

## Scenario 2: The IPWV map generation on Mt.Etna

### Scenario 2a: The IPWV map generation

Action ID	Action description	Dependencies	Challenge
<b>S2_US_INGV_01</b>	Vito logs in the VRE - Mt. Etna supersite portal.		
<b>S2_US_INGV_02</b>	Vito surfs the map browser to the area of interest		
<b>S2_US_INGV_03</b>	He chooses to open the IPWV algorithm <sup>14</sup> among a list of available algorithms.	S2_SW_INGV_01	
<b>S2_US_INGV_04</b>	Vito discovers <sup>15</sup> the following input datasets using the discovery service in the VRE portal: - the Sentinel 1 passes over Mt. Etna. One dataset (i.e. day and time of a single pass) - GPS, MODIS and Grib data (one dataset each) over the same date/time	S2_Infra_INGV_01 S2_Infra_INGV_02	
<b>S2_US_INGV_05</b>	Vito downloads the selected dataset in the VRE and the data files S2_DS_INGV_01 S2_DS_INGV_02, S2_DS_INGV_03 and S2_DS_INGV_04 are created for Sentinel 1, GPS, MODIS and Grib data respectively.	S2_DS_INGV_01 S2_DS_INGV_02 S2_DS_INGV_03 S2_DS_INGV_04	
<b>S2_US_INGV_06</b>	Vito runs the IPWV code by clicking on a button: the file S2_DS_INGV_05 and the log file are created and stored in the VRE.	S2_GUI_INGV_01 S2_SW_INGV_01 S2_DS_INGV_05	
<b>S2_US_INGV_07</b>	Vito checks / visually inspects the log file and verifies		

<sup>14</sup> The IPWV algorithm is already integrated in the VRE portal and a WPS(?) interface is available.

<sup>15</sup> The Area of Interest corresponds to the area in the map browser defined in step S2\_US\_INGV\_02



	that the resulting file has a quality sufficient for the rest of the steps (e.g. sufficient number of GPS stations with data, excessive cloud cover in MODIS imagery, etc.).		
<b>S2_US_INGV_08</b>	If the quality check fails Vito deletes the output files and may start again another trial.	S2_HI_INGV_01 S2_DS_INGV_05	
<b>S2_US_INGV_09</b>	If Vito approves the quality check of S2_DS_INGV_07, and a RO is automatically created (it is a temporary RO which may become permanent after the validation, see later). The VRC use case members are notified by email (and/or other means) that there is a new IPWV RO for Mt Etna from Vito. [The notification has a link by which VRC use case members can open directly the result.]	S2_HI_INGV_01 S2_RO_INGV_01	
<b>S2_US_INGV_10</b>	Claudia logs in the VRE and jumps to the Mt.Etna Supersite contents		
<b>S2_US_INGV_11</b>	Claudia opens and visualizes the RO and the final output results within the VRE map browser or in an Arcmap/QGis interface launched within the VM.	S2_GUI_INGV_02 S2_RO_INGV_01 S2_DS_INGV_05 S2_SW_INGV_02	
<b>S2_US_INGV_12</b>	Claudia discovers and uploads in the validation interface of her choice, other layers needed for the validation.  She carries out the validation of the final results in the RO, by comparing the results against different datasets or previous results (e.g. from past ROs), or by performing statistical analysis.	S2_GUI_INGV_03 S2_SW_INGV_03	
<b>S2_US_INGV_13</b>	If Claudia validates the final results, she changes the status of the RO from temporary to permanent. Claudia may change the access rights to the RO. The VRC use case members are notified that the validation is passed.  In addition to the permanent RO creation Claudia expects to have the dataset S2_DS_INGV_05 stored in a catalogue (internal to the VRE or external).	S2_GUI_INGV_02 S2_HI_INGV_02 S2_RO_INGV_01 S2_RO_INGV_02 S2_DS_INGV_05	
<b>S2_US_INGV_14</b>	If the validation is not passed Claudia notifies the VRC use case members that the validation has failed. The permanent RO is not created. The temporary RO can be kept or deleted with all its content.	S2_HI_INGV_02 S2_RO_INGV_01 S2_DS_INGV_05	

## Scenario 2b: The refractivity tomography map



<b>S2_US_INGV_14</b>	Massimo logs in the VRE portal, jumps to the Mt.Etna map and contents, and discovers the existing ROs of IPWV products.	S2_RO_INGV_02	
<b>S2_US_INGV_15</b>	Massimo accesses the RO of the product IPWV of Mt. Etna of a certain date and wants to use it to compute the refractivity of the atmosphere using a tomography algorithm (S2_SW_INGV_04). He opens the Tomography algorithm within a VM which allows graphical input from a Matlab interface.	S2_GUI_INGV_02 S2_GUI_INGV_01 S2_RO_INGV_02 S2_SW_INGV_04	
<b>S2_US_INGV_16</b>	From the Matalab interface Massimo selects the type of tomography (wet or wet and dry, with default values) and runs the algorithm.	S2_GUI_INGV_01 S2_SW_INGV_04	
<b>S2_US_INGV_17</b>	Massimo selects a series of other parameters from a series of Matlab interfaces and obtains the refractivity tomography product (S2_DS_INGV_06). A temporary RO is created. The VRC use case members are notified that a new temporary RO exists.	S2_GUI_INGV_01 S2_SW_INGV_04 S2_DS_INGV_06 S2_RO_INGV_03	
<b>S2_US_INGV_18</b>	Claudia opens and visualizes the S2_DS_INGV_06 within the VRE map browser or in an Arcmap/QGis interface launched within the VM. Claudia starts the validation of the S2_DS_INGV_06 results using some statistical analyses.	S2_GUI_INGV_02 S2_DS_INGV_06 S2_SW_INGV_03 S2_RO_INGV_03	
<b>S2_US_INGV_19</b>	If Claudia validates the final results, she changes the status of the RO from temporary to permanent (it becomes S2_RO_INGV_04). Claudia may change the access rights to the RO. Claudia may add some textual information to the RO. The VRC use case members are notified that the validation is passed.  In addition to the permanent RO creation Claudia expects to have the dataset S2_DS_INGV_06 stored in a catalogue (internal to the VRE or external) and exposed through web services.	S2_GUI_INGV_02 S2_GUI_INGV_03 S2_HI_INGV_03 S2_RO_INGV_03 S2_RO_INGV_04 S2_DS_INGV_06	
<b>S2_US_INGV_20</b>	If the validation is not passed Claudia notifies the VRC use case members that the validation has failed. The permanent RO is not created. The temporary RO can be kept or deleted with all its content.	S2_GUI_INGV_02 S2_HI_INGV_03 S2_RO_INGV_03 S2_DS_INGV_06	

## Scenario 2c: The Line of Sight refractivity tomography map

Action ID	Action description	Dependencies	Challenge
<b>S2_US_INGV_21</b>	Claudia logs in and browses the map to search the existing permanent ROs of refractivity tomography products for the Mt.Etna, for the period of interest.		





<b>S2_US_INGV_19</b>	Claudia accesses the RO of a certain date and wants to use it to compute the SAR LOS map. Claudia opens the VM and launches the tropospheric SAR LOS delay algorithm that reads the previously created tomography result (S2_DS_INGV_06) automatically.	S2_GUI_INGV_01 S2_DS_INGV_06 S2_SW_INGV_05	
<b>S2_US_INGV_20</b>	Claudia obtains the SAR Line of Sight MAP delay product and a permanent RO is automatically created (there is no need for the approval procedure so no need for a temporary RO). S2_DS_INGV_07 is created also as a single file. Claudia may change the access rights to the RO. The VRC use case members are notified that the validation is passed.	S2_GUI_INGV_02 S2_DS_INGV_07 S2_RO_INGV_05	
<b>S2_US_INGV_21</b>	Francesco logs in and discovers the RO in the map browser. He may want to select other RO from past processing for comparison. To compare the ROs he opens the Arcmap interface and loads previous results. He is able to save or discard the RO.	S2_GUI_INGV_02 S2_GUI_INGV_03 S2_DS_INGV_07 S2_SW_INGV_02 S2_SW_INGV_03	

## Scenario 2d: Browsing the water vapour ROs

Action ID	Action description	Dependencies	Challenge
<b>S2_US_INGV_22</b>	Francesco logs in and discovers the water vapour ROs in the map browser. He may want to select various ROs from past processing for comparison. To compare the ROs he opens the Arcmap interface, loads all layers needed and compares them.  He may want to edit the RO and add some textual information on knowledge information.	S2_GUI_INGV_02 S2_GUI_INGV_03 S2_DS_INGV_05 S2_SW_INGV_02	

## Dependencies for scenario 2a-d

### Datasets/ Products for scenario 2a-d

N°	Name	Source	Volume	Format
<b>S2_DS_INGV_01</b>	Sentinel 1	ESA SciHub	1 Gb	
<b>S2_DS_INGV_02</b>	GPS	Accessed and Downloaded from the Etna Supersite data infrastructure	50 MB	Raster, Ascii and .mat file
<b>S2_DS_INGV_03</b>	MODIS	Accessed and Downloaded from the Etna Supersite data infrastructure	10 MB	Raster and Ascii file



<b>S2_DS_INGV_04</b>	GRIB	Accessed and Downloaded from the Etna Supersite data infrastructure	10 MB	Raster and Ascii file
<b>S2_DS_INGV_05</b>	IPWV product	Result of the IPWV analysis	200 MB	Raster and Ascii file
<b>S2_DS_INGV_06</b>	tomography product	Result of the tomographic analysis	200 MB	Raster and Ascii file
<b>S2_DS_INGV_07</b>	SAR Line of Sight MAP delay product	Result of the analysis	200 MB	Raster and Ascii file

### Research Objects for scenario 2a-d

<b>N°</b>	<b>Name</b>	<b>Source</b>	<b>Volume</b>	<b>Format</b>
<b>S2_RO_INGV_01</b>	IPWV processing	IPWV code + VRE	Contains input data and output results of the IPWV processing. Should contain the IPWV log file. Contains the workflow if possible. Contains some textual description explaining the knowledge used for the processing.	Temporary
<b>S2_RO_INGV_02</b>	IPWV processing	IPWV code + VRE + user edits	Contains all information listed above plus the results of the validation procedure and some textual description explaining the knowledge used for the entire scientific process.	Permanent
<b>S2_RO_INGV_03</b>	Tomography processing	Tomography code + VRE	Contains input data and output results of the tomographic processing. Should contain the tomography code log file. Contains the workflow if possible. Contains some textual description explaining the knowledge used for the processing.	Temporary
<b>S2_RO_INGV_04</b>	Tomography processing	Tomography code + VRE + user edits	Contains all information listed above plus the results of the validation procedure and some textual description explaining the knowledge used for the entire scientific process.	Permanent



<b>S2_RO_INGV_05</b>	SAR Line of Sight processing	Tropospheric SAR LOS delay + VRE + user edits	Contains input data and output results of the SAR Line of Sight processing. Should contain the Tropospheric SAR LOS code log file. Contains the workflow if possible.	Permanent
----------------------	------------------------------	---	---	-----------

### Infrastructures/catalogues for scenario 2a-d

Infrastructure/ Catalogue Name	Infrastructure/ Catalogue Description	Access policy	Com. Protocol	Contacts
<b>S1_Infra_INGV_01</b>	ESA Catalogue	Open access	Ask ESA	Cristiano T.
<b>S2_Infra_INGV_02</b>	MEDSUV (Etna Supersite) catalogue of in situ data available for the Supersite.	Open access	Ask INGV Catania	G. P., INGV Catania

### Software requirements for scenario 2a-d

N°	Name	Description	Open source/ IPR	Contacts
<b>S2_SW_INGV_01</b>	IPWV	Software based on the IPWV Algorithm. Written in Matlab.	INGV	Claudia S.
<b>S2_SW_INGV_02</b>	GIS	ArcMap software for data visualization/analysis (can be also QGIS)	Commercial	
<b>S2_SW_INGV_03</b>	Statistical tools	Software to perform some statistical computations on the obtained results	Commercial/INGV	Claudia S.
<b>S2_SW_INGV_04</b>	Tomography	Software for tomography model		
<b>S2_SW_INGV_05</b>	SAR LOS delay	Software to perform the tropospheric SAR LOS delay		Claudia

### Algorithms for scenario 2a-d

N°	Name	Description	S2_SW_ Used/Code type	Interaction
<b>Algo_INGV_01</b>	IPWV	Suite of algorithms written by INGV for the IPWV estimate	Matlab	medium
<b>Algo_INGV_02</b>	Tomography	Suite of algorithms written by INGV for water vapour tomography	Matlab	High
<b>Algo_INGV_03</b>	SAR LOS	One or more algorithms	Matlab	low



	delay	written by INGV forLoS calculations.		
--	-------	--------------------------------------	--	--

## User stories-actions GUI requirements for scenario 2a-d

Specific GUI Requirement	Free text description/Image
<b>S2_GUI_INGV_01</b>	The user will see and use the Matlab interface launched within a graphic linux/windows VM. A final popup window appears in the VRE once the processing is over. I shall be able to interact (display/download etc.) with the results obtained with the Matlab codes, in the map browser or in Matlab or in a GIS (Arcmap/QGis).
<b>S2_GUI_INGV_02</b>	The user wants to discover the temporary and permanent ROs in the VRE map browser. The GUI should allow the user to query (spatial, temporal, author, content, etc.) the RO database of the VRC, display the ROs as symbols on the map, select them and display their content, edit the RO, save it as a permanent RO, or delete it. If the RO is saved, a list of people (the working group) should be notified via email. Something that graphically notifies the creation of a new RO (temporary or permanent) should be visible in the map browser and in the information windows of the main GUI.
<b>S2_GUI_INGV_03</b>	The User will use the map browser or a GIS interface (in the VM) to validate the results, e.g., load a previously computed dataset, open other kinds of data, perform some statistical analysis. Once the results are validated the user should be able to modify the RO using the GUI above.

## Control points for scenario 2a-d

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
<b>S2_HI_INGV_01</b>	User needs to check visually the log file of the IPWV run and assess the quality of the result.	YES: the amount and quality of the input data (e.g. number of GPS stations with data, amount if cloud cover in MODIS imagery, etc.) is sufficient NO: the amount and quality of the input data is not sufficient	YES: Press APPROVE button, the temporary RO is created and the group members are notified of the end of the IPWV step NO: All output files are deleted and the procedure can be restarted from scratch
<b>S2_HI_INGV_02</b>	User needs to validate the results of the IPWV processing.	YES: the statistical analysis of the IPWV results is providing evidence of a good result NO: the statistical analysis of the IPWV results is providing evidence of a poor result	YES: Press VALIDATED button, the permanent RO is created and the group members are notified of the existence of a permanent IPWV RO on Mt.Etna NO: All output files and the temporary RO are deleted and the procedure can be restarted from scratch



<b>S2_HI_INGV_03</b>	User needs to validate the results of the tomographic processing.	<p>YES: the statistical analysis of the tomographic results is providing evidence of a good result</p> <p>NO: the statistical analysis of the tomographic results is providing evidence of a poor result</p>	<p>YES: Press VALIDATED button, the permanent RO is created and the group members are notified of the existence of a permanent tomographic RO on Mt.Etna</p> <p>NO: All output files and the temporary RO are deleted and the procedure can be restarted from scratch</p>
----------------------	---	--	---

### Scenario 3: The ash cloud investigation on Mt.Etna

#### Scenario 3a: The SEVIRI ash cloud time series

Action ID	Action description	Dependencies	Challenge
<b>S3_US_INGV_01</b>	Luca logs in the VRE - Supersite portal, and selects the Mt.Etna Supersite as his area of interest.		
<b>S3_US_INGV_02</b>	<p>Luca discovers the input data using the discovery service in the VRE portal. The data to access are: a time series of SEVIRI data passes over Mt. Etna, the Atmospheric Profiles from WMO stations and/or from NCEP and ECMWF.</p> <p>He selects the SEVIRI dataset of interest through spatial and temporal queries. Through these queries he selects both images already acquired (and existing in the archive) and images to be acquired. Then the VRE infrastructure will automatically access the SEVIRI images as soon as they are made available in the catalogue.</p>	S3_Infra_INGV_01	
<b>S3_US_INGV_03</b>	<p>Luca creates a directory in the VRE, where all the input data for the VPR code will be placed.</p> <p>The dataset S3_DS_INGV_01_1 is created with the first image. A series of S3_DS_INGV_01_n files are then created for each of the n successive images of the time series, as they become available.</p>	<p>S3_DS_INGV_01_1</p> <p>S3_DS_INGV_01_n</p> <p>S3_GUI_INGV_01</p>	
<b>S3_US_INGV_04</b>	Stefano (who is registered as member of the same working group as Luca) is automatically notified that Luca has created a new dataset for the VPR processing for the Mt Etna volcano.	S3_GUI_INGV_01	
<b>S3_US_INGV_05</b>	Luca places in the directory reserved for the VPR input data, the land/sea and the view zenith angle images, all obtained from the first SEVIRI datasets S3_DS_INGV_01_01. These files will be used for the subsequent processings of all images in the time series.	<p>S3_DS_INGV_02</p> <p>S3_SW_INGV_02</p> <p>S3_GUI_INGV_01</p>	



<b>S3_US_INGV_06</b>	<p>Luca opens a VM to run a Matlab code to calculate the volcanic cloud mask and the average height of the volcanic cloud based on the image data.</p> <p>Luca could receive both datasets from a different source (external code, EO or in situ measurements). The volcanic cloud height is written in a parameter file for subsequent VPR run.</p>	<p>S3_DS_INGV_03 S3_SW_INGV_03</p>	
<b>S3_US_INGV_07</b>	<p>Luca runs the VPR code. The code requires as input the SEVIRI image, the volcanic cloud mask and the volcanic cloud height. These dynamical inputs can be provided either manually or automatically. The other input datasets are static, i.e. the land/sea mask, etc., do not vary for each successive SEVIRI image in the time series. There are no intermediate steps.</p> <p>When the run ends, the outputs (ash mass, effective radius and aerosol optical depth) are stored in the same VRE storage area where the input data are located, using a specific directory tree.</p> <p>A temporary RO is automatically created including the input data, the workflows used to calculate the volcanic cloud mask, the height and the VPR, and the results. This is a dynamical RO, which is updated for each new SEVIRI image in the time series.</p>	<p>S3_SW_INGV_01 S3_DS_INGV_01_1 S3_DS_INGV_02 S3_DS_INGV_03 S3_DS_INGV_04_1 S3_RO_INGV_01</p>	
<b>S3_US_INGV_08</b>	<p>The RO is automatically re-executed with the n SEVIRI images of the time series, and each time it is updated with the n output results of the VPR.</p> <p>At the end of the n runs, the RO is still temporary.</p>	<p>S3_SW_INGV_01 S3_DS_INGV_04_n S3_RO_INGV_01</p>	
<b>S3_US_INGV_09</b>	<p>Stefano (from the same working group) receives notification that a temporary VPR time series RO has been created for the Mt Etna volcanic cloud. The notification has a link by which he can open directly the results in the map browser, or in ArcMap/QGIS or ENVI/Matlab in the VM.</p> <p>He uploads and displays in the chosen window interface other data or results he needs for the validation (e.g. the S3_RO_INGV_04).</p> <p>He carries out the validation of the final results by comparing the results against different datasets or previous results, or by performing statistical analysis.</p>	<p>S3_DS_INGV_04_1-n S3_RO_INGV_01 S3_SW_INGV_04 S3_SW_INGV_05 S3_GUI_INGV_02 S3_GUI_INGV_03</p>	
<b>S3_US_INGV_10</b>	<p>If the validation process is satisfactory Stefano approves the final results and converts the RO status into permanent (it becomes S3_RO_INGV_02). He may want to edit the RO and add other information into it.</p> <p>S3_RO_INGV_01 is deleted.</p>	<p>S3_RO_INGV_01 S3_RO_INGV_02  S3_HI_INGV_01</p>	



	The scenario is ended.		
<b>S3_US_INGV_11</b>	<p>If the validation process is not satisfactory (e.g. poor accuracy, high residuals, too much noise, etc.) Stefano rejects the final results.</p> <p>He notifies Luca and the scenario should restart from the beginning.</p> <p>S3_RO_INGV_01 can be deleted or kept in the database for reference.</p>	<p>S3_RO_INGV_01</p> <p>S3_HI_INGV_01</p>	

### Scenario 3b: The MODIS ash cloud map

Action ID	Action description	Dependencies	Challenge
<b>S3_US_INGV_12</b>	Luca logs in the VRE - Supersite portal, and selects the Mt.Etna Supersite as his area of interest.		
<b>S3_US_INGV_13</b>	<p>Luca discovers the input data using the discovery service in the VRE portal. The data to access are the MODIS data passes over Mt. Etna, the Atmospheric Profiles from WMO stations, and/or from NCEP and ECMWF.</p> <p>He selects the data of interest through spatial and temporal queries.</p>	S3_Infra_INGV_01	
<b>S3_US_INGV_14</b>	Luca wants to place the dataset in a special area provided by the VRE, from where the VPR SW will read the data. The dataset S3_DS_INGV_05 is created.	<p>S3_DS_INGV_05</p> <p>S3_GUI_INGV_01</p>	
<b>S3_US_INGV_15</b>	Stefano (who is registered as member of the same working group of Luca) is automatically notified that Luca has created a new dataset for the VPR processing for the Mt Etna volcano.	S3_GUI_INGV_01	
<b>S3_US_INGV_16</b>	Luca places in the directory reserved for the VPR input data, the land/sea and the view zenith angle images, all obtained from the MODIS image.	<p>S3_DS_INGV_06</p> <p>S3_SW_INGV_02</p> <p>S3_GUI_INGV_01</p>	
<b>S3_US_INGV_17</b>	<p>Luca opens a VM to run a Matlab code to calculate the volcanic cloud mask and the average height of the volcanic cloud, based on the MODIS image data.</p> <p>Luca could receive both datasets from a different source (external code, EO or in situ measurements). The volcanic cloud height is written in a parameter file for subsequent VPR run.</p>	<p>S3_DS_INGV_03</p> <p>S3_SW_INGV_03</p>	
<b>S3_US_INGV_18</b>	<p>Luca opens a VM to run the VPR code. The code does not require other input than the data described above. There are no intermediate steps.</p> <p>When the run ends, the outputs (ash mass, effective radius and aerosol optical depth) are stored in the</p>	<p>S3_SW_INGV_01</p> <p>S3_DS_INGV_05</p> <p>S3_DS_INGV_06</p> <p>S3_DS_INGV_03</p>	



	<p>same VRE storage area where the input data are located, using a specific directory tree.</p> <p>A temporary RO is automatically created including the input data, the workflows used to calculate the volcanic cloud mask, the height and the VPR, and the results.</p>	<p>S3_DS_INGV_07</p> <p>S3_RO_INGV_03</p>	
<b>S3_US_INGV_19</b>	<p>Stefano (from the same working group) receives notification that a temporary RO has been created for the Mt Etna volcanic cloud. The notification has a link by which he can open directly the results in the map browser or in the ArcMap or ENVI windows in the VM.</p> <p>He uploads and displays in the chosen window interface other data or results he needs for the validation (e.g. the S3_RO_INGV_04).</p> <p>He carries out the validation of the final results by comparing the results against different datasets or previous results, or by performing statistical analysis.</p>	<p>S3_DS_INGV_04</p> <p>S3_RO_INGV_03</p> <p>S3_SW_INGV_04</p> <p>S3_SW_INGV_05</p> <p>S3_GUI_INGV_02</p> <p>S3_GUI_INGV_03</p>	
<b>S3_US_INGV_20</b>	<p>If the validation process is satisfactory Stefano approves the final results and converts the RO status into permanent (it becomes S3_RO_INGV_02). He may want to edit the RO and add other information into it.</p> <p>S3_RO_INGV_01 is deleted.</p> <p>The scenario is ended.</p>	<p>S3_RO_INGV_04</p> <p>S3_RO_INGV_03</p> <p>S3_HI_INGV_01</p>	
<b>S3_US_INGV_21</b>	<p>If the validation process is not satisfactory (e.g. poor accuracy, high residuals, too much noise, etc.) Stefano rejects the final results.</p> <p>He notifies Luca and the scenario should restart from the beginning.</p> <p>S3_RO_INGV_03 can be deleted or kept in the database for reference.</p>	<p>S3_RO_INGV_03</p>	

## Dependencies for Scenario 3

### Datasets/ Products for Scenario 3

N°	Name	Source	Volume	Format
<b>S3_DS_INGV_01_1</b> to <b>S3_DS_INGV_01_n</b>	Image data for VPR time series	n SEVIRI images, Atmospheric Profiles from WMO stations and/or from NCEP and ECMWF.	100 Mb * n (n could be 50-150)	Level 1 for the images
<b>S3_DS_INGV_02</b>	Ancillary data for VPR time	Volcanic cloud mask, land/sea and view zenith	100 Mb	raster





	series	angle images. They are obtained from the initial SEVIRI image of the series S3_DS_INGV_01_1.		
<b>S3_DS_INGV_03</b>	Plume height	<p>Average height of the volcanic cloud. It can be calculated from EO or in situ observations.</p> <p>If the height is calculated from EO data the action can be carried out using an automatic extraction procedure or a manual procedure using Matlab (or ENVI/IDL). These procedures could be run in the VM or on the desktop PC of the user and then the result is uploaded from there into the VRE.</p>	1kb	ascii
<b>S3_DS_INGV_04_1 to S3_DS_INGV_04_n</b>	VPR time series results	<p>n output datasets resulting from the VPR time series processing. Each dataset is composed by the maps of quantities that characterise the volcanic cloud, i.e. the aerosol optical depth (AOD), the effective radius (Re), and the columnar abundance (mass per surface unit) for the ash cloud.</p> <p>Additional information can be given as aggregated results (like total ash mass, or mean AOD and Re).</p>	10Mb * n (n could be 50-150)	binary
<b>S3_DS_INGV_05</b>	Image data for VPR	MODIS image, Atmospheric Profiles from WMO stations, and/or from NCEP and ECMWF.	200 Mb for MODIS	Level 1 for the image
<b>S3_DS_INGV_06</b>	Ancillary data for VPR time series	Volcanic cloud mask, land/sea and view zenith angle images obtained from the MODIS image.	100 Mb	raster



<b>S3_DS_INGV_07</b>	VPR results	<p>The output dataset resulting from the VPR time series processing. The dataset is composed by the maps of quantities that characterise the volcanic cloud, i.e. the aerosol optical depth (AOD), the effective radius (Re), and the columnar abundance (mass per surface unit) for the ash cloud.</p> <p>Additional information can be given as aggregated results (like total ash mass, or mean AOD and Re).</p>	10Mb	binary

### Research Objects for Scenario 3

N°	Name	Source	Description	Format
<b>S3_RO_INGV_01</b>	Time series VPR analysis	VPR code + VRE	Contains input data, and intermediate and output results of the VPR time series processing. Should contain the VPR log file. Contains the workflow to be repeated for the n images of the time series .	Temporary
<b>S3_RO_INGV_02</b>	Time series VPR analysis	VPR code + VRE + user edits	Contains all information listed above plus it contains also results of the validation procedure and some textual description explaining the knowledge used for the entire scientific process.	Permanent
<b>S3_RO_INGV_03</b>	VPR analysis	VPR code + VRE	Contains input data, and intermediate and output results of the VPR processing. Should contain the VPR log file. Contains the workflow to be repeated for the n images of the time series .	Temporary
<b>S3_RO_INGV_04</b>	VPR analysis	VPR code + VRE + user edits	Contains all information listed above plus it contains also results of the validation procedure and some textual description explaining the knowledge used for the entire scientific process.	Permanent



### Infrastructures/catalogues for Scenario 3

Infrastructure/ Catalogue Name	Infrastructure/ Catalogue Description	Access policy	Com. Protocol	Contacts
<b>S3_Infra_INGV_1</b>	SEVIRI/MODIS Catalogues	Open access		Luca Merucci

### Software requirements for Scenario 3

N°	Name	Description	Open source/ IPR	Contacts
<b>S3_SW_INGV_01</b>	Volcanic Plume Removal - VPR	Code to calculate the ash mass, effective radius and aerosol optical depth from MODIS and SEVIRI data	INGV	Luca M. /Stefano C.
<b>S3_SW_INGV_02</b>	Ancillary data extraction	Code or codes to extract from satellite images the volcanic cloud mask, the land/sea mask and the view zenith angle image. The code could use an automatic extraction procedure or a manual procedure using Matlab (or ENVI/IDL). These procedures could be run in the VM or alternatively they could be run on the desktop PC of the user and then the result is uploaded from there into the VRE.	INGV	Luca M. /Stefano C
<b>S3_SW_INGV_03</b>	Plume height extraction	Code or codes to extract from satellite images the average height of the volcanic cloud. The code could use an automatic extraction procedure or a manual procedure using Matlab (or ENVI/IDL). These procedures could be run in the VM or alternatively they could be run on the desktop PC of the user and	INGV	Luca M. /Stefano C



		then the result is uploaded from there into the VRE.		
<b>S3_SW_INGV_04</b>	Statistical tools	Software to perform some statistical computations on the obtained results (e.g. statistics of residuals, boolean operations, spatial profiling, spatial gradient analysis, time series spatial averaging, time series linear-non linear fitting, etc.).	Commercial/INGV	Luca M. /Stefano C
<b>S3_SW_INGV_05</b>	GIS	ArcMap software for data visualization/analysis (can be also QGIS)	Commercial	Luca M. /Stefano C

### Algorithms for Scenario 3

N°	Name	Description	Sw Used/Code type	Interaction
<b>S3_Algo_INGV_01</b>	VPR algorithm	Code to calculate the ash mass, effective radius and aerosol optical depth from MODIS and SEVIRI data	Matlab	medium
<b>S3_Algo_INGV_02</b>	Ancillary data extraction algorithms	Code or codes to extract from satellite images the volcanic cloud mask, the land/sea mask and the view zenith angle image. The code could use an automatic extraction procedure or a manual procedure using Matlab (or ENVI/IDL).	Matlab , IDL/ENVI	High
<b>S3_Algo_INGV_03</b>	Plume height extraction algorithms	Code or codes to extract from satellite images the average height of the volcanic cloud. The code could use an automatic extraction procedure or a manual procedure using Matlab (or ENVI/IDL).	Matlab , IDL/ENVI	High

### User stories-actions GUI requirements for Scenario 3

Specific Requirement	GUI	Free text description/Image
----------------------	-----	-----------------------------



<b>S3_GUI_INGV_01</b>	<p>A graphic window should be available to ask the user where the input data should be placed (for the processing SW to use them) and at the end of the processing the same window should display the location and directory tree pointing to the final results.</p> <p>The user must see and use the processing interface launched within a graphic linux/windows VM.</p> <p>A notification window appears in the VRE once the processing is over. The user shall be able to interact (display/download/validate, etc.) with the results obtained by the processing in the map browser or in ENVI/Matlab or in Arcmap/QGis.</p>
<b>S3_GUI_INGV_02</b>	<p>The user wants to discover the temporary and permanent ROs in the VRE map browser. The GUI should allow the user to query (spatial, temporal, author, content, etc.) the RO database of the VRC, display the ROs on the map, select them and display their content, edit the RO, save it as a permanent RO, or delete it. If the RO is saved the working group should be notified via email. Something that graphically notifies the creation of a new RO (temporary or permanent) should be visible in the map browser and in the information windows of the main GUI.</p>
<b>S3_GUI_INGV_03</b>	<p>The User will use the map browser or a GIS interface (in the VM) to validate the results, e.g., load a previously computed dataset, open other kind of data, perform some statistical analysis. Once the results are validated the temporary RO should be saved or discarded (see above).</p>

### Control points for Scenario 3

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
<b>S3_HI_INGV_01</b>	User validates the VPR results	<p>YES: User validates the VPR results.</p> <p>NO: User does not validate the VPR results.</p>	<p>YES: Press OK button and converts the temporary RO to permanent and then publishes it.</p> <p>NO: start the VPR analysis once again</p>

### Scenario 4: The generic scientific data analysis investigation

#### User actions for Scenario 4

Action ID	Action description	Dependencies	Challenge
<b>S4_US_INGV_01</b>	Simone logs in the VRE - Supersite portal, and selects the Mt.Etna Supersite as his area of interest.		
	Simone opens the VM and uploads from his desktop PC to a specific area in the VM an IDL code for data modeling/analysis, and the test data needed to run the code.	<p>S4_GUI_INGV_01</p> <p>S4_DS_INGV_01</p> <p>S4_SW_INGV_01</p>	



<b>S4_US_INGV_06</b>	Simone edits the IDL code, and then runs the code on the test data. The code opens various windows to input processing parameters.	S4_DS_INGV_01 S4_SW_INGV_01	
<b>S4_US_INGV_07</b>	When the run ends, the output results are stored in the same VRE storage area where the test data are located, using a specific directory tree.  Simone displays the output results in the VRE map browser, or in ArcMap/QGis or ENVI in the VM, and compares them with previous results or runs statistical analyses on them.	S4_DS_INGV_02 S4_GUI_INGV_03	
	Simone decides that his code is providing good results and is ready to be shared with the community.  He creates some graphical forms to pass input data to his code, and an executable version of the code.  He then publishes it in the open access SW section of the VRE, from which other scientists can run it.	S4_GUI_INGV_02 S4_SW_INGV_02	
<b>S4_US_INGV_08</b>	Elisa enters the Mt.Etna VRE and runs the code created by Simone from the public access SW tool window of the VRE.  She does not have to access the VM interface to run the code.  She manually inputs data and processing parameters using the graphic forms created by Simone.	S4_SW_INGV_02 S4_GUI_INGV_04 S4_DS_INGV_03 S4_Infra_INGV_1	
<b>S4_US_INGV_08</b>	Elisa run ends and a temporary RO is created.	S4_DS_INGV_04 S4_RO_INGV_01	
<b>S4_US_INGV_09</b>	Elisa opens the RO and the processing results in the VRE map browser, or in ArcMap/QGis or in ENVI/Matlab in the VM.  She uploads and displays in the chosen window interface other data or previous models she needs for the validation (e.g. from a catalogue of models).  She carries out the validation of the final results by comparing the results against different datasets or previous models, or by performing statistical analysis.	S4_DS_INGV_04 S4_DS_INGV_05 S4_RO_INGV_01 S4_SW_INGV_04 S4_SW_INGV_05 S4_SW_INGV_06 S4_GUI_INGV_03 S4_GUI_INGV_05	
<b>S4_US_INGV_10</b>	If the validation process is satisfactory Elisa approves the final results and converts the RO status into permanent (it becomes S4_RO_INGV_02). She may want to edit the RO and add other textual information into it.  The scenario is ended.	S4_RO_INGV_01 S4_RO_INGV_02  S4_HI_INGV_01	
<b>S4_US_INGV_11</b>	If the validation process is not satisfactory (e.g. poor model accuracy, high residuals, too much noise,	S4_RO_INGV_01	



	etc.) Elisa rejects the final results. The scenario should restart from the beginning. S4_RO_INGV_01 can be deleted or kept in the database for reference.	S4_HI_INGV_01	
--	--	---------------	--

## Dependencies for Scenario 4

### Datasets/ Products for Scenario 4

N°	Name	Source	Volume	Format
S4_DS_INGV_01	Test data for modeling code	Synthetic dataset or test dat	100 Mb	varies
S4_DS_INGV_02	Output results of modeling code test run	Modeling/analysis code	100 Mb	varies
S4_DS_INGV_03	Real data for modeling code	Real dataset to be modeled	100 Mb	varies
S4_DS_INGV_04	Output results of modeling code	Modeling/analysis code	100 Mb	varies
S4_DS_INGV_05	Data and/or models for validation	A catalogue of previous models, or data catalogues	100 Mb	varies

### Research Objects for Scenario 4

N°	Name	Source	Description	Format
S4_RO_INGV_01	Modeling/ analysis results	Modeling code	Contains input data, processing parameters and output results of the modeling processing. Should contain the modeling log file. Contains the workflow.	Temporary
S4_RO_INGV_02	Modeling/ analysis results	Modeling code	Contains all information listed above plus it contains also results of the validation procedure and some textual description explaining the knowledge used for the entire scientific process.	Permanent

### Infrastructures/catalogues for Scenario 4



Infrastructure/ Catalogue Name	Infrastructure/ Catalogue Description	Access policy	Com. Protocol	Contacts
S4_Infra_INGV_1	Catalogue of geophysical models of the Mt.Etna Supersite	Open to the community		Simone

## Software requirements for Scenario 4

N°	Name	Description	Open source/ IPR	Contacts
S4_SW_INGV_01	Modeling/ analysis software	INGV code for data analysis or modeling, working version, running only in an INDL/Matlab environment	INGV	Simone
S4_SW_INGV_02	Modeling/ analysis software	INGV code for data analysis or modeling, running from the SW tool window of the VRE.	INGV	Simone
S4_SW_INGV_04	Statistical tools	Software to perform some statistical computations on the obtained results (e.g. statistics of residuals, boolean operations, spatial profiling, spatial gradient analysis, time series spatial averaging, time series linear-non linear fitting, etc.).	Commercial /INGV	Simone
S4_SW_INGV_05	GIS	ArcMap software for data visualization/analysis (can be also QGis)	Commercial	Simone
S4_SW_INGV_06	Programmin g language	ENVI/IDL/Matlab	Commercial	Simone

## Algorithms for Scenario 4

N°	Name	Description	Sw Used/Code type	Interaction
S4_Algo_INGV_01	Modeling algorithm	Code to invert or analyse data	Matlab/IDL	High

## User stories-actions GUI requirements for Scenario 4

Specific GUI	Free text description/Image
--------------	-----------------------------





Requirement	
<b>S4_GUI_INGV_01</b>	<p>A graphic window should be available to ask the user where the modeling/analysis code and the input data should be placed in the VM file system.</p> <p>At the end of the processing the same window should display the location and directory tree pointing to the final results.</p> <p>The user must see and use the processing interface launched within a graphic linux/windows VM.</p> <p>A notification window appears in the VRE once the processing is over. The user shall be able to interact (display/download/validate, etc.) with the results obtained by the processing in the map browser or in ENVI/Matlab or in Arcmap/QGis.</p>
<b>S4_GUI_INGV_02</b>	A GUI is needed to create graphic windows to pass manually input parameters to IDL/Matlab codes to be run by the community.
<b>S4_GUI_INGV_03</b>	The User will use the map browser or a GIS interface (in the VM) to validate the results, e.g., load a previously computed dataset, open other kind of data, perform some statistical analysis. Once the results are validated the temporary RO should be saved or discarded (see above).
<b>S4_GUI_INGV_04</b>	A GUI placed in the main VRE window in which open access SW can be accessed for download, edit and run. In case of run, the GUI will manage the input and output parameters of the code.
<b>S4_GUI_INGV_05</b>	The user wants to discover the temporary and permanent ROs in the VRE map browser. The GUI should allow the user to query (spatial, temporal, author, content, etc.) the RO database of the VRC, display the ROs on the map, select them and display their content, edit the RO, save it as a permanent RO, or delete it. If the RO is saved the working group should be notified via email. Something that graphically notifies the creation of a new RO (temporary or permanent) should be visible in the map browser and in the information windows of the main GUI.

### Control points for Scenario 4

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
<b>S4_HI_INGV_01</b>	User validates the VPR results	<p>YES: User validates the VPR results.</p> <p>NO: User does not validate the VPR results.</p>	<p>YES: Press OK button and converts the temporary RO to permanent and then publishes it.</p> <p>NO: start the VPR analysis once again</p>

### Scenario 5: The Campi Flegrei Volcano Source Modelling

#### User actions for Scenario 5

Action ID	Action description	Dependencies	Challenge
<b>S5_US_INGV_01</b>	Cristiano logs in the VRE - Supersite portal, and selects the Campi Flegrei Supersite as his area of interest.		



<b>S5_US_INGV_02</b>	Elisa logs in the VRE - Supersite portal, and selects the Campi Flegrei Supersite as his area of interest.		
<b>S5_US_INGV_03</b>	<i>Cristiano wants to place the dataset in a special area provided by the VRE, from where Elisa and him are able to collaborate as a group.</i> Cristiano uploads the deformation data of 2011-2013 at Campi Flegrei from COSMO-SkyMed InSAR analysis. Two datasets (ascending/descending orbits) S5_DS_INGV_01 and S5_DS_INGV_02 are created.		
<b>S5_US_INGV_04</b>	Elisa is automatically notified that Cristiano has created new datasets of the Campi Flegrei volcano.	S5_DS_INGV_01 S5_DS_INGV_02	
<b>S5_US_INGV_05</b>	Elisa opens a VM to perform volcano source modelling. She needs a VM with limited access to his research group, including Cristiano.	S5_GUI_INGV_01	
<b>S5_US_INGV_06</b>	Elisa launches ArcGis and visualizes the two datasets. She decides the data show remarkable deformation (uplift) and good spatial coverage. She wants to proceed to model these data to infer the volcanic source characteristics.	S5_SW_INGV_01 <b>S5_HI_INGV_01</b>	
<b>S5_US_INGV_07</b>	Elisa verifies, as usually happens ( <i>therefore is not a control point</i> ), that the number of InSAR data points is too large. Therefore, she downsamples both datasets using the shape file downsampling tool in GIS, creating S5_DS_INGV_03 and S5_DS_INGV_04.	S5_SW_INGV_01 S5_DS_INGV_01 S5_DS_INGV_02 S5_GUI_INGV_01	
<b>S5_US_INGV_08</b>	She launches the Volcano Source Modelling SW and starts the analysis.  The Volcano Source Modelling (VSM) tool begins requiring input data selection and parameter definition (number of datasets, type of dataset, weight among the different datasets).	S5_SW_INGV_02 S5_DS_INGV_03 S5_DS_INGV_04 S5_GUI_INGV_02	
<b>S5_US_INGV_09</b>	The VSM tool requires to choose the type of volcanic source and the range of its parameters' search. The spherical source is chosen, and the extreme values of its parameters are set.	S5_SW_INGV_02 S5_GUI_INGV_02	
<b>S5_US_INGV_08</b>	There are advanced parameters of the inversion algorithm to be set, but Elisa decides to use the default ones. Elisa clicks the inversion button and execution starts.	S5_SW_INGV_02 S5_DS_INGV_03 S5_DS_INGV_04 S5_GUI_INGV_02	
<b>S5_US_INGV_10</b>	While running the inversion tool within the VSM, the ongoing parameters' search is visualized by graphical windows. The results are stored in S5_DS_INGV_05 and S5_DS_INGV_06 (one for each input dataset) and the parameters' sampling (ascii file) S5_DS_INGV_07.	S5_DS_INGV_03 S5_DS_INGV_04 S5_SW_INGV_02 S5_GUI_INGV_02	
<b>S5_US_INGV_11</b>	A RO is automatically created including reference to	S5_SW_INGV_02	



	all the input and output files. The VSM log file is embedded in the RO. Possibly also the workflow. Within the RO the most important files (the final results) are the shape files of the mean ground velocities and a vector file of the ground displacement at each coherent pixel.	S5_DS_INGV_05 S5_DS_INGV_06 S5_DS_INGV_07 S5_GUI_INGV_02 S5_RO_INGV_01	
<b>S5_US_INGV_12</b>	Elisa checks the results of the inversion: the parameters' search histograms, the best-fit, the data residuals. She downloads locally S5_DS_INGV_05, S5_DS_INGV_06 and S5_DS_INGV_07 to check the results with her previous attempts.	S5_SW_INGV_01 S5_SW_INGV_02 S5_DS_INGV_05 S5_DS_INGV_06 S5_DS_INGV_07 S5_GUI_INGV_01 S5_GUI_INGV_02 S5_RO_INGV_01	
<b>S5_US_INGV_13</b>	If the validation process is satisfactory, Elisa approves the final results of S5_RO_INGV_01, she converts its status into permanent. She may want to edit the RO and add other information into it. S5_RO_INGV_02 is created and S5_RO_INGV_01 is deleted. The scenario is ended.	S5_RO_INGV_01 S5_RO_INGV_02  S5_HI_INGV_02	
<b>S5_US_INGV_14</b>	If the validation process is not satisfactory (e.g. poor accuracy, high residuals, too much noise, etc.) Elisa rejects the final results of S5_RO_INGV_01. She notifies Cristiano and the scenario should restart from S5_US_INGV_08. S5_RO_INGV_01 can be deleted or kept in the database for reference (with additional information).	S5_RO_INGV_01  S5_HI_INGV_03	

## Dependencies for Scenario 5

### Datasets/ Products for Scenario 5

N°	Name	Source	Volume	Format
<b>S5_DS_INGV_01</b>	InSAR image	Result of SAR interferometry	50-500 Mb	Raster and vector
<b>S5_DS_INGV_02</b>	InSAR image	Result of SAR interferometry	50-500 Mb	Raster and vector
<b>S5_DS_INGV_03</b>	Downsampled InSAR data	GIS tool or other downsampling tools	< 100 Mb	Shape - ascii
<b>S5_DS_INGV_04</b>	Downsampled InSAR data	GIS tool or other downsampling tools	< 100 Mb	Shape - ascii



<b>S5_DS_INGV_05</b>	Computed data	Result of inversion tool VSM	< 100 Mb	Shape - ascii
<b>S5_DS_INGV_06</b>	Computed data	Result of inversion tool VSM	< 100 Mb	Shape - ascii
<b>S5_DS_INGV_07</b>	Parameters' search	Result of inversion tool VSM	< 100 Mb	Shape - ascii

## Research Objects for Scenario 5

N°	Name	Source	Description	Format
<b>S5_RO_INGV_01</b>	Non-validated VSM of Campi Flegrei	VSM tool	Contains input data, and intermediate and output results of the VSM processing. Should contain the log file. Contains the workflow if possible. Contains some textual description explaining the knowledge used for the VSM processing.	Temporary
<b>S5_RO_INGV_02</b>	Validated VSM of Campi Flegrei	VSM + validation	Contains all information listed above plus the results of the validation procedure and some textual description explaining the knowledge used for the entire scientific process.	Permanent

## Infrastructures/catalogues for Scenario 5

Infrastructure/ Catalogue Name	Infrastructure/ Catalogue Description	Access policy	Com. Protocol	Contacts

## Software requirements for Scenario 5

N°	Name	Description	Open source/ IPR	Contacts
<b>S5_SW_INGV_01</b>	GIS	ArcMap and/or QGis software for data visualization/analysis	Commercial	
<b>S5_SW_INGV_02</b>	VSM	Software to perform inversions of geodetic data (e.g., InSAR, GPS, etc.) to infer the volcanic source parameters (depth, shape, potency, etc.).	Open/INGV	Elisa Trasatti

## Algorithms for Scenario 5



N°	Name	Description	Sw Used/Code type	Interaction
<b>S5_Algo_INGV_01</b>	Mogi	analytical solution of a sphere dilating in a half-space (Mogi, 1958).	Fortran	low
<b>S5_Algo_INGV_02</b>	NA	Too to perform data inversions of any kind. It resolce the non-linear inversion problem (Sambridge, 1999).	Fortran	low

### User stories-actions GUI requirements for Scenario 5

Specific GUI Requirement	Free text description/Image
<b>S5_GUI_INGV_01</b>	The user must see and use the GIS interface launched within a graphic linux/windows VM. A graphic window should be available to inform the user where the input data have been placed.
<b>S5_GUI_INGV_02</b>	The user is able to discover data, set up the input parameters for the VSM tool, open the input/computed.

### Control points for Scenario 5

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
<b>S5_HI_INGV_01</b>	User checks the suitability of the dataset DS_INGV_01 and DS_INGV_02 to perform data inversion Uses GUI_INGV_01.	YES: User verifies good coverage and remarcable deformation, absence of InSAR processing artifacts. NO: The data is not suitable for inversions (e.g., reduced spatial coverage, low displacements, presence of ramps).	YES: proceed to data downsampling. NO: end of the work! Or choose to use a single dataset if only one of the two is not suitable for inversions.
<b>S5_HI_INGV_02</b>	User checks the temporary RO content. Uses GUI_INGV_01 and GUI_INGV_02.	YES: User verifies completeness of temporary RO. NO: The temporary RO was not correctly created and is incomplete.	YES: Press OK button and publishes the temporary RO NO: manually edits the temporary RO and fills the gaps. Then GOTO YES



<b>S1_HI_INGV_03</b>	User validates the final results. Uses GUI_INGV_01 and GUI_INGV_02.	YES: The final inversion results are validated NO: The inversion results are not validated	YES: Press OK button and converts the temporary RO into permanent (an editing of the RO content may be necessary using GUI_INGV_02). NO: The VSM processing is restarted from the beginning or from an intermediate step within action US_INGV_08
----------------------	--	---	--

### Scenario 6: The generic reference search generation

Action ID	Action description	Dependencies	Challenge
<b>S6a_US_INGV_01</b>	Giuliana logs in the VRE		
<b>S6a_US_INGV_02</b>	Giuliana accesses the VRE of the Supersite of interest (e.g. Mt. Etna)		
<b>S6a_US_INGV_03</b>	Giuliana accesses a subarea of the VRE (see below the GUI) where she performs a set of “smart” queries in selected on line repositories, as Google Scholar, Earth-prints, Isi Web of Science. For example she searches for “InSAR resources about Mount Etna”, or other topics like “2012 eruption” or “deformations”.		
<b>S6a_US_INGV_04</b>	Giuliana checks the resulting lists and verifies that they have a sufficient quality for the subject (e.g. relevance for the subject,...). If the queries are OK, a (temporary) RO is created.	Access to Internet is required as basic infrastructure. S6a_Infra_INGV_1 S6a_Infra_INGV_2 S6a_Infra_INGV_3	
<b>S6a_US_INGV_05</b>	The VRC use case members are notified by email (and/or other means) that there is a new Bibliographic RO for Mt Etna from Giuliana. The VRC members may login and find the RO in the specific Supersite, e.g., Mt Etna. [The notification has a link by which VRC use case members can open directly the result.]		
<b>S6a_US_INGV_06</b>	Elisa is one of the VRC members that is being notified. Since Elisa is working on the recent eruptions at Mt Etna, she is interested in the RO created by Giuliana.		



	Elisa access the Mt Etna Supersite VRE and finds the RO.		
<b>S6a_US_INGV_07</b>	Elisa clicks on the RO button and finds in the web browser the bibliographic search already set up by Giuliana.		
<b>S6a_US_INGV_08</b>	Elisa checks the papers listed, finding many works pertinent to the “2012 Eruption – deformation” search and choose the last one published, of which she had lost the publication notification.		
<b>S6a_US_INGV_09</b>	Elisa validates the RO.		
<b>S6a_US_INGV_10</b>	Stefano is another member of VRC community who is interested in publications about (say) Mt. Etna and InSAR data. He is notified by email that the new RO has been created. He activates the RO, retrieved the publications and adds his comments to some publications. This can be seen as scenario S6b. <sup>16</sup>		

## Dependencies for Scenario 6

### Datasets/ Products for Scenario 6

	Name	Source	Volume	Format
<b>S6a_DS_INGV_01</b>	Google Scholar result list	<a href="https://scholar.google.it/scholar?q=insar+mount+etna+&amp;btnG=&amp;hl=it&amp;as_sdt=0%2C5">https://scholar.google.it/scholar?q=insar+mount+etna+&amp;btnG=&amp;hl=it&amp;as_sdt=0%2C5</a> (see on Rohub the RO object “testbo”/Scholar)	<1Mb	Web page or other format
<b>S6a_DS_INGV_02</b>	Earth-prints result list	<a href="http://www.earth-prints.org/simple-search?query=%2B%28%28InSAR%29+AND+%28Mount+Etna%29+OR+%28Mt+Etna%29%29&amp;rpp=10&amp;sort_by=0&amp;order=DESC&amp;etal=6&amp;submit_search=Update">http://www.earth-prints.org/simple-search?query=%2B%28%28InSAR%29+AND+%28Mount+Etna%29+OR+%28Mt+Etna%29%29&amp;rpp=10&amp;sort_by=0&amp;order=DESC&amp;etal=6&amp;submit_search=Update</a> (see Rohub the RO object “testbo”/Simple Search)	<1Mb	Web page or other format
<b>S6a_DS_INGV_03</b>	ISIwebofscience result list	<a href="http://apps.webofknowledge.com/Search.do?product=WOS&amp;SID=S15zm3oCznRACwRHv3g&amp;search_mode=GeneralSearch&amp;prID=ef2abe1e-1d70-4692-bd9e-089f3496ca07">http://apps.webofknowledge.com/Search.do?product=WOS&amp;SID=S15zm3oCznRACwRHv3g&amp;search_mode=GeneralSearch&amp;prID=ef2abe1e-1d70-4692-bd9e-089f3496ca07</a> (see Rohub the RO object “testbo”/Search.do)	<1Mb	Web page or other format

<sup>16</sup> Please note that we identify at least two situations: 1) the first one is that of GENERATING the Reference search and its related RO, which may include also a process of validation and temporary/permanent biblio RO creation; 2) the second one is that of RE-USING the RO. Furthermore we can add the Reference Analysis, which is the activity of analyzing and commenting results. We have to choose if this is a further scenario or it is an option of the previous one 6a.



## Research Objects for Scenario 6

N°	Name	Source	Description	Format
<b>S6_RO_INGV_01</b>	Non-validated reference search for a given subject (e.g. INSAR and Mt.Etna)	GoogleScholar Earth-prints ISIwebofscience	Contains the queries/URLs to the online repositories. Contains some textual description explaining the knowledge used for the bibliographic search.	Temporary
<b>S6_RO_INGV_02</b>	Validated reference search for a given subject (e.g. INSAR and Mt.Etna)	GoogleScholar Earth-prints ISIwebofscience + validation	Contains all information listed above plus the results of the validation procedure and some textual description explaining the knowledge used for the refinement of the query.	Permanent

Rereference Search Object URI:

<http://sandbox.wf4ever-project.org/portal/ro?ro=http://sandbox.wf4ever-project.org/rodl/ROs/testbo/>

Infrastructure/ Catalogue Name	Infrastructure/ Catalogue Description	Access policy	Com. Protocol	Contacts
<b>S6a_Infra_INGV_1</b>	Earth –Prints repository	Open access		INGV Library
<b>S6a_Infra_INGV_2</b>	Google Scholar	Open access		
<b>S6a_Infra_INGV_3</b>	Isi Web of Science	Requires login/licence (e.g. via IP address)		Single research institution's refereee for licence

## Software requirements for Scenario 6

N°	Name	Description	Open source/ IPR	Contacts
<b>S6_SW_INGV_01</b>	Generic web browser	A generic web browser is sufficient to access online repositories and perform queries of interests.	Free download from Internet	
<b>S6_SW_INGV_02</b>	Comment module	Facility to comment references about a given subject.	Available in the VRE	

## Algorithms for Scenario 6





No algorithms or computations are foreseen at the moment, since this RO has been classified in the taxonomy as a *bibliographic* RO (“e.g. I aggregate literature, my comments, SW, other resources, and make a bundle to disseminate my views on a certain subject. Also available for the public”).)

## User stories-actions GUI requirements for Scenario 6

Specific GUI Requirement	Free text description/Image
<b>S6_GUI_INGV_01</b>	Giuliana accesses an area of the VRE where she has a form with at least two fields to fill in: search terms and selection of the repository and three buttons: search, save, validate. May this area be placed in the tool window of the main VRE GUI?
<b>S6_GUI_INGV_02</b>	Results of the reference search are displayed in the information window of the main VRE GUI.
<b>S6_GUI_INGV_03</b>	Availability of reference search (permanent RO) may be visualized in the tool window of a given Supersite VRE or on a map see also S4_GUI_INGV_05.

See also:

<b>S4_GUI_INGV_05</b>	The user wants to discover the temporary and permanent ROs in the VRE map browser. The GUI should allow the user to query (spatial, temporal, author, content, etc.) the RO database of the VRC, display the ROs on the map, select them and display their content, edit the RO, save it as a permanent RO, or delete it. If the RO is saved the working group should be notified via email. Something that graphically notifies the creation of a new RO (temporary or permanent) should be visible in the map browser and in the information windows of the main GUI.
-----------------------	---

## Control points for Scenario 6

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
<b>S6_HI_INGV_01</b>	User checks the temporary RO content.	YES: User verifies completeness of temporary RO. NO: The temporary RO was not correctly created and is incomplete.	YES: Press OK button and publishes the temporary RO NO: manually edits the temporary RO and fills the gaps. Then GOTO YES



## **Appendix C – Land Monitoring VRC Use Case Scenario**

### Personas

Personas	Representing group	Description
Paulo	Developer	Person developing the service
Michele	Operator	Person operating and creating the requests
Sergio	Collaborator	Person collaborating with the operator
Emanuele	User	Person notified of new items



## Detailed Use Case scenarios

Scenarios	Free text description (max 50 words)
1. Land Monitoring	SatCen wants to detect changes on selected areas of interest using image pairs from a specific dataset.
2. Land Monitoring Time Series	SatCen wants to detect changes on selected areas of interest using time series from an heterogeneous dataset.
3. Land Monitoring supported by Social Media	SatCen wants to detect and characterise changes on selected areas of interest using time series from an heterogeneous dataset and verify them with the support of ancillary information from social media.

### Scenario 1: Land Monitoring

Action ID	Action description	Dependencies	Challenge
US_LM_01	Paulo wants to add a new application with a change detection map algorithm.		<b>EVER-FUN –030</b> <b>EVER-COL-010</b> <b>EVER-RE-100</b> <b>EVER-RE-140</b>
US_LM_02	Paulo wants to update the application with the change detection map algorithm.	US_LM_01	<b>EVER-FUN –030</b> <b>EVER-RE-120</b>
US_LM_03	Paulo wants to define the access rights to the change detection map algorithm <sup>17</sup> .	US_LM_01	<b>TBD</b>
US_LM_04	Michele wants to define a specific area of interest within a specific time and search for products.	DS_LM_01	<b>EVER- GEN-010</b> <b>EVER-RE-010</b>
US_LM_05	Michele wants to discover the exact pairs of a data product from a specific dataset. An exact pair is a product with the same spatial resolution, polarization, orbit (i.e. asc. or desc.) and view angle from the Sentinel 1 archive.	DS_LM_01	<b>EVER-FUN –030</b> <b>EVER-RE-010</b>
US_LM_06	Michele wants to apply a specific processing chain to the selected pair.	US_LM_03 Algo_LM_01 Algo_LM_02 Algo_LM_03	<b>EVER-FUN –030</b> <b>EVER-COM-010</b> <b>EVER-COM-030</b> <b>EVER-RE-020</b> <b>EVER-RE-060</b> <b>EVER-RE-080</b>
US_LM_07	Michele wants to visualize the change detection map and auxiliary information (e.g. logs, metrics).	US_LM_06	<b>EVER-RE-020</b> <b>EVER-RE-030</b>

<sup>17</sup> This step may include the creation of a live RO which is linked to the observation we want to make.



			<b>EVER-RE-090</b>
US_LM_08	Michele wants to publish the change detection map including the metadata of the input products and the service used to produce the map <sup>18</sup> .	US_LM_07	<b>EVER- GEN-010</b> <b>EVER-COM-020</b> <b>EVER-RE-040</b> <b>EVER-RE-050</b> <b>EVER-RE-150</b> <b>EVER-RE-160</b>
US_LM_09	Michele wants to define the access rights to the published change detection map <sup>19</sup> .	US_LM_08	<b>TBD</b>
US_LM_10	Michele wants to define the access rights to the input data products of the change detection map.	DS_LM_01	<b>TBD</b>
US_LM_11	Sergio wants to discover a change detection map.	US_LM_09	<b>TBD</b>
US_LM_12	Michele (or Sergio) wants to upload an auxiliary information (e.g. raster, vector, another change detection) layer to visualize together with change detection map.	US_LM_07	<b>TBD</b>
US_LM_13	Sergio wants to annotate the discovered change detection map.	US_LM_10	<b>EVER- GEN-010</b> <b>EVER-GEN-020</b> <b>EVER-COM-010</b> <b>EVER-COL-010</b> <b>EVER-RE-70</b>
US_LM_14	Sergio wants to reproduce the change detection map.	US_LM_06	<b>EVER-RE-110</b>
US_LM_15	Emanuele wants to subscribe to the change detection service.	US_LM_09	<b>TBD</b>
US_LM_16	Emanuele is notified when a given change detection is produced.	US_LM_13	<b>TBD</b>

## Scenario 2: Land Monitoring Time Series

Action ID	Action description	Dependencies	Challenge
US_LM_01	<i>Paulo wants to add a new application with a change detection map algorithm.</i>		
US_LM_02	<i>Paulo wants to update the application with the change detection map algorithm.</i>		
US_LM_03	<i>Paulo wants to define the access rights to the change detection map algorithm.</i>		
US_LM_17	Michele wants to define a specific area of interest within a specific time and search for products.	DS_LM_01 DS_LM_02	<b>EVER- GEN-010</b> <b>EVER-RE-010</b>

<sup>18</sup> This step may require the creation of a snapshot of the RO.

<sup>19</sup> This action is done when creating the RO in the ROHub



US_LM_18	Michele wants to discover the data products from an heterogeneous dataset. For Sentinel 1, the products should have the same spatial resolution, polarization, orbit (i.e. ascending or descending) and view angle. For Sentinel 2, data should be provided at Level 1c (atmospherically corrected) and have a minimum cloud coverage <sup>20</sup> .	DS_LM_01 DS_LM_02	<b>EVER-FUN –030</b> <b>EVER-RE-010</b>
US_LM_19	Michele wants to apply a specific change detection map algorithm to the selected time series of Sentinel 1 and Sentinel 2 and fuse the output maps.	US_LM_03 Algo_LM_01 Algo_LM_02 Algo_LM_04 Algo_LM_06 Algo_LM_08	<b>EVER-FUN –030</b> <b>EVER-COM-010</b> <b>EVER-COM-030</b> <b>EVER-RE-020</b> <b>EVER-RE-060</b> <b>EVER-RE-080</b>
US_LM_07	<i>Michele wants to visualize the change detection map and auxiliary information (e.g. logs, metrics).</i>		
US_LM_08	<i>Michele wants to publish the change detection map (i.e. create the RO) including the metadata of the input products and the service used to produce the map.</i>		
US_LM_09	<i>Michele wants to define the access rights to the published change detection map.</i>		
US_LM_10	<i>Michele wants to define the access rights to the input data products of the given change detection map.</i>		
US_LM_11	<i>Sergio wants to discover a change detection map.</i>		
US_LM_12	<i>Michele (or Sergio) wants to upload an auxiliary information (e.g. raster, vector, another change detection) layer to visualize together with change detection map.</i>		
US_LM_13	<i>Sergio wants to annotate the discovered change detection map.</i>		
US_LM_14	<i>Sergio wants to reproduce the change detection map.</i>		
US_LTS_15	<i>Emanuele wants to subscribe to the change detection service.</i>		
US_LTS_16	<i>Emanuele is notified when a given change detection is produced.</i>		

### Scenario 3: Land Monitoring Supported by Social Media

Action ID	Action description	Dependencies	Challenge
US_LM_01	<i>Paulo wants to add a new change detection map algorithm.</i>		

<sup>20</sup> All search operations are done by a specialized EO catalogue.



US_LM_02	<i>Paulo wants to update the change detection map algorithm.</i>		
US_LM_03	<i>Paulo wants to define the access rights to the change detection map algorithm.</i>		
US_LM_17	<i>Michele wants to define a specific area of interest within a specific time and search for products.</i>		
US_LM_18	<i>Michele wants to discover the data products from an heterogeneous dataset. For Sentinel 1, the products should have the same spatial resolution, polarization, orbit (i.e. ascending or descending) and view angle. For Sentinel 2, data should be provided at Level 1c (atmospherically corrected) and have a minimum cloud coverage.</i>		
US_LM_20	Michele wants to apply a specific processing chain to the selected time series of Sentinel 1 and Sentinel 2 and fuse the output maps.	US_LM_03 Algo_LM_01 Algo_LM_02 Algo_LM_04 Algo_LM_05 Algo_LM_06 Algo_LM_07 Algo_LM_08	<b>EVER-FUN -030</b> <b>EVER-COM-010</b> <b>EVER-COM-030</b> <b>EVER-RE-020</b> <b>EVER-RE-060</b> <b>EVER-RE-080</b>
US_LM_07	<i>Michele wants to visualize the change detection map and auxiliary information (e.g. logs, metrics).</i>		
US_LM_21	Michele wants to add the information coming from Social Sensing sources <sup>21</sup> related to the specified area of interest.	US_LM_07 DM_LM_03	
US_LM_08	<i>Michele wants to publish the change detection map (i.e. create the RO) including the metadata of the input products and the service used to produce the map.</i>		
US_LM_09	<i>Michele wants to define the access rights to the published change detection map.</i>		
US_LM_10	<i>Michele wants to define the access rights to the input data products of the given change detection map.</i>		
US_LM_11	<i>Sergio wants to discover a change detection map.</i>		
US_LM_12	<i>Michele (or Sergio) wants to upload an auxiliary information (e.g. raster, vector, another change detection) layer to visualize together with change detection map.</i>		
US_LM_13	<i>Sergio wants to annotate the discovered change detection map.</i>		
US_LM_14	<i>Sergio wants to reproduce the change detection map.</i>		
US_LTS_15	<i>Emanuele wants to subscribe to the change detection</i>		

<sup>21</sup> Social sensing sources are shown in GUI\_LM\_04



	<i>service.</i>		
US_LTS_16	<i>Emanuele is notified when a given change detection is produced.</i>		

## Future Services

Possible future scenarios might be related to the activation of the above scenarios in a multiple automatic way, where users receive alerts on detected changes on specific defined testing areas.

## Dependencies

### Datasets/ Products

N°	Name	Source	Volume	Format
DS_LM_01	Sentinel-1A/B	ESA SciHub	TBD	
DS_LM_02	Sentinel-2A	ESA SciHub	TBD	
DS_LM_03	Social Sensing Events <sup>22</sup>	Big Data Europe Platform	TBD	

### Infrastructures/catalogues

Infrastructure/ Catalogue Name	Infrastructure/ Catalogue Description	Access policy	Com. Protocol	Contacts
Infra_LM_01	ESA SciHub	Registered users		
Infra_LM_02	Collaborative Data Hub			
Infra_LM_03	ROHub			
Infra_LM_04	Big Data Europe Platform			

## Software requirements

N°	Name	Description	Open source/ IPR	Contacts
SW_LM_01	SNAP Toolbox	SNAP Toolbox is including Sentinel 1/2/3 toolboxes	Open Source	

<sup>22</sup> They can be originated from RSS feeds or from social media streams (e.g. Twitter), but they will be aggregated according to their related "subject" and geolocation.



SW_LM_02	Apache Spark	SNAP operators (processing algorithms) parallel processing are implemented using Spark (tile based parallel processing – n nodes – implementing SNAP toolboxes operators)		
SW_LM_03	JAVA	Java Runtime Environment (JRE)		

## Algorithms

N°	Name	Description	Sw Used / Code type	Interaction
Algo_LM_01	SNAP Calibration S1	Calibration algorithm using SNAP-ESA module on each downloaded image	SW_LM_01 SW_LM_02 SW_LM_03	Automatic
Algo_LM_02	SNAP Co-registration S1	Co-registration algorithm using SNAP-ESA to create a stack for S1	SW_LM_01 SW_LM_02 SW_LM_03	Automatic
Algo_LM_03	SNAP Change Detection S1	Change Detection using SNAP-ESA change detection module on the co-registered stack for S1	SW_LM_01 SW_LM_02 SW_LM_03	Automatic
Algo_LM_04	SatCen Change Detection S1	Change Detection using SatCen internal change detection module on the co-registered stack for S1	SW_LM_02 SW_LM_03	Automatic
Algo_LM_05	SatCen Change Characterisation S1	Change Characterisation using SatCen internal module on the co-registered stack for S1	SW_LM_02 SW_LM_03	Automatic
Algo_LM_06	SatCen Change Detection S2	Change Detection using SatCen internal change detection module on S2	SW_LM_02 SW_LM_03	Automatic
Algo_LM_07	SatCen Change Characterisation S2	Change Characterisation using SatCen internal module on S2	SW_LM_02 SW_LM_03	Automatic
Algo_LM_08	SatCen S1-S2 data fusion	Innovative algorithm to fuse the Change Detection Maps from S1 and S2	SW_LM_02 SW_LM_03	Automatic





## User stories-actions GUI requirements

Specific GUI Requirement	Free text description/Image
GUI_LM_01	The search user interface allows a keyword, spatial and temporal search
GUI_LM_02	The search user interface allows selection of products
GUI_LM_03	The search user interface allows the search of exact pairs and time series
GUI_LM_04	The user interface shows the change detection map directly in the basemap
GUI_LM_05	The user interface allows to retrieve the Social Sensing Events
GUI_LM_06	The user interface gives access to logs, metadata and SS events of the change detection map
GUI_LM_07	The user interface allows the download of the (enriched) change detection map
GUI_LM_08	The user interface allows the user to subscribe to the notification service
GUI_LM_09	The user interface allows the user to upload maps to verify the changes

## Control points

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
HI_LM_01	User needs to check visually the quality of the change detection map Uses GUI_LM_01	User checks for any geospatial incongruences (land, lake) or co-registration issues (too much land use change)	YES: Press OK button and is allowed to register the product and send alert to the user NO: Press delete or generate an issue to the operations (that will send it to the developer)
HI_LM_02	User would like to verify the info coming from Sentinel 1 and Sentinel 2 Uses <b>TBD in the next phases</b>	User checks for any incongruences between changes on the maps generated by the two sensors.	YES: Press OK button and is allowed to register the product and send alert to the user NO: Press delete, exclude either the map generated with S-1 and S-2 or generate an issue to the operations (that will send it to the developer)
HI_LM_03	User would like to verify the info coming from satellite images and social sensing Uses <b>TBD in the next phases</b>	User checks for any incongruences between changes on the maps and related news from media	YES: Press OK button and is allowed to register the product with the related info from social sensing and send alert to the user NO: Press delete, exclude the social media part or generate an issue to the operations (that will send it to the developer)

## **Appendix D – Natural Hazards VRC Use Cases Scenario**

### Use Case Scenario Detailed Description

#### Personas

Personas	Representing group	Description
Hazel N.	NERC-BGS Keyworth	Coordinator of the NHP Hazard Impact Modelling group
Timothy A.	HSL Buxton	Develops impact modelling methodology with scientific hazard specialists
Oliver G.	HSL Buxton	Develops impact modelling methodology with scientific hazard specialists (Bob, Steve, Tom, Katy, Graeme)
Robert (Bob) M.	NERC-CEH Wallingford	Hydrological modeller and G2G surface water flooding expert who works closely with Tim
Steve C.	NERC-CEH Wallingford	Hydrological modeller and G2G surface water flooding expert who works closely with Bob.
Tom D.	NERC-BGS Keyworth	Landslide hazard modelling lead
Katy F.	NERC-BGS Keyworth	Landslide hazard modelling specialist who works closely with Tom.
Graeme B.	Flood Forecasting Centre	Provides scientific expertise on flooding and stakeholder feedback to the HIM group. Works closely with Bob, Steve, Tim and Oliver.
Becky H.	Met Office	

Scenarios	
<b>1. Surface Water Flooding hazard impact modelling case studies</b>	HSL, CEH and FFC work together to test different surface water flooding (SWF) hazard and impact modelling case studies for different times/locations based on past observed events. A model is loaded each time with a different time and location to derive results.

### Scenario 1: Surface Water Flooding

Action ID	Action description	Dependencies	Challenge
US_NHP_0	Bob/Steve complete a G2G process locally (outside the	DS_NHP_01	



1	VRE) to produce a SWF hazard footprint (or a set of footprints) for a given time and location in ASCII grid format.	DS_NHP_02 DS_NHP_03	
US_NHP_02	Bob/Steve will zip the SWF hazard footprint and want to upload it to the VRE (DS_NHP_04).		
US_NHP_03	Bob/Steve will need to be able to record what Met Office data they have used e.g. observed rainfall via radar/rain gauge and/or ensemble (multiple) forecast data. Type of input data is an input to tool configuration e.g. if forecast data used as input, the resulting impact output at US_NHP_14 will include a measure of risk.		
US_NHP_04	Tim wants to be automatically notified each time a footprint is uploaded to the VRE (via GUI_NHP_01)	US_NHP_02	
US_NHP_05	Tim needs to extract the zipped data as an ASCII grid using an R script.	US_NHP_01 US_NHP_02 US_NHP_03 US_NHP_04 Software: SW_NHP_02 Data: DS_NHP_04 Algorithm: Algo_NHP_01	
US_NHP_06	Tim needs to pre-process the ASCII grid using an R script. Results: DS_NHP_05	US_NHP_05 Software: SW_NHP_02 Data: DS_NHP_04 Algorithm: Algo_NHP_02	
US_NHP_07	The result will be identified as having been 'post-processed'	US_NHP_06 Data: DS_NHP_05	
US_NHP_08	The result will be saved in a different location on the VRE	US_NHP_06 Data: DS_NHP_05	
US_NHP_09	Tim needs to set different input parameters for the impact modelling that will be applied by the R script. It is expected that the UI will enable him to input the parameters.	US_NHP_08 Software: SW_NHP_02	



		Data: DS_NHP_05 GUI: GUI_NHP_02	
<b>US_NHP_10</b>	<p>Tim wants to run another R script that will apply the impact modelling.</p> <p>He will need to call different input data. Configuration defines which input data are required.</p> <p>Results: DS_NHP_07</p>	US_NHP_09 Software: SW_NHP_02 Data: DS_NHP_05 DS_NHP_06 Algorithm: Algo_NHP_03	
<b>US_NHP_11</b>	Output is in the form of 1km grids which need to be saved in a different location	US_NHP_10 Data: DS_NHP_07	
<b>US_NHP_12</b>	Tim will define input parameters for aggregation of 1km grid data to county level (see US_NHP_09. This will be input via the UI and used by the R script in US_NHP_11.	GUI: GUI_NHP_02	
<b>US_NHP_13</b>	<p>Tim will need to aggregate the 1km grid output from US_NHP_11 to county level outputs e.g. 1 impact value for each UK county. The approach for this will be configured in US_NHP_10.</p> <p>Results: DS_NHP_09</p>	US_NHP_10 US_NHP_11 US_NHP_12 Software: SW_NHP_02 Data: DS_NHP_07 DS_NHP_08 Algorithm: Algo_NHP_04	
<b>US_NHP_14</b>	<p>If initial G2G input includes ensemble forecasting then the county level output will model the risk from surface water flooding.</p> <p>Results: DS_NHP_10</p>	US_NHP_03 US_NHP_13 Software:  SW_NHP_02 Data: DS_NHP_09 Algorithm: Algo_NHP_05	
<b>US_NHP_15</b>	User check of outputs	US_NHP_13 US_NHP_14 Data: DS_NHP_09	



		DS_NHP_10 Software: SW_NHP_01/02/ 03	
US_NHP_1 6	A number of different impact outputs will need to be saved and stored in a separate location with metadata.	Data: DS_NHP_09 DS_NHP_10	
US_NHP_1 7	The team (Oliver, Graeme, Bob and Steve) need to be notified that new outputs are available.	US_NHP_16	
US_NHP_1 8	The team (as above) will check the outputs. If useful output – go to scenario 2. Otherwise, notify Tim and discard. (HI_NHP_04)	Data: DS_NHP_08 DS_NHP_09 Software: SW_NHP_01 SW_NHP_03	
US_NHP_1 9	The HIM group want to be notified when a new impact output has been validated and stored.	US_NHP_18 Control Point: HI_NHP_04	

## Scenario 1: Dependencies

### Datasets/ Products

N°	Name	Source	Volume	Format
DS_NHP_01	Observed rainfall	Met Office Rain gauge		ASCII grid
DS_NHP_02	Observed rainfall	Met Office Radar		ASCII grid
DS_NHP_03	Ensemble (multiple) Forecast Data	Met Office Forecasting		ASCII grid
DS_NHP_04	SWF hazard footprint	G2G		ASCII grid (currently 1km)
DS_NHP_05	Post-processed SWF hazard footprint	US_NHP_06		ASCII grid (currently 1km)
DS_NHP_06	Impact Library	FFC/HSL		ASCII grid (currently 1km)
DS_NHP_07	Impact grids	US_NHP_10		ASCII grid (currently 1km)



<b>DS_NHP_08</b>	County areas	level	FFC reporting areas for county summary		ESRI Shapefile (county level)
<b>DS_NHP_09</b>	County impacts	level	US_NHP_13		ESRI Shapefile (county level)
<b>DS_NHP_10</b>	Risk summary		US_NHP_14		ESRI Shapefile (county level)

## Infrastructures/catalogues

Infrastructure/ Catalogue Name	Infrastructure/ Catalogue Description	Access policy	Com. Protocol	Contacts
<b>Infra_NHP_1</b>	Impact Library			

## Software requirements

N°	Name	Description	Open source/ IPR	Contacts
<b>SW_NHP_01</b>	ArcGIS	GIS software	Off –the-shelf	Tim A.
<b>SW_NHP_02</b>	R	Statistical/programming	Open Source	Tim A.
<b>SW_NHP_03</b>	Python	Scripting/programming	Open Source	Tim A.

## Algorithms

N°	Name	Description	Sw Used/Code type	Interaction
<b>Algo_NHP_01</b>		Zipped data script	R	
<b>Algo_NHP_02</b>		SWF footprint pre-processing script	R	
<b>Algo_NHP_03</b>		Impact Modelling Script	R	
<b>Algo_NHP_04</b>		County Impacts Aggregation script	R	
<b>Algo_NHP_05</b>		Risk calculation script	R	



## User stories-actions GUI requirements

Specific GUI Requirement	
<b>GUI_NHP_01</b>	The user interface will notify me when a new G2G output is available. Date, location and name of event will be attached in some way.
<b>GUI_NHP_02</b>	The user interface will allow me to input and verify input parameters/configurations that will be used by the impact modelling scripts – via pop-ups?
<b>GUI_NHP_03</b>	The user interface will provide a pop-up window to notify me when the impact modelling process is complete. The outputs will be put into a new location/folder following some simple rules.
<b>GUI_NHP_04</b>	The user interface will notify the users when a validated impact output has been created.

## Control points

N°	Description	Conditions (YES/NO)	If YES/NO - Go to
<b>HI_NHP_01</b>	G2G output needs to be uploaded to the VRE		
<b>HI_NHP_02</b>	Once notified that the G2G output has been uploaded, the user runs the impact modelling script		
<b>HI_NHP_03</b>	User inputs relevant parameters when prompted by UI		
<b>HI_NHP_04</b>	User checks the process has run correctly and that the final output has been stored in correct destination folder	YES: output is stored correctly NO: Process has experienced an error	NO: check for errors and re-run process
<b>HI_NHP_05</b>	User checks output	YES: Output suggests parameters used have produced a valid result  NO: Output suggests the parameters used did not produce a valid result	YES: Complete metadata to show result has been validated and is stored as an RO complete with metadata, input parameters, date, location etc. enabling a full re-run of the process (with options for operator-defined modifications)  NO: Complete metadata to show result has been checked and shown not to have produced a valid output. Store as an RO complete



			with metadata, input parameters, date, location etc. in a separate folder. Re-run process with different parameters