REVISED USER REQUIREMENTS

Deliverable D1.2.3

Circulation:	PU: Public
Lead partner:	FÖMI
Contributing partners:	All
Editors:	Márta Belényesi (FÖMI) Dániel Kristóf (FÖMI)
Quality controller:	Ewald Quak
Version:	1.0
Date:	06.06.2014

Consisting of

SINTEF	STIFTELSEN SINTEF, Department of Applied Mathematics, Oslo, Norway
Fraunhofer	Fraunhofer Institute for Computer Graphics Research (IGD), Darmstadt, Germany
CNR-IMATI-GE	Institute for Applied Mathematics and Information Technologies of the National Research Council (CNR-IMATI), Genova, Italy
MOSS	M.O.S.S. Computer Grafik Systeme GmbH (MOSS), Munich, Germany
HRW	HR Wallingford Ltd (HRW), Wallingford, UK
FOMI	Hungarian National Mapping and Cadastral Agency (FOMI), Institute of Geodesy, Cartography and Remote Sensing, Budapest, Hungary
UCL	University College London (UCL), Research centre for Photogrammetry, 3D Imaging and Metrology, London, UK
TUDelft	Delft University of Technology (TUDelft), Department of Earth and Climate Sciences & Man-Machine Interaction Group, Delft, The Netherlands
IGN	Institut National de l'Information Géographique et Forestière (IGN), Paris, France
UBO	Université de Bretagne Occidentale (UBO), European Institute for Marine Studies, Brest, France
Ifremer	L'InstitutFrancais de Recherche pour l´Exploitation de la Mer (Ifremer), Brest, France
Liguria	Regione Liguria, Genova, Italy

DOCUMENT HISTORY

Version[1]	Issue Date	Stage	Content and Changes
1.0	06.06.2014	100%	Version submitted to the Project Officer

[1]Integers correspond to submitted versions

Authors:

Bajjouk, Touria
Belényesi, Márta
Boehm, Jan
Brédif, Mathieu
Brodtkorb, André R.
Dokken, Tor
Fekete, István
Gera, Dávid Ákos
Giachetta, Roberto
Giannoni, Francesca
Gierlinger, Thomas
Kovács, Tamás
Kraut, Volker
Kristóf, Dániel
Lehoczki, Róbert
Lindenbergh, Roderik
Martina, Federica
Michel, Frank
Mikus, Gábor
Millard, Keiran
Molnár, András
Nádor, Gizella
Nguyen, Thai Binh
Olasz, Angéla
Patanè, Giuseppe
Pittaluga, Simone
Rudolf, Heino
Saur, Sébastien
Sobrero, Davide
Spagnuolo, Michela
Stumpf, André
Szabó, Judit

EXECUTIVE SUMMARY

This report reflects in a unified document the project strategy towards the implementation of validation scenarios that will demonstrate the functionalities offered by IQmulus. It contains a clear mapping between user requirements and processing services/visualization modalities, as required during the first project review. According to the service release timelines presented in the document, some of the services and visualization functionalities will be fully integrated in the first functional prototype of IQmulus that is due at the end of the second project year.

In the previous WP1 deliverables on *User Requirements* (D1.1, D1.2.1 and D1.2.2), altogether 139 User Stories (short natural language sentences of the intended functionality of a software system) have been analysed, filtered and prioritized. The outcome of this process was then used to define initial 'Showcases', i.e., basic sets of requirements to drive the early development of the infrastructure, prototypes and basic services in the first phase of the project.

From the beginning the Showcases were expected to evolve further during the next project years, guiding the development of services (WP4) and visualization modalities (WP5), and the implementation of demonstration scenarios (WP6) towards their validation (WP7).

In the first project year one Marine and two Land Showcases were elaborated, containing a synthesis of multiple original User Stories (see D1.2.2 and its extension prepared after the first project review). As a result of the on-going requirement consolidation process, from the perspectives of both users and developers, the Showcases were revised for project year 2. Note that the showcases are expected to evolve again also in the second half of the project.

This report (as D1.2.3 *Revised User Requirements*) now presents three Showcases: an Integrated Land Showcase (merging and slightly modifying the previous two Land Showcases), the Marine Showcase (basically unchanged) and a new Urban Showcase (covering arising user needs and available functionalities), relating them again to existing User Stories. The document then proceeds to higher levels of detail concerning the definition of the capabilities intended for the first functional prototype[2] and beyond. The overall approach to address the Showcases is presented in Section 2, followed by the results for each Showcase being described in Section 3.

Each Showcase is decomposed into workflows (five for the Integrated Land Showcase, four for the Marine Showcase, and two for the Urban Showcase), based on the already existing project outcomes and experiences and providing the framework for the year two developments.

For each separate workflow an *As-Is Analysis* is given, discussing the current methods of the project-internal users to execute this workflow. Then a summary of the user perspectives and expectations and the potential for innovation is provided. The "*Big Data*" issues of *Volume*, *Variation*, *Velocity* and *Analytics* are commented on for each workflow using a predefined table.

The *proposed IQmulus solutions* are then presented against the background of the As-Is Analysis. Each workflow is typically decomposed into the four components of preparation, feature detection, change detection and presentation (visualization). The necessary software tools within each workflow component are classified in three main groups: generic system functionalities (WP2/WP3), WP4 services for data integration and processing (functional and domain), and WP5 visualization functionalities. Finally, input and reference datasets provided for the testing of each workflow are listed in tables. *Input data* are to be processed in the IQmulus workflow. *Reference data* result from the current (non-IQmulus) workflow and are to be used as processing baselines for benchmarking.

[2] For the revised Showcases the updated WP4 services are included as tables at the end of the present document, with the same service numbering as in the eRoom table "WP4 Services and Representations".

TABLE OF CONTENTS

1 INTRODUCTION

This document is a working document, connected to the IQmulus Task 1.4 "Requirement Consolidation", which is designed to run continuously during the first 30 months of the project.

The previous deliverable (Consolidated User Requirements, D.1.2.2[3]) documented the specification and consolidation of requirements collected in previous stages of the IQmulus project (published in D1.1[4] and D1.2.1[5]). The process had to be complemented by the categorisation and prioritisation of requirements, and was carried out by keeping in mind that the well-structured requirement specification must be usable for many Work Packages.

Altogether 139 User Stories (short natural language sentences of the intended functionality of a software system) have been analysed, filtered and prioritized, and the selected User Stories have been translated into Use Cases.

The outcome of the above analysis was used to define 'Showcases', i.e., subsets of user requirements to drive the early development of the infrastructure, prototypes and basic services in the first phase of the project. It is important to mention here that the development is an iterated procedure, based on the feedback of the users, testers and developers.

The term 'Showcase' was used because we wanted to make sure each development iteration had something we could show to users, particularly in early iterations where the users may not fully appreciate how developments are fully aligned with their specific needs.

In the first project year one Marine and two Land Showcases were elaborated, with all of them containing a synthesis of multiple original User Stories to simplify development work, help testing and clarify further tasks while keeping in mind specific user needs.

Now, in Year 2, interaction with professional users is crucial for the further refinement and consolidation of user requirements. As a result of the on-going requirement consolidation process, a decision was made on the revision of these Showcases since the last Deliverable (D.1.2.2), and the further work was adapted to the revised Showcases.

This document analyses the workflows extracted from the revised Showcases, introduces the current solutions applied by the users in the form of an 'As-Is Analysis' per workflow, compares them to the proposed IQmulus solutions for each workflow, and summarizes the services required per workflow.

The Showcases are expected to evolve further during the next project years, and they will be used to guide the development of services (WP4) and visualization modalities (WP5), and to guide the implementation of the demonstration scenarios (WP6). They will also be used during the validation phase (WP7). The requirement collection process will be concluded later in Project Month 30, documented in D1.2.4 "Finalized User Requirements".

[3] D1.2.2 „Consolidated User Requirements" Link on IQmulus website: http://www.iqmulus.eu/images/material/deliverables/Version14Jan_D1.2.2.pdf
[4] D1.1 "State of the art analysis" Link on IQmulus website: http://www.iqmulus.eu/images/material/deliverables/D1.1_State-of-the-art-analysis.pdf
[5] D1.2.1 "Initial User Requirements" Link on IQmulus website: http://www.iqmulus.eu/images/material/deliverables/D1.2.1_Initial_User_Requirements_Version_1.0.pdf

During the preparation of the present deliverable, the following documents were taken into consideration (links are related to eRoom):

- Description of Work document (https://project.sintef.no/eRoom/math/IQmulus/0_2dadb) (in particular the IQmulus system architecture and design)
- D.1.2.2 Consolidated User Requirements (http://www.iqmulus.eu/images/material/deliverables/Version14Jan_D1.2.2.pdf)
- D4.2.1 Spatio-temporal Data Fusion Toolkit (Version 1) (http://www.iqmulus.eu/images/material/deliverables/D4.2.1_Spatio_Temporal_Data_Fusion_Toolkit_Month_12.pdf)
- D4.3.1 Feature Extraction and Classification Toolkit (Version 1) (http://www.iqmulus.eu/images/material/deliverables/D4.3.1_Feature_Extraction_and_Classification_Toolkit_Month_12.pdf)
- D4.4.1 Multivariate Surface Generation Toolkit (Version 1) (https://project.sintef.no/eRoomReq/Files/math/IQmulus/0_33efb/D4.4.1_Multivariate_surface_generation.pdf)
- D4.5.1 Change Detection and Dynamics Toolkit (Version 1) (https://project.sintef.no/eRoomReq/Files/math/IQmulus/0_33efc/D4.5.1_Change_detection_and_dynamics_toolkit.pdf)

2 METHODOLOGY

This document summarizes the results of a deep analysis of user requirements, which was carried out taking into account the suggestions gathered during the first review meeting. Therefore, we decided to include in this document not only the analysis of the requirements per se, but also the analysis of the functionalities (processing services and visualization) that need to be developed to realize the user expectations, focusing mainly on the user's perspectives and the generic workflow development while taking into account the development trends in the project.

Besides WP1, other WPs, namely WP2, WP3, WP4, and WP5 contributed to the preparation of the document. Showcase leaders were responsible for collecting all the necessary information needed from the participants. The document was prepared following guidelines prepared beforehand.

The work has been based on the already existing project outcomes, the requirements gathered for scenarios and the development capacity for the second year. The chapter "Formulation of Showcases and revising User Stories connected to them" introduces the changes of the Showcases established last year, according to the user needs and developer suggestions. The chapter "Decomposing workflows from each Showcase" describes how the process steps and services will be combined into workflows and showcases suitable to fulfill the requirements of the scenarios. This chapter provides an 'As-Is Analysis' for each workflow separately, summarizes user perspectives and expectations towards the project, and innovation aspects of IQmulus according to the workflows. The aim of the chapter "Proposed IQmulus solutions" is providing a methodology for the following purposes: (1) introducing the IQmulus solutions for the current workflows listed in the 'As-Is Analysis', (2) providing appropriate bases for testing services and workflows during the development work of the IQmulus platform further on, (3) in a way fulfilling the most urgent user needs.

The main steps of the whole process are listed below, with details presented in the next sections:

1) Revision of existing Showcases:
    a) Re-formulating the Land Showcase.
    b) Elaborating a new Urban Showcase.
    c) Detailing the Marine Showcase.
    d) Revising the corresponding User Stories per Showcase.
2) Examination of the current solutions for the tasks arising from the Showcases:
    a) Decomposing Showcases into workflows.
    b) Performing an 'As-Is Analysis' per workflow, providing short descriptions of the steps setting up the current workflows.
    c) Summarizing user perspectives and expectations, innovation aspects and the relevance to the 'Big Data' issues of each workflow.
3) Detailed descriptions of the workflows as they are planned to be realized in Year 2 (Proposed IQmulus solutions):
    a) Introducing IQmulus solutions per workflow.
    b) Defining workflow components. Revising corresponding services per workflow components according to D1.2.2.

c) Characteristics of sample and reference data available for testing the services listed in the workflows.

This generic approach provides the framework for the year two developments and will be updated and extended further along the project runtime, particularly by allowing more complexity.

2.1 FORMULATION OF SHOWCASES AND REVISING USER STORIES CONNECTED TO THEM

As a result of discussions among the consortium partners, the three Showcases (Super User Stories) formulated in the first project year – and published in D1.2.2 – have been slightly re-formulated, and a new Urban Showcase was created according to the user needs and developer suggestions received thus far, as follows:

1) The Marine Showcase was not changed.
2) The Land 1 and Land 2 Showcases were merged and slightly modified as a single Integrated Land Showcase.
3) A new Urban Showcase was formulated to cover arising user needs and available functionalites.

Each Showcase is intended to collect and integrate User Stories and their related Use Cases from the Land and Marine Testbed, respectively. The lists of User Stories connected to the Showcases have already been published in D1.2.2. In the current document (D1.2.3) revised and updated lists adapted to the re-formulated Showcases are introduced, but there are no significant changes regarding the list of User Stories.

2.2 DECOMPOSING WORKFLOWS FOR EACH SHOWCASE

Formulation of workflows

Each showcase was analysed in order to identify its main *workflows,* that is, the logical components needed to realize the showcase. As a rule of thumb, a workflow produces a relevant piece of output which contributes to the showcase. Typically, users may want to interact/visualize, or store for further usage, intermediate results produced by workflows. The number of workflows per Showcase depends on the complexity of the showcase: thus four workflows are included in the Marine Showcase, five are included in the Integrated Land Showcase, and the Urban Showcase consists of two workflows. Workflow relations within the showcase are also mapped, showing what kinds of dependencies exist, which workflow provides input to another one, etc.

This generic approach provides the framework for the year two developments that will be updated and extended further along the project runtime, particularly by allowing more complexity.

Performing “As-Is Analysis” and summarizing user perspectives per workflow

After the basic formulation of the workflows per showcase, it is important to introduce the user perspectives and the “As-Is Analysis” for each workflow separately.

"As-Is Analysis" is an important step in the development process. This type of analysis provides an understanding of the current state, covers the current solution for the tasks described by the workflows, and is a good starting point for our solutions. As-is analysis can also help to identify the added value in a project. It can provide insight into the scope of the problems by helping to identify gaps and processes that require further analysis.

The main steps of preparing the 'As-Is Analysis' per workflow in IQmulus are:

1) *Introduction and main steps of the current solution:* short summary of the current workflow solution we would like to improve with the IQmulus tools and environment developed during the project. The summary refers to the existing software environment and implementation, the degree of automation in the current solution (automated, semi-automated, ...), etc. The main steps of the solution are also listed in most cases, grouped by the following topics:

   a) pre-processing /processing/results, or

   b) following any other logical aspects, which are important from the user point of view.

   A short description of the steps making up the current workflow is documented, and a Use Case Diagram is provided where available. A Use Case Diagram (UCD) can be used to describe a system or sub-system functionality. A UCD does not show the details of the use case: it only summarizes some of the relationships between use case, actors, and system. In particular, the diagram does not show the order in which steps are performed to achieve the goals of each use case; the present use cases only deal with functional requirements for a system. For a detailed description of setting up Use Case Diagrams we refer to D1.2.2.

2) *Summarizing the user perspective and expectations as well as the innovation aspects*: the aim here is to summarize the disadvantages or shortcomings of the current solution, introducing innovation aspects that IQmulus could provide concerning the given workflow, and highlighting the "Big Data" aspects of the workflow. "Big Data" issues (Volume, Variation, Velocity, Analytics) are synthesized in a predefined table for each workflow.

Proposed IQmulus solutions per workflow

The aim of this section is to introduce what IQmulus will provide against the background explained in the As-Is Analysis. For each workflow a general description is presented, and the novelty compared to the "As-Is Analysis" is highlighted.

For the sake of the workflow description, it was suggested to structure them into (1) preparation, (2) feature detection, (3) change detection, and (4) presentation components. For some of the workflows, some of the components might be missing or not meaningful.

The working showcases to be presented at the end of year 2 therefore consist of several executions of workflows in a sequence. For the second year the project infrastructure (WP 3) will allow the execution of showcases following this concept. The visualization tools of the project infrastructure will at this stage only be used in step 4 (presentation) of the process.

The workflows describe the goals we are targeting for the next project years in terms of user requirements to be satisfied. The showcases are described so as to reflect properly how IQmulus can be used to create complex analysis pipelines using simpler processing services and visualization functionalities. The showcase analysis also inspired the identification of the processing services needed to realize them, and suggested practical examples against which the

infrastructure components will be tested (data storage and access, data sharing and browsing, interaction and visualization capabilities, workflow management, service interface,..).

The results obtained by IQmulus, to be presented at the end of year 2, will be measured against the targets delineated by the workflows, and all services contributing to the implementation of the showcases will be discussed, and possibly presented in an integrated fashion into the IQmulus platform.

For consistency, the description of the workflows in the current document typically has the following structure

1) Preparation component

   The preparation step covers all necessary work to provide the available data in a suitable structure. This covers e.g. transformation and format change but also covers changing the representation of the data. Additional aspects of the preparation step are spatial and thematic filters and/or segmentation functions.

2) Feature Detection component

   Feature Detection covers the identification of information in the dataset provided by the preparation step. Again the step can vary from low to high complexity starting with template-based feature detection from only one data source, up to multi- sourced processes combining the different information available from different data sources.

3) Change Detection component

   The change detection step in the early stage of the project will be carried out by comparing the development of a region over time or from different sources. Change detection might result in an indication of a change (e.g. coordinates where something changed) or a description/representation of the change (e.g. volume representation of the differences between two surface models).

4) Presentation/Visualization component

   The end of the workflow is the presentation step. The presentation step will in most cases be the visualization of the workflow result. Thus it might be a visualization using the tools provided by WP 5, showing a list representation of the results and offering data provision to third party tools,.

Software tools involved in the given IQmulus workflow are described following the grouping above, and each of the tool is also classified according to the work package responsible for their development, namely, *generic system functionalities* (WP2/WP3), *processing service* (WP4), and visualization functionalities (WP5), supporting users on visual decision making and interactive visual communication.

After the revision of Showcases, the WP4 Service Table will be updated accordingly. The updated services are included as tables in an annex of the present document and the numbering of the services is the same as in the eRoom table “WP4 Services and Representations” (https://project.sintef.no/eRoom/math/IQmulus/0_32a05).

We note at this point that the WP4 service development plan will be refined and revised in D4.1.3, and the Service Table in the eRoom system will always maintain the updated status of discussions about services and their development stage. The detailed description of the services in the annex covers the following aspects:

- the name and description of the service;
- the name of the component the service belongs to within the given workflow;

- type of the input needed, and output generated;
- responsible partner;
- steering parameters needed;
- status/deliverable date of the service (if the services are not released yet, the exact month of the development deadline has to be indicated).

Input and reference data provided for workflow development and testing

Finally, input and reference datasets provided for the testing of each workflow are listed and described in tables. **Input data** are data to be processed in the IQmulus workflow. By **reference data** we mean data resulting from the current (non-IQmulus) workflow to be used for setting the processing baseline for benchmarking.

Besides the name, ID and short description of the datasets, one can get information about the spatial resolution, spectral resolution and other data characteristics like dimensionality. In the descriptive part information on the acquisition technique, source, licence needs, reference system, size and coverage, data structure and format are shown. The ID comes from the table *"Metadata table for sample datasets"* in eRoom, which also summarizes the most important information about the datasets (https://project.sintef.no/eRoom/math/IQmulus/0_2e395).

2.3 DOCUMENTATION

The present Deliverable 1.2.3 contains the following results:

- Three showcases, from which one is original from D1.2.2, one is arising from D1.2.2 and was re-formulated, and one is a new additional showcase which is focusing on building detection in urban areas;
- Updated list of User Stories connected to each showcase;
- Workflows that make up each showcase;
- As-Is analysis of each workflow(the 'As-Is' analysis is the assessment of the current situation, meaning the description of the present solution for the same task we would like the IQmulus system to work on);
- User perspective and expectations as well as innovation aspects concerning the given workflow;
- Textual descriptions of the workflow components;
- Use Case Diagrams illustrating the stepwise activities and actions in a workflow, where available;
- Updated list of services connected to each workflow in the present document;
- List of datasets provided for testing per workflow in the present document;
- Updated list of datasets provided for testing in eRoom (https://project.sintef.no/eRoom/math/IQmulus/0_2e395).

Methodological choices have been described for and shared with all IQmulus partners from the WPs involved, partly through documents sent via email, and/or uploaded to the eRoom folder "Showcases" (https://project.sintef.no/eRoom/math/IQmulus/0_33659).

Methodological guidelines were provided for each showcase leader and participant for preparing the "As-Is" analysis, writing about proposed IQmulus solutions and mapping the workflow relations within a showcase in order to provide a consistent documentation.

The guidelines are uploaded to eRoom:
(https://project.sintef.no/eRoomReq/Files/math/IQmulus/0_33cb7/Showcase_user_perspective_guidance+example_FOMI_2014-04-17.docx).

3 RESULTS

3.1 SHOWCASES

In the Description of Work it was stated: "Within the IQmulus project, two test beds are built up to evaluate the approach and resulting technology in real-world scenarios". These are based on 'as is' scenarios that could be improved by IQmulus and are user-driven based on the users in IQmulus. The consortium has the data and models for these test beds and to operate the 'as is' workflows for these test beds.

The two test beds and associate trials in the DoW are as follows:

- Land

  Trial 2.1 Rapid response for floods, flash floods and landslides

  Trial 2.2 Rapid response for Industrial accidents

  Trial 2.3 Disaster management in urban areas

  Trial 2.4 Territorial Management

- Marine

  Trial 1.1 Morphological coastal changes

  Trial 1.2 Identification of underwater seabed features

  Trial 1.3 Scour Risk Forecast

The test beds are labelled as 'marine' and 'land' but this is simply a label related to the user domain they refer to. There is clearly too much complexity to deliver an IQmulus solution for all of the trials all at the same time, so for this reason the concept of showcases was realised that would implement the key (generic) functionality applicable to each of the test beds in the early various stages of the project. Three Showcases were formulated in Year 1, which were slightly re-formulated at the beginning of Project Year 2. The three resulting showcases are:

- Integrated Land Showcase
- Marine Showcase
- Urban Showcase

The detailed description of the showcases is given below.

The Integrated Land Showcase (1.2.2_SC2) and corresponding User Stories

The showcase exemplifies how IQmulus could respond to the requirements and expectations of expert users (typically hydrologists or geologists) supporting decision makers in civil protection in events related to flooding. The services integrated in the showcase represent the IQmulus answer both to explicit user requirements (e.g. how to handle data sets that are too big for current techniques) and to unexpressed ones (e.g. users may not be aware of available visualization or data management techniques and therefore unable to express a specific requirement in that direction). In this sense, the showcase integrates in a unique scenario the user stories selected in the original land-related showcases, and enriches them with additional

and "futuristic" aspects. The partners contributing to the showcase, from a user perspective, are mainly Regione Liguria, CNR-IMATI, FÖMI, and UBO. The showcase reflects therefore the expectations of the users they represent, and also the expertise they can contribute.

The text below describes a usage scenario of IQmulus from the perspective of expert users who support decision makers for risk management in critical hydro-meteo events.

> *As hydrologist or geo-morphologist supporting decision makers in civil protection, I want to analyse data measured during critical events to prepare better prediction and monitoring of floods and landslides. To this end, I want to study the evolution of measured precipitation data as well as slope deformation from optical images, compute parameters to produce high-quality input for hydrological and mechanical modelling and simulation, and compare the results to reference measurements obtained for flooding events and landslides.*

The goal of the scenario above is to demonstrate the usefulness of IQmulus to compute rapidly, even roughly, parameters needed to run flooding simulation or landslide models. To this end, it is necessary to elaborate quickly large and heterogeneous datasets, the solution being not only implementing a high-performance computing environment, but primarily devising an intelligent handling and storage of the data.

In the future, experts supporting the preparation of emergency plans could experiment in a reasonable computing time and within an easy-to-use platform with studies of:

- Historical data, e.g.:
  a) How much rain was collected in this basin?
  b) What was the expected flooding?
  c) What did actually happen?
- Different scenarios, e.g.: how much water will there actually be?
- Different actions, e.g.:
  a) place sandbags in location A or in location B:
  b) what is the effect?

This requires rapid configuration of scenarios for running (efficient) simulations and interaction to enable experts to try different settings with realistic data.

The User Stories strongly related to this showcase are as follows:

- 1.1_52: As a hydrologist, I want to extract significant breaklines from high resolution DTMs or point clouds, such as dikes, dams, banks, ditches so that I can improve my hydraulic model and make more accurate simulations.
- 1.2.1_62: I want to create an automatic dam recognition algorithm (automatic feature extraction, eCognition) especial for dam break lines, to be able to monitoring the water height and the dam height differences.
- 1.2.1_64: As hydrologist supporting the decision maker in civil protection I would like to use radar reflectivity field to interpolate observed rainfall data from traditional rain gauge network (punctual data)
- 1.2.1_65: As hydrologist supporting the decision maker in civil protection I would like to have a shape file of the Ligurian small basins (< 10 km2) derived from a high resolution DTM exploiting the available Lidar data of Regione Liguria to derive the DTM and a cut interpolated rainfall field over the shape file.

- 1.2.1_50: I want to be able to derive landslide parameters (geometry / speed / volume / depth / type etc.) from multi-temporal aerial photography and DTM data.
- 1.1_54: As a hydrologist, I want to combine flood simulation models with all kind of topographic data in order to know the flooded area and all the human and economic stakes impacted so that I can provided decision makers with accurate and reliable information.
- 1.1_33: As a GIS expert, I want to delineate the extent and the maximum possible extent of the area potentially flooded in case of an industrial disaster (e.g. dam burst of a waste reservoir) so that I can provide the polygons of the risk area to the disaster management authorities.
- 1.2.1_16: I want to get an overview map of areas affected potentially by flood, in order to make the best decision for minimize the potential impacts of the disaster.
- 1.2.1_31: I want to interactively modify a water level predicted height and see the affected land parcels on the cadastre so that I can expose maps of submersion areas
- 1.2.1_67: I would like to have easy tools to integrate up-to-date data about hydraulic, hydro-geological, seismic, industrial risk for the production of multi-risk maps
- 1.1_53: As a hydrologist, I want to combine flood simulation model with high resolution DTM so that I can publish a water height map on the flooded area.
- 1.1_30: As a GIS expert I want to generate a categorized waterlog map (with at least the following categories: open water surface, flooded vegetation, wet soil) in a particular area in order to provide data for a quicker and more reliable agricultural damage assessment and yield loss estimation.
- 1.2.1_18: I need to get the boundary-polygon of the affected area, and the location of the dam break on a quick and automated way, in order to provide the overview map and the basic statistics for the decision-maker in case of a disaster.
- 1.2.1_13: I have to generate a categorized inland inundation map in order to provide data for agricultural field experts and office experts.
- 1.1_28: As a field expert in disaster management, I want to perform real-time updating of road status attributes in an emergency situation (flooding or high-level inland excess water) on my mobile device so that I can make route planning by taking into account the flooded areas or other barriers.
- 1.1_32: As a Remote Sensing expert I want to receive answers about the damage of a land parcel (natural: water, drought, hail, disease; damage caused by game/ animals; agrotechnical errors), in order to be able to estimate the scope and the extent of the damage.

The showcase, it’s current as-is solutions and the proposed IQmulus solutions are described as split into the following workflows, where the names of the reference partners are also indicated:

- (LS1) Organization of the terrain data according to their membership to main alert area (terrain model preparation) [Regione Liguria];
- (LS2) Analysis and modelling of the observed rain (precipitation analysis) [Regione Liguria];
- (LS3) Detection and characterization of flood and waterlogging [FÖMI];
- (LS4) Detection and characterization of landslides [UBO];
- (LS5) Simulation of floods with observed rain data [FÖMI, Regione Liguria, UBO].

The Marine Showcase (1.2.2_SC1) and corresponding User Stories

The Marine Showcase is formulated as follows:

> *I want to create a seamless land/underwater elevation model by the integration of land and underwater data sources (topographic and bathymetric LIDAR and SONAR point clouds, existing digital elevation models and surface models in different formats - data models - resolutions), and want to extract the shoreline based on it, to obtain a seamless data product that can also be used in further analysis and processing tasks.*

The showcase is demonstrating how the current process for fusing multiple surveys and surface generation can be improved both in terms of automation and processing spent. These improvements will enable surfaces to be generated for much larger areas and also in response to user control. *The user control requires a new approach to the current deconfliction method so surveys selected by users can be combined without the need for human interaction.* The *user – controlled* or *interactive deconfliction* planned to be implemented in IQmulus will combine the metadata of the different datasets with geometrical properties of the different datasets as well as their spatial relationships.

An integrated approach for *user-guided interactive deconfliction* can by itself already considerably facilitate the work of expert marine users. Still, in addition, it is necessary to tackle the growing computational challenges of the methods involved in this setting. For this purpose we envisage to introduce a hybrid and fairly compact representation of the shapes occurring in a marine DTM, allowing the reduction of "big data to small data".

The shape representation in the marine showcase is thus envisaged to be a combination of point sets, triangulations and spline (piecewise polynomial) representations.

- Spline representations are very well suited for smooth shapes, and large areas of the sea bottom are indeed smooth.
- Triangulations are very well suited for the representation of the non-smooth areas of the sea bottom.
- Point sets are then still used as a representation of areas where a good spline representation or triangulation is not yet available. An example of such areas is a smooth muddy sea bottom with scattered boulders. A spline surface will fit the bottom well, while the boulders will probably be a set of very small collections of points. These small sets can be represented as triangulations or just as point sets. A similar situation might arise for a triangle representation where a small number of points is not consistent with the major trend of the triangulation.

As the processing starts from point clouds, either general ones or in a raster structure, it is natural to identify portions of the point clouds that in an incremental process gradually can be represented in a compact way either by splines or triangulations. The final result will be a hybrid representation of point clouds, triangulations and splines, where the remaining point sets are possibly semantically tagged. The approach will draw heavily on the longstanding experience of some consortium partners in shape representations, especially from CNR-IMATI (for triangulations) and SINTEF (for piecewise polynomial splines).

The following aspects of IQmulus functionality are being demonstrated in this showcase:

- (MS1) Marine DEM generation: rapid and flexible generation of a single seamless surface from multiple disparate point cloud source data and associated visualization;

- (MS2) Inspection of the quality of the representation;
- (MS3) The use of a single seamless smooth spline surface DEM for identifying parts of the point data not behaving according to the smoothness hypothesis defined by minimal wavelengths of variation and accuracy thresholds. This part of the data either represents features, or parts of the model that the smooth representation cannot accurately represent. The classification of these parts of the point data into feature categories is not addressed in the second year.
- (MS4) Change detection by comparing DEMs from different dates. In year 2 the focus will be on raster represented DEMs.

The need for seamless elevation models of the seabed is a well-established user requirement identified in Deliverable D1.2.2 (Consolidated User Requirements). It is also a key input dataset to the three marine trials identified (morphological change, identification of underwater features and scour risk). It also has direct synergy with the land DEM. This showcase also informs other parts of IQmulus such as input data storage, intermediate product data storage and visualization.

The marine DEM generation advances current approaches to marine DEM generation by removing the level of human-computer interaction required in the data preparation steps and reducing the time taken to generate a surface. Both of these are steps towards interactive and on-demand surface generation in response to user inputs.

The feature extraction based on splines advances current approaches by allowing the non-smooth parts of the point data to be sorted out from the smooth parts, thus enabling the actual feature classification to later work on small excerpts from the large point data.

The related User Stories for the Marine Showcase are listed below:

- 1.1_47: As a GIS expert in a data production institute, I want to co-register topographic LiDAR data and bathymetric LiDAR data acquired at different times so that I can perform sea/land classification.
- 1.1_48: As a GIS expert in a data production institute, I want to perform accurate sea/land classification of combined LiDAR datasets so that I can produce a continuous sea/land DTM.
- 1.2.1_ 42: I want to merge topographic and bathymetric data on the land-sea interface (data coming from different horizontal and vertical datums)
- 1.2.1_60: I want to make a data fusion in between SONAR, LiDAR and satellite images (boundary of the waterbed) and point dataset to get the best digital surface model (underwater, seamless DEM) to be able to measure in the section of the riverbed.
- 1.2.1_69: I would like to have tools for integrating data about the coast line and seafloor in a unique surface model which allow me to work with large data (Lidar), with historical data (validation, reliability of time series), with better integration results at the border of tiles.
- 1.2.1_75: I would like to have tools to study the dynamics of sediments in the seafloor (volumes of sediments moved in temporal series).

The Urban Showcase (1.2.2_SC3) and corresponding User Stories

In this chapter a general description of the new Urban Showcase is given, followed by a list of related user stories.

I want to update my existing 3D catalogue of urban topographic objects given a new data set. For that purpose I want to be able to

- *remove non-static objects (like cars, rubbish bins, bikes, people,...),*
- *characterize changes (addition, removal, deformation) in static objects (like trees, bus-stops, facade elements, chimneys,...),*
- *possibly include new or exclude existing categories (e.g. remove category of roof-top antennas; add category of charge points for electric cars).*

The following workflows were generated to demonstrate IQmulus functionality in this showcase (clearly without trying to be exhaustive):

- (US1): Detection of buildings for monitoring and cadastral updating (long workflow)
- (US2): Individual tree extraction from urban LMMS data (short workflow).

User Stories related to the Urban Showcase are listed below.

- 1.2.1_22: As a GIS expert (geodata provider/integrator) working at local level, I want to capture topographic objects (such as cable networks, street edges, urban furniture, traffic lights...) from data acquired by mobile mapping systems (LiDAR point clouds, images) so that I can create or update topographic city maps.
- 1.2.1_20: As a GIS expert (geodata provider/integrator) working at local level, I want to provide architects and urban planners online access to city 3D model so that they are able to integrate their projects into the model and share it with administration and citizens for communication and project assessment purposes.
- 1.2.1_19: As a GIS expert (geodata provider/integrator) working at local level, I want to provide citizens with access to virtual navigation tool through 3D city model using a simple light web client embedded in any kind of web browser.
- 1.2.1_21: As a GIS expert (geodata provider/integrator) working at local level, I want to automatically detect individual trees from a Lidar point cloud in a urban area, so that I can monitor growth and foresee pruning works.

Additional example user stories for Trial 2.4 Territorial management, in the field of spatial planning (these User Stories are not registered in RedMine. IDs are eRoom-IDs):

- (#1) I want to compare the development plan with the implemented results in a simple and quick way to explore contradictions, differences and similarities.
- (#2) I want to compare the development planning stages with different format/resolution/origin on a simple and quick way to explore contradictions, differences and similarities.
- (#3) I want to compare time series air photographs, high resolution satellite images, DTM/DSM with high spatial accuracy and vector layers owned by the Institute (VÁTI) on a simple and quick way in order to clean the data and monitor changes in the spatial structure and land use.

3.2 WORKFLOWS

Workflows of the Integrated Land Showcase

The workflows envisaged to realize the Integrated Land Showcase are:

- Land workflow 1 (LS1): Multi-resolution model for land monitoring;
- Land workflow 2 (LS2): Analysis of observed rain;
- Land workflow 3 (LS3): Flood and waterlogging detection;
- Land workflow 4 (LS4): Detection and characterization of landslides;
- Land workflow 5 (LS5): Comparison of simulated floods/landslides with observed data.

Workflow relations within the showcase

For the Integrated Land Showcase a workflow-relation diagram is presented in Figure 1.

The result of the workflow LS1 can be re-used to analyse the same area by the same users, exploiting the novel organization of terrain data in a multi-resolution structure which is more efficient and effective for this context of users and analysis.

In parallel, LS2 will exemplify how time series, or even real-time measures, of rain data can be used to model and analyse the precipitation field over an area of interest.

The effects of heavy rain, as captured by satellites for instance, will be integrated in the showcase by LS3, which will provide users the possibility to detect floods and waterlogging, and by LS4, which will provide support for the detection and characterization of landslides.

Finally, LS5 will demonstrate how IQmulus could be used to build scenarios where observed data are used to run simulations and eventually compare the results of the simulation with observed events.

The multi-resolution model for land monitoring (LS1) and the precipitation field model (LS2) will be used by LS5 for producing the input for a flood simulation model. For year 2, LS5 will be limited to the production of input for the flood simulation, based on the results of LS1 and LS2. We will work further on this workflow in the next project periods. The title of the LS5 is kept general, for consistency with its role in the showcase.

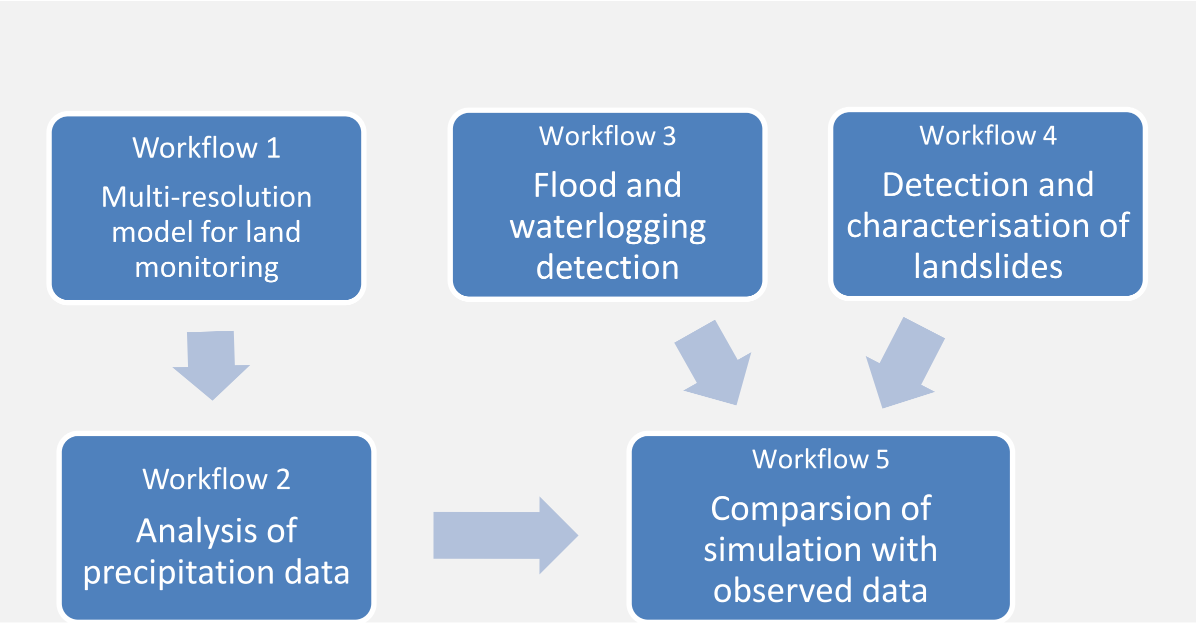


Figure 1. *Workflow-relation diagram for the Integrated Land Showcase*

LS1 - Land Workflow 1: Multi-resolution model for land monitoring

Goal: preparation of terrain data in a multi-resolution model constructed in order to optimize the analysis of hydro-meteo events

Description and As-Is Analysis

Introduction and main steps of the current solution

Currently, Regione Liguria is the owner of a lot of local data and information. The software and methodologies used by Regione Liguria for the preparation of terrain data are in line with standard approaches and practices for terrain data modelling, namely:

1) State-of-the-art GIS systems are used (e.g. GEOMEDIA, ERDAS) to produce DTM and maps.
2) The maps produced are validated and published as reference topographic models of the Regione Liguria.
3) The scale at which the official models have been released is up to the scale 1:5000, with a DTM built on a regular grid (5mx5m) used to extract other thematic maps such as for drainage basins and networks.
4) The validation process includes in situ measures undertaken by municipalities.

User perspective and expectations, innovation aspects

The resolution available is not sufficient to characterize the morphology of small scale basins (area of 10km^2): the Liguria territory is characterized by many small-scale basins that highly influence the response of the territory to heavy rain events. Recently, Lidar techniques have made available data at a resolution which could help experts concretely: If scalable and robust methods to extract small-size basins could be made available, they could help experts understand better the behaviour of the territory with respect to flooding and landslides. But the handling of big Lidar data (now covering about 65% of the Liguria region, expected to rise up to 85% in the next year) is still problematic with current technologies; also, drainage basin extraction could be improved by utilizing a triangulated model instead of a grid, to better model the details of the terrain.

The relevance of the workflow to the "big data" issues is synthesized in Table 1.

Table 1. Big Data issues for the LS1 workflow

Indicator	Relevance	Comment
Volume	High	Time series of acquired data (large volumes of small-sized data) Lidar datasets (for the Regione Liguria, approx. 1 Tera)
Variation	High	Rain data (representative of data streams measured every 5minutes, different sources of data) Lidar, DTM, cartography Image data (optical, radar) Simulation data (hydrological and mechanical)
Velocity	Medium	Rain data are “real time”
Analytics	High	Identification of small scale drainage basins

Proposed IQmulus solutions

The solution proposed within IQmulus is to cope with the "big data" trying to devise an intelligent data management approach, which is tailored to the requirements of the users in this specific context. A hydrologist is interested in having reliable instruments and models to make predictions and carry out monitoring on river basins. On the other hand, a geo-morphologist is focused on change detection of a specific area, for example to evaluate land-slide movements. Both these types of experts need effective, simple and fast applications able to elaborate a huge amount of data and to give an easy-to-understand result in order to take crucial decisions during events.

In this showcase, the requirement is to be able to handle high-resolution data about the terrain, but these data are likely to change slowly, for instance, when new surveys are run. The need for an increased performance of the processing here is not so much in terms of terrain model creation or updating, but rather on the efficient consumption of these big models, either in terms of efficient visualization or in terms of efficient geomorphologic analysis.

In this showcase, there is an interesting *mix of scales and resolutions* required by the users. The analysis of precipitation or landslides may occur at regional or local scale. For instance, to analyse the precipitation field evolution, one would look at the regional scale (in this case, the area of Liguria is approx. 5500km^2), with rain gauge networks providing sparse data over the whole area. For the analysis of precipitation effects in relation to flooding and landslides, the focus is instead usually localized within a single basin whose scale may be small, down to approx. 2km^2. The user therefore needs to switch quickly and easily from regional scale to small basin scale, and the resolution of the data necessary at each step varies.

The novelty proposed by IQmulus is to prepare *terrain data in a model which supports effectively the multi-resolution and multi-scale analysis described above*. We propose to build a terrain model based on a triangle mesh but organized in a multi-resolution structure which reflects the hierarchy of drainage basins. Technically, we assume that high-resolution land data are given as Lidar point clouds, as in the Regione Liguria case. As the representation and drainage basin extraction over the whole Regione Liguria at the full resolution is unfeasible, we plan to re-organize the Lidar data into sub-sets according to "semantically relevant" criteria, namely, according to the membership of the points to areas of relevance for the application domain. We have identified two possible solutions to explore: areas defined by the Catchment Area Plan (*ambito di bacino* in Italian*)* or areas defined as Warning Areas . The Catchment Area Plan groups all the neighbouring basins with similar administrative and geomorphological characteristics. Over the entire region there are about 20 catchment areas for the slopes towards the Tyrrhenian sea. The northward slopes are included in the River Po catchment plan. The warning area is a homogeneous meteo-hydrologically responsive area. This uniformity of

response is due to the similarity of both environmental and administrative constraints. The boundaries of the warning area are overlapping with the boundary of municipalities.

These boundaries are available from the related cartography owned and maintained by the Civil Protection group (shape file). See Figure 2 for an example of such a map.

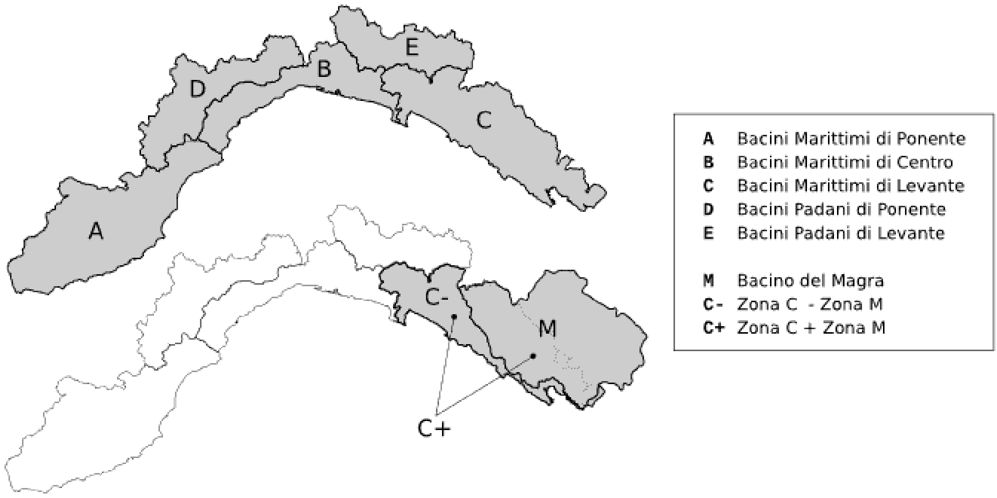


Figure 2. *Liguria Warning Area map*

This partitioning allows us to work independently on each sub-set of data. Within each sub-set, we plan to build the multi-resolution representation, where the organization of the resolution/level of details will reflect the hierarchy of the drainage basins[6] of smaller scale. The preparation of the terrain data involves therefore a sequence of processing steps related to point re-ordering the point cloud(s) and splitting them into the semantic areas, filtering, generating a triangle mesh for each area, extracting the drainage basins, and finally building the multi-resolution model with a final step of surface generation.

Workflow components

The components of the workflow and the corresponding services are described below. Note that in this workflow the feature detection step is actually needed for setting up the land model. The context of usage indeed provides the right "semantics" for organizing large point clouds in an effective manner. Note also that the services proposed for this workflows are general: in this case, we will be using drainage-related areas as thematic information for splitting the point cloud, but any other thematic layer could be used is the user interest is different (e.g. risk map)

Preparation component

Input terrain data (e.g. LIDAR point clouds) are re-organized according to the areas defined by the boundaries of the areas with a clear semantic meaning and usage in this specific usage scenario (Catchment Area Plan or Warning Areas, which are named hereafter *semantic areas*).) After the partitioning, data filtering and reduction is performed. Then, a triangle mesh constrained to the semantic area's boundaries is built for each of the partitioned data sets (note: the triangulated basins will match nicely at the borders, no need to explicitly stitch them together).

[6]http://en.wikipedia.org/wiki/Strahler_number

Generic system functionalities required:

Requirement name	*Functionality*
Definition of area of interest (Select AOI)	The user has to input an area of interest.
Import data	The users select and imports data belonging to the area of interest (point clouds and vector layer with semantic area boundaries)

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Vector layer partitioning of a point cloud (Splits/orders the available terrain data in areas defined by drainage boundaries)	LS1_1_1	New (Service proposed)	Second year PM24
Outlier Classification in Point Cloud (Detect (classify) outlier points in point cloud)	LS1_1_2	Service #10	Released
Resampling of Point Cloud (Resample an unordered point cloud to match a given point density)	LS1_1_3	Service #35	Released
Sub-sampling (thinning) of Point Cloud, preserving high-level features	LS1_1_4	Service #36	Released
Constrained triangulation Create a triangulation of a point cloud preserving user-defined linear constraints (e.g. boundaries, feature lines)	LS1_1_5	Service #49	Released

WP5 visualization functionalities required:

The users may want to see the result of this step, selecting a single semantic area; the point cloud within the selected area is visualized.

Requirement name	*Functionality*
Visualize Vector Layer partitioning of a Point Cloud	**Data to be visualized**: - vector layer used by LS1 to partition the original point cloud - the point clouds within each closed boundary in the vector layer resulting from WP4 services 35 and/or 36. **Output:** 3D visualization of the point cloud within the selected "semantic area"

Feature detection component

For each surface patch, the detection of drainage basins at the smallest scale possible is performed. The hierarchy of drainage basins and the river network is extracted, and the mesh is annotated accordingly; possible labels are: `basin_id` for facets, `river_id` for edges, `branch_id` for vertices. The labelling is used to store the terrain model in a multi-resolution structure (LOD related to the scale of drainage basins).

Generic system functionalities required: N/A

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Detection of flow lines and drainage basins from triangle meshes (From the output meshes - vector layer – generated in LS1_5, drainage basins of each area detected)	LS1_2_1	Service #92	Second year PM24
Multi-resolution triangulation for land monitoring (Building of the multi-resolution mesh together with the drainage basins hierarchy. Storage to the IQmulus platform needed.)	LS1_2_2	Service # 48	Third year PM30

WP5 visualization functionalities required: see the presentation component

Change detection component: none

Presentation/visualization component

For devising visualization services for the multi-resolution model, it is necessary to implement services that extract the triangle mesh at the level of detail required by the user. The visualization may include in this case also the colour coding of the drainage basins, on which the multi-resolution is built.

Generic system functionalities required:

Requirement name	*Functionality*
Multi-resolution model visualizer	The user visualizes the multi-resolution model for land

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Extract the triangle mesh to be visualized at the given LOD (Extract the triangle mesh for visualization at a user-selected level of detail, with colour coding of drainage basins; the scale/level of detail is a parameter of this service)	LS1_4_1	New (Service proposed)	Third year (PM30)

WP5 visualization functionalities required:

Requirement name	*Functionality*
Visualize Triangle Mesh	**Data to be visualized**: - triangle meshes extracted by LS1_4_1, with colour coding/labelling of drainage basins **Output:** 3D visualization of the triangle mesh, with colour coding/labelling of drainage basin

Input and reference data provided for workflow development and testing

Input data

Dataset ID	Dataset title	Dataset short description	Spatial resolution: Resolution distance (m) or scale	Spectral resolution: Number of bands (for RS data)	Spatial resolution: Point density (points/m2)	Data characteristics: Dimensionality
#36	Liguria-DEM	DEM of the whole Regione Liguria. Source: Regione Liguria Reference System: EPSG:3003 Format: GEOTIFF Size: 1.6 GB	5x5	-	-	-
#40	Liguria-LAS	Dataset from airborne lidar acquisition. Source: Italian Ministry of Environment, Regione Liguria, CNR IMATI Coverage: 65% of Liguria area License: Creative Commons BY-SA-NC v 3.0 (for IQmulus partners) Reference System: WGS84 Size: 1Tb Contents: - raw data from the acquisition process [LAS files] - DSM first and DSM last elaborated from lidar data [.asc files] - XYZ and DTM elaborated from lidar data [.asc and .xyz files] - Envisat Ascending and Descending (PS Interferometry) - ERS Ascendind and Descending (PS Interferometry) - Metadata - Framework of Union [ESRI Shapefile]			2	
#45	Basin Boundary Map	Regione Liguria official boundary of watershed				

Reference data

The main source of data for this scenario is the dataset #40 in the table above. This is a complex dataset, which contains both the original LiDar acquisitions and the DTM elaborated with the techniques described in the as-is-analysis. The DTM belonging to this dataset will be used as reference data to evaluate the goodness of the multi-resolution model constructed in the scenario. Also, the DSM first and last will be used to measure the quality of the representation at varying levels of details of the multi-resolution model.

LS2 – Land Workflow 2: Analysis of precipitation data

Goal: analysis of the observed rain, consisting of the approximation of the observed rain with various techniques, the extraction of the critical points (in particular maxima), and the tracking of the path of the precipitation maxima.

Description and As-Is Analysis

Introduction and main steps of the current solution

Currently, the analysis of precipitation data is done in two settings:

1) statistical analysis of time series (analysis of past events, statistics) or
2) monitoring in "real-time" the evolution of the precipitation.

Typically, the analysis of the precipitation data consists of approximating the precipitation field over the area covered by the measuring instruments using accumulated values of rain in standard time intervals (30min, 1h, 2h, 3h, 6h, 12h, 24h). The approximated field is typically visualized as a 2D map; possibly with a contouring of the accumulated rain values (see Figure 3 for an example).

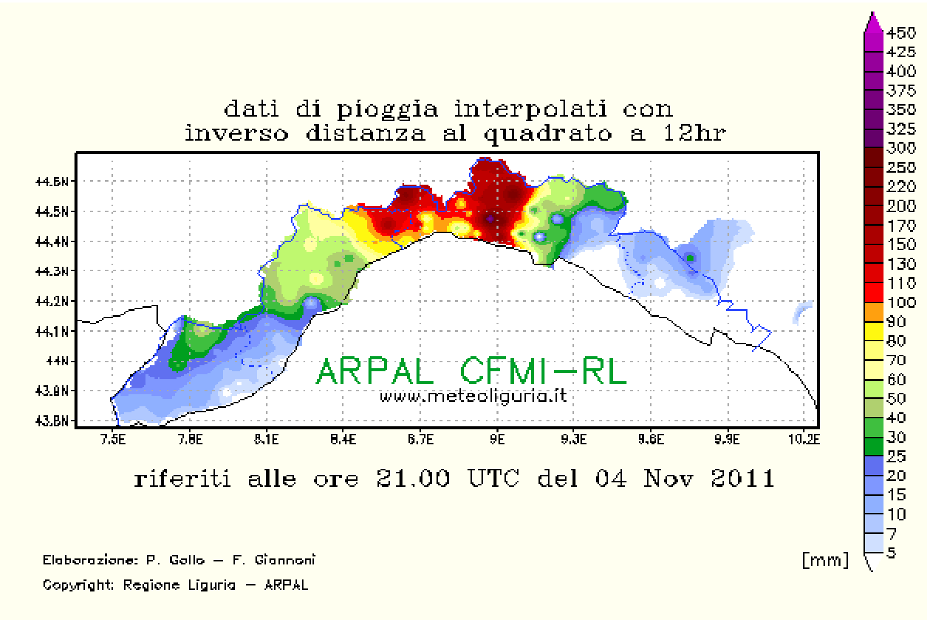


Figure 3. *12-hour cumulated rain map.*

The precipitation field is interpolated using the inverse weighted distance (IWD) schema. The unknown value z, e.g. the rainfall depth, at location x_0 can be estimated as a linear combination of nearby measured values, with the weights being inversely proportional to some power of the distance between observations and x_0:

$$z(x_0) = \sum_{i=1}^{N} \omega_i z(x_i)$$

where the weights ω_i are expressed as functions of distances as follows:

$$\omega_i = \frac{{d_{i0}}^{-p}}{\sum_{i=1}^{N} {d_{i0}}^{-p}}$$

The basic idea for the IDW method is that observations that are close to each other on the ground tend to be more similar than those further apart (Tobler, 1970)[7], hence observations closer to x_0 receive a larger weight. This exact interpolation method requires the choice of the exponent p and of a search radius R or alternatively the minimum number N of points required for the interpolation. Commonly the exponent is equal to 2. Greater values of the exponent p assign greater influence to values closest to the interpolated point, with the result turning into a mosaic of tiles (like a Voronoi diagram) with nearly constant interpolated value for large values of p.

Rain measures available correspond to users/data providers with two slightly different perspectives: experts in charge of monitoring rain up to the regional level and part of the national monitoring network, and experts working for the municipality of Genova. Each of them own a network of rain gauges whose technical specification and spatial distribution matches the needs of their specific role in civil protection:

1) ARPAL professional network, 145 stations, measures every 5-20 minutes, GPRS and radio link connection, data size: 2 MB per day, coverage Regione Liguria (Figure 4)
2) Genova Municipality, 30 stations, measuring every 3 minutes, GPRS or LAN connections, data size: 1Mb per day, coverage Genova municipality. Beside these two direct measures, radar data could be very well used to complement the acquisition of information on rain data.

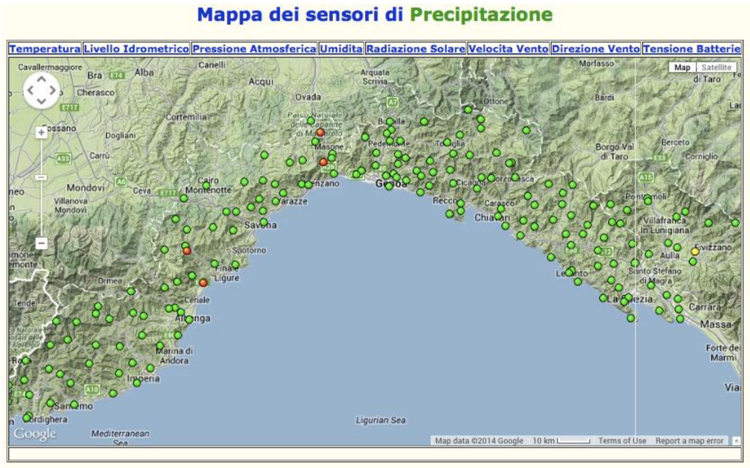


Figure 4. *Distribution of the rain gauges of ARPAL/Regione Liguria*

User perspective and expectations, innovation aspects

Currently, the various data available are not used in an integrated manner, due to technical difficulties and to the lack of methods to fuse the two sources of data while keeping the accuracy distinct, so that the proper reasoning on the approximated precipitation field can be made. In particular, the network owned by the Genova municipality is only semi-professional and therefore the accuracy of the measures is not guaranteed; conversely, the network owned by ARPAL is professional and its installation adheres to the standards of the WMO (World Meteorological Organization). However, the possibility to include in the approximation process also semi-professional data is important and could open the way to integrate in the future also

[7] Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography 46(2): 234 - 240.

user-provided content giving to this type of data the correct interpretation and accuracy weights.

In summary it is felt by the expert users that it is really important to have rain precipitation monitoring systems that provide *evidence of the uncertainty/reliability of the approximated data*, and methods to cope with the *resilience* of the approximation to temporary false rain data or pending communications from measuring stations. Another crucial user request is the *effectiveness of the visualization* beyond standard 2D maps of the results of the precipitation analysis.

Table 2. Big Data issues for the LS2 workflow

Indicator	Relevance	Comment
Volume	High	Time series of acquired data (large volumes of small-sized data) Lidar datasets (for the Regione Liguria, approx. 1 Tera)
Variation	High	Rain data (representative of data streams measured every 5minutes, different sources of data) Lidar, DTM, cartography Image data (optical, radar) Simulation data (hydrological and mechanical)
Velocity	Medium	Rain data are “real time”
Analytics	High	Information about the dynamics of critical rain events

Proposed IQmulus solutions

The novelty we plan to bring with IQmulus concerns:

- testing of approximation methods more suitable to sparse data approximation and carried out taking into account the heterogeneity of data sources;
- coding of the uncertainty/reliability as weights associated to observed measures to reflect their accuracy, and adoption of corrections to cope with failures of stations;
- improvement of the precipitation field analysis by the development of a service for tracking the path of precipitation maxima;
- effective visualization of the precipitation field, including the rendering of the accuracy.

In year 2 of IQmulus, we will concentrate on services to support historical data analysis. We will integrate the real-time monitoring settings in the following project period, but we will work on the historical data analysis taking into account the real-time perspective, meaning that the analysis services should provide an output in a time frame suitable for the real-time monitoring. In year 3, moreover, we will consider how the approximation could take into account the different contextual information and adopt corrective methods to properly model precipitation in urban areas as compared to rural ones.

Workflow components

The workflow components and the corresponding services are listed below.

Preparation component

The year 2 target is concerned with historical data analysis. The preparation phase concerns the selection of the time interval for which the analysis has to be done, and the set-up of all cumulated values over the time interval required. After this data preparation, the surface model for the approximation of the observed rain is called. The output will be a time-varying surface of

the form RainField(P,R(P,t)) where R(P,t) is the value for the approximated rain at point *P=(x,y)* at time *t*. Three alternative methods for approximating the observed rain are considered.

Generic system functionalities required:

Requirement name	*Functionality*
Selection of time interval	The user has to select the temporal interval for the analysis of the time series (necessary for the service LS2_1_1 below)

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Extraction of rain data (Extract rain data for the time interval and prepare the cumulated rain values for the time interval required)	LS2_1_1	New (Service proposed)	Second year PM24
Approximation by Kriging (Approximation of heterogeneous rainfall data (punctual gauge measurements or radar measures) with ordinary Kriging)	LS2_1_2	Original service: #40	Released
Approximation by RBF (Approximation of heterogeneous rainfall data (punctual gauge measurements or radar measures) with radial basis functions)	LS2_1_3	Original service #67	Released
Approximation by LR-splines (Approximation of heterogeneous rainfall data (punctual gauge measurements or radar measures) with LR-splines)	LS2_1_4	Service #58	Second year PM18

WP5 visualization functionalities required:

Requirement name	*Functionality*
Visualize precipitation field over the multi-resolution model	**Data to be visualized:** - precipitation field (output of LS2_1_2 or LS2_1_3, LS2_1_4) represented as a surface above the DTM, not as a simple overlay. Moreover, a suitable rendering method will be used to make clear the the uncertainty of the approximation, - multi-resolution model at a given resolution (as produced by LS1_4_1) **Output:** 3D visualization of approximated rain and the accuracy of the approximation

Feature detection component

Given an approximated rain field, the maxima of rain will be detected and their path over the time interval will be tracked.

Generic system functionalities required: N/A

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Critical Points Extraction of critical points (isolated and degenerate) from grids and triangulations	LS2_2_1	Service #44	Released
Tracking of critical points Given two rain field, with critical points extracted, match the corresponding critical points. It assumes that the two precipitation field are very close in time	LS2_2_2	New (Service to be proposed)	Second year PM24

WP5 visualization functionalities required: N/A

Change detection component: none

Presentation/visualization component

Visualization of the approximated rain field as an animation or other graphical metaphors, showing the evolution of the rain field over time. The rain field could for example be displayed as a triangle mesh (low resolution) over the terrain, with an offset defined at each point P by the rain height/quantity at that location (possibly, scaled to be well visible).

Generic system functionalities required: N/A

WP4 services required: N/A

WP5 visualization functionalities required:

Requirement name	*Functionality*
Visualize the movement of the maxima of precipitation over the multi-resolution model	**Data to be visualized:** - output of LS2_2_2 (path of maxima of the precipitation field) - the underlying approximated rain fields (time steps) - multi-resolution model at a given resolution (as produced by LS1_4_1) **Output:** (animated/interactive) 3D visualization of rain fields and critical points (maxima)

Input and reference data provided for workflow development and testing

Input data

The terrain data are the same as for the LS1 workflow, they are not repeated here. For the analysis of the observed rain data, we will use the set of observations measured during two critical events in Regione Liguria: 20 September 2013 (12 hours of rain observations); and the time slot from 16 to 20 January 2014 (96 hours of observations).

For the sampling of the precipitation fields, we will use triangle meshes extracted from the LS1 multi-resolution model at varying resolutions.

Dataset ID	Dataset title	Dataset short description	Spatial resolution: Resolution distance (m) or scale	Spectral resolution: Number of bands (for RS data)	Spatial resolution: Point density (points/m2)	Data characteristics: Dimensionality
#37	Observed rainfall data	Liguria Region: - ARPAL dataset: historical rainfall data over Liguria Region [time-varying data set, overall size 10Tera, one-day measures account for about 2MB, .txt]	irregularly spaced	na	145 stations over 5410 km^2, irregularly spaced	time series; time resolution: 5 -20 minutes professional rain gauges
#42	Observed rainfall data	Liguria Region: - Genova Municipality dataset: rainfall data over the municipality of Genova [time-varying data set, overlall size 1GB, one-day measures account for about 1,5MB, .txt]	irregularly spaced	na	30 stations over 243 km^2, irregularly spaced	time series; time resolution: 3 minutes semi-professional rain gauges

Reference data

Regione Liguria - ARPAL will provide the maps of the precipitation field computed with their techniques as reference datasets to be used during the testing phase (similar to the dataset #24 in the eRoom data table, but for the critical rain events selected in the showcase).

LS3 - Land Workflow 3: Flood and waterlogging detection

Goal: preprocessing and classification of satellite images for detection of flooded and waterlogged areas

Description and As-Is Analysis

Introduction and main steps of the current solution

The currently used flood detection method developed by FOMI is based on satellite image analysis, and is running in the ERDAS Imagine software environment. A model is built with the software's own model builder to support easy-handling of data and processing, but there are some partly-, or even non-automated steps in the process still, where IQmulus could provide a solution beside other benefits.

The main steps of the current existing solution are listed below:

- Preprocessing of satellite images, including:
  - geometric transformation, reprojection
  - cloud and cloud shadow filtering for producing a cloud mask
  - calibration of the images (Top of the Atmosphere / ToA reflectance calculation) and
  - calculation of spectral indices, namely: NDVI, NDSI, NDWI (NDxI)
- Processing – classification, where several input data are processed, namely:
  - ToA reflectance calibrated image, the previously calculated spectral indices and cloud mask, the mask of natural waters (rivers and lakes), and the mask of not supported areas in Land Parcel Identification System (LPIS).
  - The result is a thematic map with the following 9 categories:
    - natural waters
    - waterlog
    - seriously affected soil
    - moderately affected soil
    - weakly affected soil
    - vegetation in waterlog
    - dry areas
    - clouds
    - not supported areas.

Preprocessing

Cloud and cloud shadow masking is an important step in the pre-processing phase, because cloud shadow and water could be similar spectrally. There is no automated method for filtering, for images from different satellites different aspects and thresholds are applied manually. Because automation of filtering cloud and cloud shadow masking is not part of the development process in IQmulus, we do not go further in detail concerning the methodology here.

For ***ToA reflectance calculation*** a model was built with ERDAS Imagine (Figure 5), but it is not a fully automated method for calibration. For images from different satellites different parameters have to be set manually in the model.

The TOA reflectance represents the solar radiation incident on the instrument in standard unitless terms, independent of the position of the sun with respect to the earth:

$$R_{\lambda} = \pi * D^{2} * L_{\lambda} / (Esun_{\lambda} * COS(\Theta))$$

where:

R_{λ} : calibrated reflectance on the top of the atmosphere referred to the λ wavelength range of the specific band

D: earth-sun distance (in astronomical units, changes within the year)

$L_{\lambda} = DN * A_{\lambda}$: radiance calculated from DN values, for each λ wavelength range of the specific band

DN_{λ} : recorded DN values (digital number) for each λ wavelength range of the specific band

A_{λ}: calibration constant for λ wavelength range of the specific band (from *.hdr)

$Esun_{\lambda}$: the mean exoatmospheric solar irradiance at the specific band. It is defined on the basis of the solar radiation spectrum and the satellite sensor's spectral sensitivity curve

Θ: Sun zenith angle (from *.hdr)

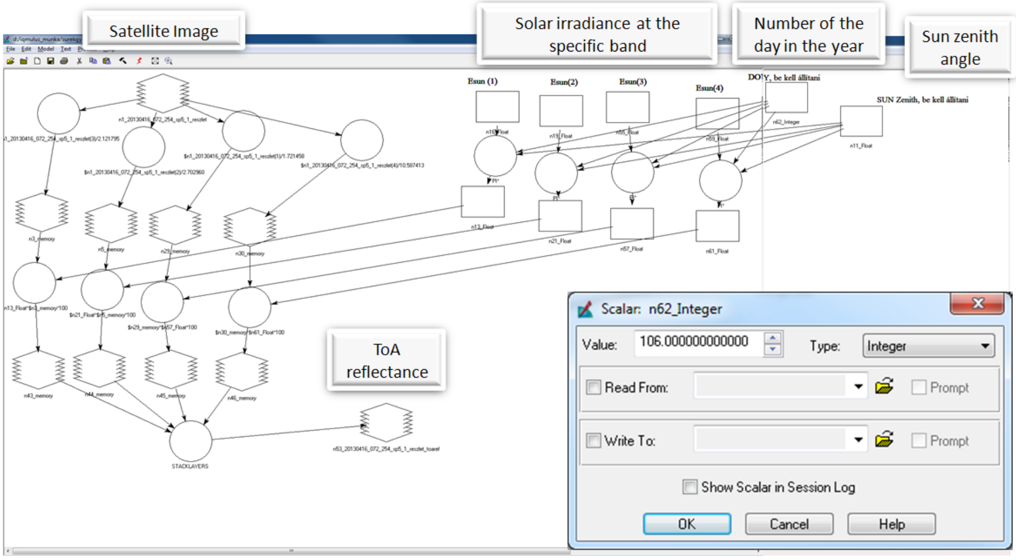


Figure 5. *TOA reflectance calculation model in Erdas Imagine*

The method for the ***calculation of spectral indices*** is currently also a semi-automated process (manual corrections are applied if necessary). The following indices are calculated:

1. $NDVI=(R_{NIR}-R_{RED})/(R_{NIR}+R_{RED})$: normalized difference vegetation index,
2. $NDSI=(R_{SWIR}-R_{NIR})/(R_{SWIR}+R_{NIR})$: soil index (important in special cases like peat)
3. $NDWI=(R_{SWIR}-R_{RED})/(R_{SWIR}+R_{RED})$: water index

where:

NIR: near infrared

RED: red

SWIR: short wave infrared

NDxI = (NDVI, NDSI, NDWI).

If necessary, intercalibration (linear transformation) of NDxI values is carried out. NDxI is defined based on different types of satellite images taken at about the same time for the same area. Finally generation of multitemporal spectral indices by layer stack of NDxI layers is

performed. From the calculated spectral indices for plant growth (NDxI) three spectral index time series are derived (…_ndvi, .._ndsi, …_ndwi) as a result.

Processing

The processing workflow is also set in a model (built in ERDAS Imagine). All the above-mentioned generated data are present as inputs to the model, namely:

- ToA reflectance calibrated image
- spectral indices
- cloud mask
- mask of natural waters and
- mask of areas not supported in LPIS.

Thresholds are set manually in the model for ToA reflectance for each band, and for spectral indices. The model parameters have to be re-set for each image. Figure 6 illustrates the classification model and Figure 7 the results.

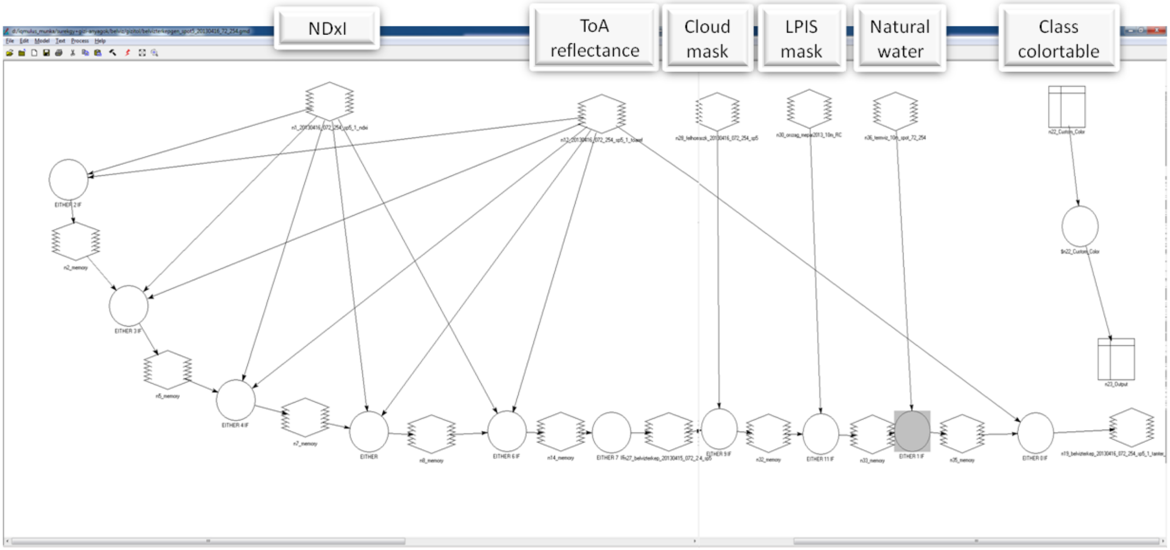


Figure 6. *Graphical model for classification of waterlogging in Erdas Imagine*

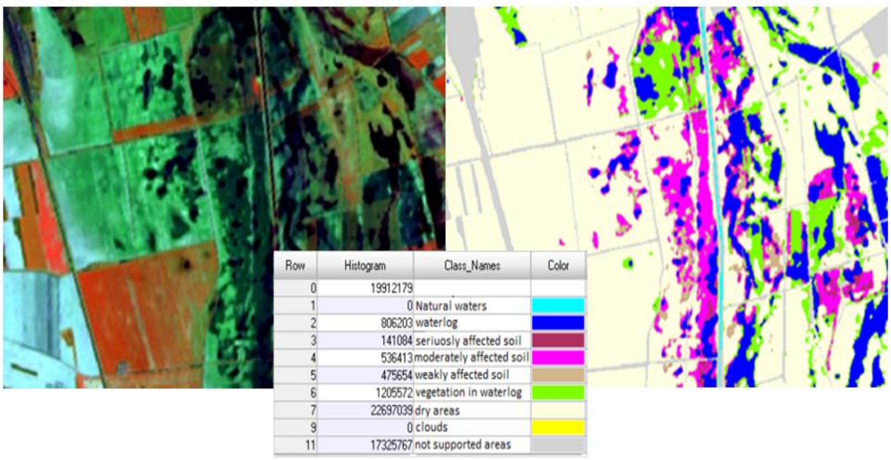


Figure 7. *Result of classification (example)*

Figure 8 illustrates the Use Case Diagram of the current workflow.

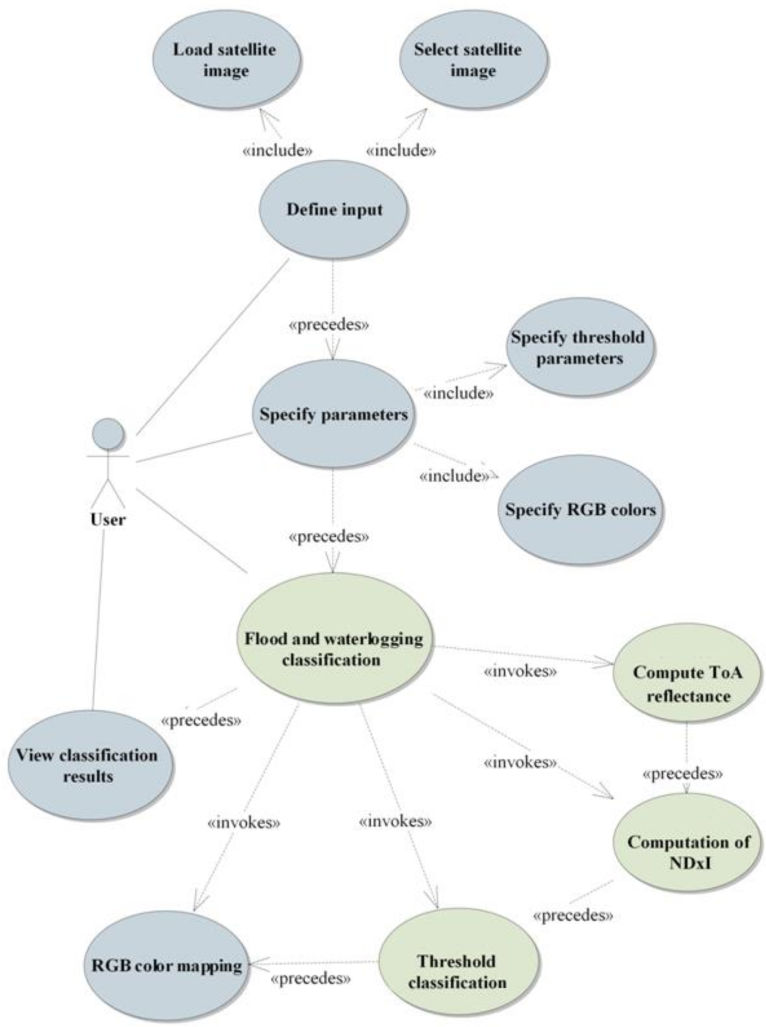


Figure 8. *Use Case Diagram of LS3 (Flood and waterlogging detection workflow)*

User perspective and expectations, innovation aspects

IQmulus could provide several solutions for the automation of pre-processing and classification procedures based on heterogeneous big data.

1. In the pre-processing phase ***automation of cloud and cloud shadow masking*** could be very helpful, especially because such algorithms are available (e.g. fMASK for Landsat 4, 5, 7, and 8 images; SPOTCASM for Spot5 HRG) and could be built in IQmulus as well.
2. ***Automation of ToA reflectance calculation*** is also an available solution, because hierarchical data storage formats with sub-datasets are available, where all parameters necessary for calculation are included (e.g. *.hdf for Landsat images; *.dimap (xml) for Spot images). GDAL which is a translator library for geospatial raster data formats could be a possible choice for *.hdf processing.
3. ***Automation of spectral index calculation*** per image would be very useful, because indices could thus be available for many other tasks as well.
4. In the ***processing/ classification*** phase the automation possibilities for masking and classification are grouped as follows:

i. Threshold settings:
   - manually digitized reference data could be used for threshold settings
   - automated learning from the results of the previous processing could also be a solution.

ii. Classification: using decision tree learning classification methods, e.g. Classification And Regression Tree (CART) approach during the process.

iii. Integrating terrain data (Digital Terrain Model) into the processing chain:
   - currently we can say that there is a lack of accurate data;
   - developing services for the „best interpolation”, the best surface representation, could be a priority, and those services would also be very useful for many other tasks in IQmulus as well.

IQmulus will play a major role increasing the degree of automation of this workflow, but it is important to highlight the relevance of the showcase to the "big data" issues of the project as well.

Detecting flood and waterlogging is usually a task, which requires series of acquired satellite images for different areas, but almost at the same time. Spatial resolution could vary from medium (30 m/pixel) to high resolution (1-5 m/pixel), which affects the data volume. In the case of Hungary, one full non-calibrated Landsat coverage is at least 9,6 Gb. Using time series data and high-resolution satellite images, data volume quickly increases from medium to high. Quick response in a situation like that could be important, thus the processing time of the satellite image series in a given time is a main issue in this workflow. Original data needs calibration; other data has to be derived from the original (in this case spectral indices), and the images could come from different sensors, thus variation of the data in the processing chain is at least medium. Calibration and other pre-processing steps (e.g. cloud filtering) could multiply the original data volume. Assuming that all the results are stored, this data-volume could vary from 870 Gb up to 2 Tb/year for the case of Hungary (with the area of 93 000 km^2), depending on the spatial resolution and the density of the time series. The capacity and velocity of extracting information from the data has to be at least medium due to the quick response expected in a flood situation, and to the simultaneous data processing of all input images for different areas. It means velocity of thematic classification has to be 1-10 minutes per image, but not more. The issues mentioned above are synthesized in Table 3.

Table 3. Big Data issues for the LS3 workflow

Indicator	**Relevance**	**Comment**
Volume	Medium/High	Series of acquired satellite images for different areas at the same time. Different images from different sensors (medium or high resolution) will affect the data volume. Some numbers on the example of Landsat-8 scenes over Hungary: - 1 scene (185 x 185 km, 30/15m resolution, 8 bands) is roughly 1.2 Gb - Hungary is covered by 8 scenes, which is 9.6 Gb - Storing intermediate results (calibration, cloud filtering, resampling, spectral indices) adds an additional volume of ~50 Gb, so the total sums up to about 60 Gb – for a single observation date (coverage) - For multitemporal analysis or monitoring, this data volume has to be multiplied by the number of acquisition dates. Landsat satellites have 16-day orbit cycle (= return time), and the data is most often

		complemented by other kinds of satellite imagery (high-resolution, different sensors, etc.)
Variation	Medium	Original image data (optical, from different satellites); Calibrated data; Calculated spectral indices
Velocity	Medium/high	Quick response in case of flood. Simultaneous data processing for different areas is needed.
Analytics	High	Higher level of automation via smarter algorithms

Proposed IQmulus solutions

Workflow components

The "flood and waterlogging detection" workflow is a complex workflow containing multiple algorithms (some of which are also available as separate services). Compared to the current solution, it provides a higher level of automation via smarter algorithms; therefore, it improves overall processing time and implies a better use of human resources. The user provides reference data instead of the manual setting of thresholds, and derived reference statistics are used for setting the parameters of classification algorithms.

Figure 9 illustrates the Use Case Diagram of the IQmulus solution workflow.

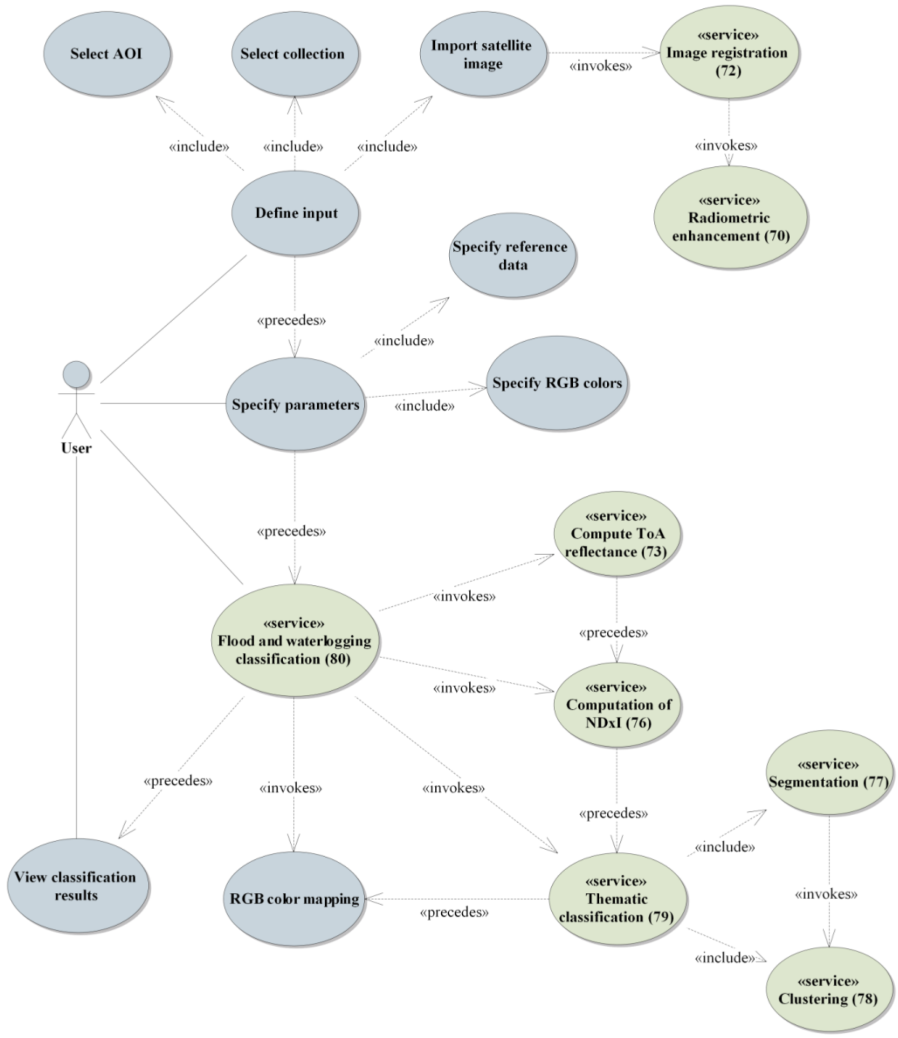


Figure 9. *Use Case Diagram for Flood and Waterlogging Detection, version 2 (PY2, PM24).*

Preparation component

The input satellite image is loaded from the distributed file system. After selection of the Area of Interest (AOI), input multispectral satellite data has to be pre-processed, which is a crucial step before executing any kind of quantitative analysis (spectral index calculation in this case), especially for monitoring or comparing purposes. The Area of Interest is selected by multi polygons by the user, the areas out of the AOI will be masked out of the image.

The Top of the Atmosphere (ToA) reflectance calculation is necessary to complete, using satellite image metadata file. If the input has no ToA reflectance data (this can be determined by inspecting file metadata, or specified using a service parameter), it is computed as the first operation of the service (the required satellite data may be contained in the metadata, or can be specified as parameters of the service). As hierarchical data storage formats with sub-datasets are available with all parameters necessary for calculation, automation of this step can be achieved. Finally spectral indices are calculated as part of the preparation phase of the workflow. The NDxI values for the raster are computed from either the original image or the computed ToA reflectance data.

Generic system functionalities required:

Requirement name	*Functionality*
Import data	The user selects and imports data (satellite image)
Definition of area of interest (Select AOI)	The user has to input an area of interest.
Area extraction	Specified area is extracted from the images (performed by applying a mask to the image, no image transformation is required)

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Computation of Top Of the Atmosphere reflectance	LS3_1_1	Part of the Service #73	Second year PM22
Computation of spectral indices Computation of NDVI, NDSI, NDWI. NDxI=(NDVI, NDSI, NDWI). NDxI is defined based on different types of satellite images taken at about the same time for the same area.	LS3_1_2	Service #76	Second year PM19

WP5 visualization functionalities required: N/A

Feature detection component

Feature detection consists of a segmentation process, which provides the input data for the next step, which is clustering. The resulting cluster map, the image containing the spectral indices, and the reference classification data serve as input for the classifier that produces the classes (water categories) as multipolygons (vector data). Water information is computed using thresholding based on the specified service parameters resulting in the water categories map.

Generic system functionalities required:

Requirement name	*Functionality*
Specify reference data	The user selects and imports reference data (e.g. classified image) for classification

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Raster segmentation Raster segmentation, performed using a predefined or specified algorithm	LS3_2_1	Service #77	Second year PM22
Clustering of raster data Generating clustered raster image based on pre-defined algorithm	LS3_2_2	Service #78	Second year PM22
Thematic classification of raster data Match regions of the images to preselected categories to produce water categories	LS3_2_3	Service #79	Second year PM22

WP5 visualization functionalities required: N/A

Change detection component: none

Presentation/visualization component

If specified by the parameters, the result can be converted to RGB (false colour) representation using the specified category-colour mapping. The output is a GeoTIFF file.

Generic system functionalities required:

Requirement name	*Functionality*
Specify RGB colours	The user selects and imports a pre-prepared RGB colour table

WP4 services required: N/A

WP5 visualization functionalities required

Requirement name	*Functionality*
Visualize classification results	**Data to be visualized**: - initial raster image (GeoTiff) - extracted classification/categories (multipolygons) **Output**: False colour representation of the results using the specified category-colour mapping (RGB colour table is imported previously)

Input and reference data provided for workflow development and testing

Input data

Dataset ID	Dataset title	Dataset short description	Spatial resolution: Resolution distance (m) or scale	Spectral resolution: Number of bands (for RS data)	Spatial resolution: Point density (points/m2)	Data characteristics: Dimensionality
#5	Orthophotos	Multichannel Imagery (RGB+NIR) Ortophotos produced from aerial photographs taken 2000, 2005 2007, 2008, 2009, 2010, 2011, 2012 with native resolution of 0,5m/px.	0.5x0.5	VNIR+RGB		2.5D (2D + height/elevation)
#47	Landsat-5 imagery	Landsat image for flood detection, 2010/156. Raster..	30 m	7 spectral bands (visible, NIR, MIR, TIR: 6 reflective + 1 emissive)		2D

Reference data

Input data for the LS3 (Flood and waterlog detection) workflow testing

Dataset ID	Dataset title	Dataset short description	Spatial resolution: Resolution distance (m) or scale	Spectral resolution: Number of bands (for RS data)	Spatial resolution: Point density (points/m2)	Data characteristics: Dimensionality
#48	Reference map for flood detection	Reference thematic map for flood detection: validated result of operational flood detection, from the current workflow using ERDAS Imagine. To be used for benchmarking.	30	1		2D

LS4 – Land Workflow 4: Detection and characterization of landslides

Description and As-Is Analysis

Introduction and main steps of the current solution

Slow-moving landslides are important geomorphologic processes in many mountain environments, and in areas with potential impact on infrastructure and human settlements continuous monitoring is required. Aerial photographs and especially very-high resolution satellite images are becoming more and more frequently available and can be exploited to measure and monitor the surface deformation. Currently available solutions for displacement measurements depend on specialized commercial software solutions (ENVI+IDL) and are not well adapted for the analysis of large very-high resolution images at full resolution.

The main steps of the currently existing solution are the following:

1) Resample and crop input images to the same resolution and extent.
2) Run image correlation with selected input images and parameters.
3) Filtering of the resulting displacement fields.
4) Generate visualization of the vector field.

The following paragraphs describe each step listed above.

Resample and crop input images

For the application of the correlation algorithm only images with the same extent and aligned pixels are accepted. If the input images do not satisfy this condition resampling and cropping is required which can be realized with most professional GIS and image processing software. Commonly used tools in this context are ArcGIS (Esri) or ENVI (Exelis).

Image correlation

CosiCorr[8] is the IDL library implementing among other tools a sub-pixel image matching technique that operates in the frequency domain. The technique was developed at the California Institute of Technology and the compiled library is available freely for download. The execution of the library requires ENVI (Exelis), which is proprietary software for image processing. The method follows a window-based approach where matching is performed through phase correlation after a Fourier transform with sub-pixel precision. A scheme for masking high frequencies from the normalized cross-spectrum additionally provides better robustness to noise, and a hierarchical sequence with iteratively decreasing window size can be used to measure large displacement. The implemented algorithm can be considered as the current state-of-the-art and is widely used in studies of coseismic slip, glacier monitoring or the investigation of sand dunes[9,10,11]. The algorithm is computationally expensive and measurements are therefore often performed at equally spaced intervals which reduces the computational load but also the spatial resolution of the output. Only two images can be processed at a time.

---

[8] Leprince, S., Barbot, S., Ayoub, F., Avouac, J.P., 2007. Automatic and Precise Orthorectification, Coregistration, and Subpixel Correlation of Satellite Images, Application to Ground Deformation Measurements. Geoscience and Remote Sensing, IEEE Transactions on, 45(6), 1529-1558.

[9] Barisin, I., Leprince, S., Parsons, B., Wright, T., 2009. Surface displacements in the September 2005 Afar rifting event from satellite image matching: Asymmetric uplift and faulting. Geophysical Research Letters, L07301, 6.

[10] Necsoiu, M., Leprince, S., Hooper, D.M., Dinwiddie, C.L., McGinnis, R.N., Walter, G.R., 2009. Monitoring migration rates of an active subarctic dune field using optical imagery. Remote Sensing of Environment, 113(11), 2441-2447.

[11] Scherler, D., Leprince, S., Strecker, M.R., 2008. Glacier-surface velocities in alpine terrain from optical satellite imagery - Accuracy improvement and quality assessment. Ibid., 112(10), 3806-3819.

Filtering of the displacement fields

The derived displacement fields typically comprise an important fraction of noise and false matches in areas with vegetation and low contrast. Consequently, post-filtering is usually necessary and may include manual removal of false matches, masking, or smoothing with non-local means filters. However, this step depends on the specific applications and currently no standard procedure exists. Image-correlation and post-processing are often performed in different software environments hindering stream-line processing of large quantities of images.

Visualization of the vector field

The final outputs are three separate raster containing respectively the x-component, the y-component, and the signal-to-noise ratio of the measurement. Commonly those outputs are imported into Matlab or ArcGIS to generate colour coded representations of the displacement magnitude overlaid with arrow vectors to indicate the direction of the displacement.

User perspective and expectations, innovation aspects

The outlined processing chain comprises several independent steps that can currently not be performed in the same environment hindering batch mode processing of longer time series in one go. In addition the correlation algorithm allows to process only one pair at a time and requires manual intervention for each pair to be processed. The algorithm is not parallelized and therefore not well adapted to process large very-high resolution images at full resolution. As VHR satellite images become available at short intervals (a few days are possible) a streamlined, reproducible, efficient and largely automated processing would be desirable while the accuracy and spatial resolution of the results should not be compromised.

The relevance of the workflow to "big data" issues is synthesized in Table 4.

Table 4. Big Data issues for the LS4 workflow

Indicator	Relevance	Comment
Volume	Medium	Historical archives of aerial images and the increasing fleet of VHR satellites yield important image archives spanning over several decades and typically exceed several GB.
Variation	High	Archives typically comprise images from different systems since there is no long-term continuity of aerial cameras and VHR satellite systems. Since the illumination and the surface appearance can change significantly a high robustness of the matching algorithm is required.
Velocity	Medium	While VHR images are currently typically acquired with a time period of several months, acquisitions at time intervals of a few days are already feasible and will become more common in the next years.
Analytics	High	Displacements may range from several tenths of meters to a few centimetres and hence high robustness and precision of the matching is required.

Proposed IQmulus solutions

The proposed IQmulus workflow targets to combine several components for the quantification of surface displacement from multi-temporal VHR optical images. The employed tools comprise algorithms for harmonization of the input data, measurements of the horizontal displacement field, filtering of outliers and false matches and the visualization of the derived displacement fields.

Workflow components

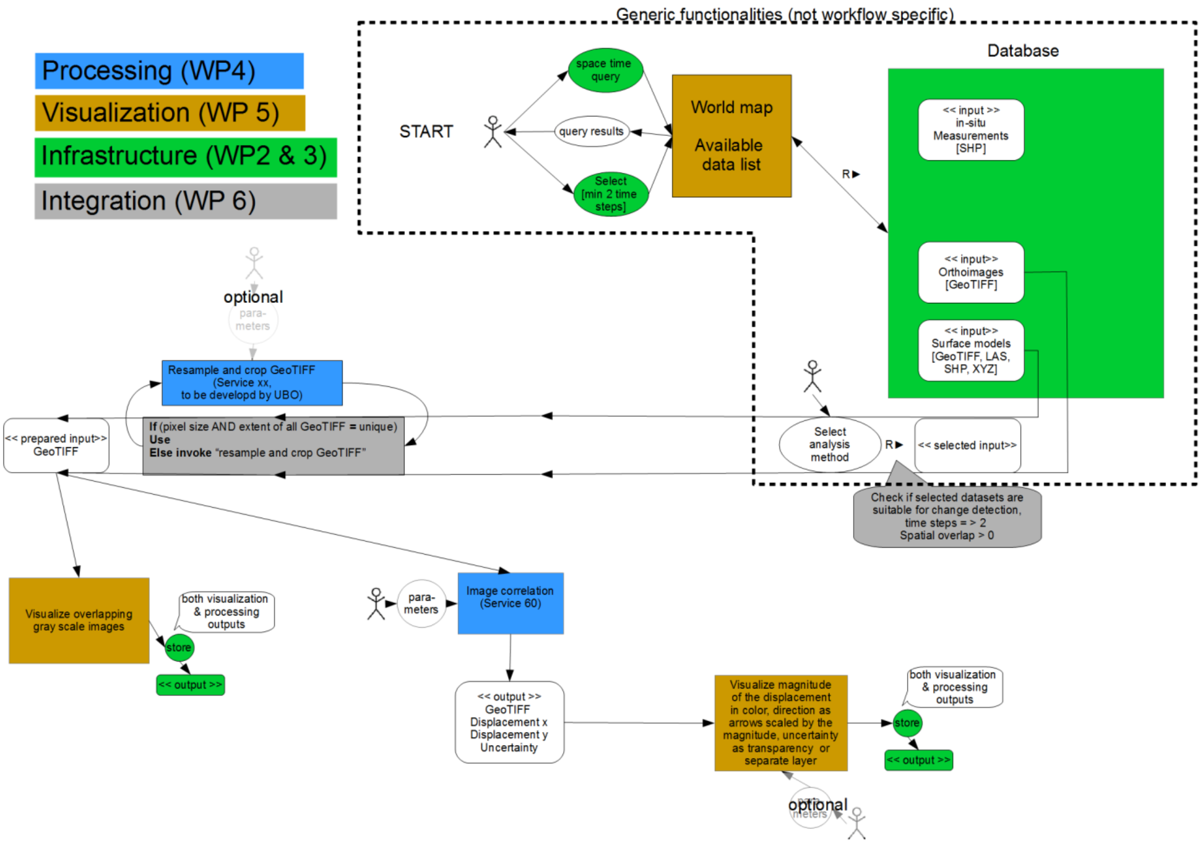


Figure 10. *Overview for the land workflow 4 showing the interactions of the user, infrastructure, processing services and visualization.*

Preparation component

As input data for the workflow two or more single band GeoTIFF images with at least partial overlap are expected. The images should be already orthorectified since parallax shift resulting from the topography is not taken into account. If the images do not have the same spatial resolution a resampling component will be invoked automatically to resample the coarser resolution image to the resolution of the higher resolution image. The algorithm uses a sinc kernel interpolation to reduce aliasing effects and will automatically align the pixel grid of the input images.

Generic system functionalities required:

Requirement name	***Functionality***
Import data	The user selects and imports data (satellite image)
Definition of area of interest (Select AOI)	The user has to input an area of interest.
Area extraction	Specified area is extracted from the images (performed by applying a mask to the image, no image transformation is required)
Storage of resulting raster	Store the resampled raster as GeoTIFF along with the original data. Temporal storage is mandatory and permanent storage will be possible if desired by the user.

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Raster resampling A sinc kernel interpolation is used to harmonize the resolution of the input images	LS4_1_1	New service (to be proposed)	Second year PM 24

WP5 visualization functionalities required: N/A

Feature detection component

For areas with low contrast and significant vegetation coverage image matching will typically yield important errors and such zones therefore require special treatment. VHR satellites typically record panchromatic and multi-spectral data and both can be exploited to locate areas where the image matching is likely to fail. For this purpose the multi-spectral data is analysed using a series of band math, thresholding, and morphological filtering operations to generate a binary mask which will be used as an input in the subsequent change detection step.

Generic system functionalities required: -

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Computation of spectral indices Band math to compute the NDVI	LS4_2_1	Service #76	Second year PM 19
Raster thresholding Generate a binary mask marking all areas where the NDVI is above a certain threshold OR the brightness is below a certain threshold	LS4_2_2	Service #69	Second year PM 19
Convolutional filtering Combination of morphological closing and opening to eliminate isolated pixel	LS4_2_3	Service #71	Second year PM 22

WP5 visualization functionalities required: N/A

Change detection component

Given the input raster, hierarchical cross-correlation with spatial regularization will be invoked to measure the horizontal component of the surface motion at sub-pixel accuracy. At this step several parameters can be adjusted by the user or a predefined parameter set adapted to measurements of landslide deformation from optical images can be invoked. The binary mask generated in the feature extraction step will be used to leave areas unconsidered that typically yield spurious results.

Generic system functionalities required: -

WP4 services required

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Measuring surface displacement Hierarchical cross-correlation is used to find homologous points in two images	LS4_3_1	Service #60	Released PM 18

WP5 visualization functionalities required: N/A

Visualization component

Two different outputs should be visualized. First, the derived displacement components (given as two GeoTIFF raster representing x and y components) should be converted into directed vectors. The vectors representing the direction of displacement should be displayed e.g. as on overlay on a colour coded image representing the magnitude of the displacement. Optionally it should be possible to display the vectors as an overlay on the input data from the earlier time step. Second, the correlation coefficient will be displayed for example with a colour scale reaching from green (strong correlation) to red (weak correlation).

Generic system functionalities required:

Requirement name	*Functionality*
Storage of resulting raster and graphics	Store the resampled raster as GeoTIFF along with the original data. Temporal storage is mandatory and permanent storage will be possible if desired by the user.

WP4 services required: N/A

WP5 visualization functionalities required:

Requirement name	*Functionality*
Visualization of displacement fields	**Data to be visualized:** - derived displacement (GeoTIFF) - derived correlation coefficient (GeoTIFF) **Output:** visualization of the calculated displacement field and the correlation coefficient.

Input and reference data provided for workflow development and testing

Input data

Dataset ID	Dataset title	Dataset short description	Spatial resolution: Resolution distance (m) or scale	Spectral resolution: Number of bands (for RS data)	Spatial resolution: Point density (points/m2)	Data characteristics: Dimensionality
#39	Orthorectified Pléaides Satellite Images	Panchromatic Pléiades images that were orthorectified using surface models extracted from corresponding Pléiades stereo-pairs. Surface movement of active landslides in the valley can be derived from those images.	0.5	panchromatic	N/A	2D + time
		Additional dataset from Regione Liguria will be made available in the next weeks (areas affected by heavy rain which caused landslides)				

Reference data

Dataset ID	Dataset name	Dataset short description	Format
#44	Ground reference measurements for landslide displacement	Continuous measurements at the surface of two landslides	TXT

LS5 - Comparison of simulated floods/landslides with observed data

Description and As-Is Analysis

Introduction and main steps of the current solution

For this last workflow of the showcase, there is no actual As-Is analysis to be referred to. At the Regione Liguria, monitoring of the hydrological status is a crucial activity carried out by various experts and requiring simulation and forecast models customized for the specific geographical area. These processes are extremely complex and require long and detailed tuning. Moreover, the development of flood simulation models, per se, is not within the scope of the project. IQmulus is indeed primarily addressing advancements in terms of efficient handling of large heterogeneous data, and here comes the clear link to the effective analysis of environmental data (observed rain and geo-morphology, as proposed in LS5) finalized to the extraction of parameters needed to initialize flood simulation models.

LS5 therefore concentrates on this aspect of the futuristic pipeline delineated by the showcase. The comparison between simulated and observed events will be also considered, even if we cannot plan now a sound validation scenario as the simulation model used is not tuned to the specific area of interest for the showcase.

User perspective and expectations, innovation aspects

The workflow demonstrates how to support users in the extraction of input parameters to be fed to hydrological models, using observed rain cumulated in a drainage basin to start the simulation of flooding. From a user perspective the impact of this usage scenario is very high; the workflow could be completed with the comparison of the differences between the flooded area resulting from the simulation with the actual flood observed during critical events (for instance, as derived from LS3). Users point to improvements in terms of reducing the computing time but also in terms of adopting advanced 3D visualizations of the flooding event. 3D animation of flooding could have a high positive impact on the communication to the general public of the potential risks that heavy rain could cause.

The relevance of the workflow to the "big data" issues, synthesized in Table 5, groups somehow all the aspects that were discussed for the other workflows in this showcase. LS5 indeed, is an example of how all the intermediate results may be used in the future to support really complex analysis and monitoring systems

Table 5. Big Data issues for the LS5 workflow

Indicator	Relevance	Comment
Volume	High	Time series of acquired data (large volumes of small-sized data) Lidar datasets (for the Regione Liguria, approx. 1 Tera)
Variation	High	Rain data (representative of data streams measured every 5minutes, different sources of data) Lidar, DTM, cartography Image data (optical, radar) Simulation data (hydrological and mechanical)
Velocity	Medium	Rain data are "real time" Simulation modes are computationally very expensive
Analytics	High	Identification of small scale drainage basins Evaluation of simulation models

Proposed IQmulus solutions

For this case, we will concentrate on the computation of the parameters needed for running the GPU implementation of the flooding simulation model developed by SINTEF[12] (Figure 11). The hydrological model is based on the shallow water equations, including friction source terms. The model runs on a regular grid, and needs as input an initial estimate of precipitation, which quantifies the water content over the surface. The precipitation estimate will be computed from the precipitation fields, together with any other value useful to characterize friction locally (e.g., land use, local roughness of the terrain). Since the terrain model is assumed to be stored in a multi-resolution structure, we will compute the input grid needed for the simulation so that it fits the resolution expressed by the LOD requested by the user.

After the simulation is run, it could eventually be possible to compute the difference between simulated data and real events, where the simulation is run starting from real precipitation data during the event.

It must again be underlined that the role of this demonstration is to show users how advanced graphics (3D simulation of flooding, based on scientific modelling of the flooding) could be embedded in IQmulus. Also, the demonstrator will show how efficient processing of large DTMs can quickly provide input for the simulation models that is related to the morphological analysis of the DTM. The specific simulation model that will be used in the demonstrator has not been designed or adjusted specifically for the Regione Liguria territory and therefore the results of the simulation itself are not meant to be correct in a hydrological sense, however, as stated before, the role of the demonstrator at this stage is simply to show how the processing pipeline could contribute in the future to the analysis in this field and improve time-to-simulation and visualization.

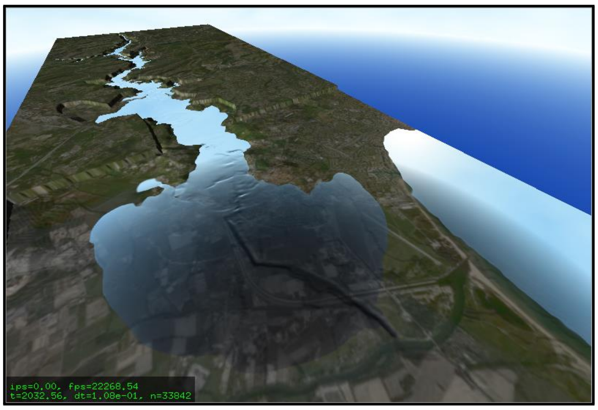


Figure 11. *Example of flood simulation.*

We will not be able to perform the complete loop of operations needed in the scope of IQmulus, still the idea pushed forward by the showcase is to suggest how the efficient computation of environmental data could pave the way towards the evaluation of simulation methods.

[12] A. R. Brodtkorb, M. L. Sætra, and M. Altinakar, **Efficient Shallow Water Simulations on GPUs: Implementation, Visualization, Verification, and Validation**, Computers & Fluids, 55, (2011), pp 1-12

Workflow components

Preparation component

Generic system functionalities required: N/A

WP4 services required

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Compute cumulated precipitation over multi-resolution for hydrology The quantity of cumulated observed precipitation is computed for the time interval and area selected	LS5_1_1	New service (to be proposed)	Third year PM30
Convert multi-resolution for hydrology to regular grid The simulation model runs on a regular grid, the conversion prepares the terrain model so that it can be fed to the simulation	LS5_1_2	New service (to be proposed)	Third year PM30
Convert cumulated precipitation to regular grid The cumulated observed precipitation, computed over the multi-resolution for hydrology, is mapped onto the regular grid, so that it can be given as input to the simulation	LS5_1_3	Service #35	Released

WP5 visualization functionalities required: N/A

Requirement name	*Functionality*
Simulate flood over terrain	The input parameters generated by LS5_1_3 are used to run a simulation using existing simulator developed by SINTEF

Feature extraction component: N/A

Change detection component: N/A

Presentation component

Generic system functionalities required: N/A

WP4 services required: N/A

WP5 visualization functionalities required:

Requirement name	*Functionality*
Visualize flood over the terrain	The simulation results are visualized in 3D in an interactive manner (with animation) for user inspection.

Input and reference data provided for workflow development and testing

Input data

Same as for LS1 and LS2

Reference data

The meaning of this workflow is to demonstrate how advanced and efficient simulation could be plugged into the platform. There is no way to evaluate the quality of the result of the simulation per se.

Workflows of the Marine Showcase

The Marine Showcase includes the following workflows:

1) Marine Workflow 1 (MS1): Marine DEM generation – elevation model generated from point cloud data.
2) Marine Workflow 2 (MS2): Error checking – distance of survey data from a surface.
3) Marine Workflow 3 (MS3): Simple feature extraction –identifying parts of the point data not behaving according to the smoothness hypothesis defined by minimal wavelengths of variation and accuracy thresholds.
4) Marine Workflow 4 (MS4): Change detection – measuring submarine dune migration.

All these workflows assume that the data used is already tiled, so tiling is not part of any of these workflows.

Workflow relations within the showcase

The workflows MS2 and MS3 are built on workflow MS1. Workflow MS3 assumes that both workflows MS1 and MS2 have already been run. Multiple surfaces generated according to the marine workflow MS1 can serve as input for the displacement measurements in MS4.

MS1 – Marine Workflow 1: Elevation model from point cloud data

Goal: to generate a single seamless surface from multiple disparate point cloud source data in a rapid and flexible way, and to visualize it.

Description and As-Is Analysis

Introduction and main steps of the current solution

This 'as is' analysis is based on current practices used within HR Wallingford and Ifremer.

Currently, at HR Wallingford, the process for generating the marine DEM is performed using workflows developed in the COTS software FME (Feature Manipulation Engine) developed by Safe Software (the current workflow for creating a surface is shown in Figure 19). FME provides a library of geospatial data processing tools and an environment in which they can be scripted as workflows. These workflows can call external libraries or user-generated code.

There are four main stages of the process: (1) data load, (2) deconfliction, (3) product generation and (4) quality control. In addition there is an ancillary process that provides supporting input.

Process: Data Load

Individual source data files are loaded into a database. The data load uses the PostGres point cloud storage model. The PDAL library is used to convert XYZ data to LAS format for import into the data base. Data that is not in LAS format (e.g. XYZ) is first converted to LAS, again using the PDAL library. As part of the loading process, data files are split into tiles (or blocks) with each tile holding a similar volume of data. Figure 12 shows an example of the survey coverage loaded into the database. Some areas are only covered by one survey while other areas are covered by numerous surveys.

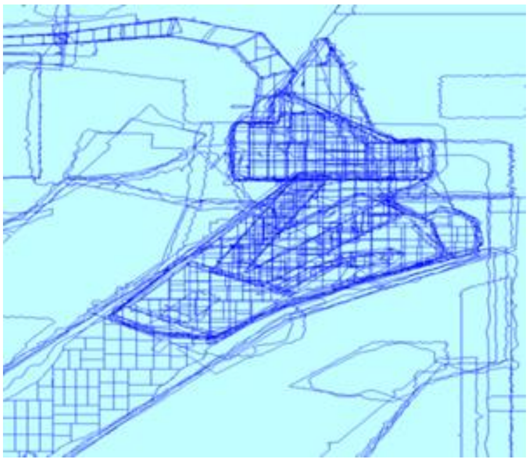


Figure 12. *A section of the sea bottom where some tiles are covered with many surveys while other tiles have just one survey*

Process: Deconfliction

To make a good elevation model, the relationships in time and space between the different overlapping survey tiles have to be understood to properly prioritize and select which survey to use in each area, and to properly combine the surveys in accordance with user needs (for example a user may want to create a surface using only particular data). This process is denoted as "deconfliction".

An example of the source data resulting from a deconfliction process is shown in Figure 13. Each polygon represents a particular data tile selected for creating the DEM and the colored area represents data from the same survey split into multiple tiles. Where no survey data exists, vector data is taken as input from navigation charts to fill the gaps as best as possible.

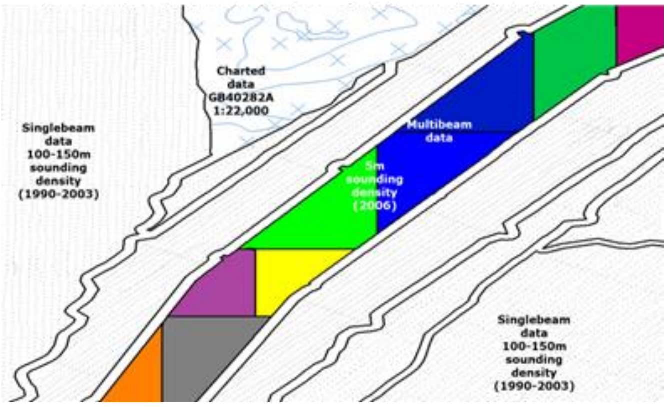


Figure 13. *An example of the result of a deconfliction process, and the data sets selected for each tile.*

Deconfliction is more complex than a simple rule-based selection of appropriate surveys. This is because of edge matching and topological consistency between survey tiles. This is presently overcome by manual intervention. This approach is acceptable where there is only one target product specification, however, interactive deconfliction (as requested by users) is potentially allowing for multiple target specifications.

The current deconfliction algorithms are custom written and contain a rule set that is designed to create the best, most up to date data from which to generate a surface. This uses the metadata about each survey to prioritise which surveys to use. Buffers are also established around deconclicted data extents to reduce edge match problems and data 'slithers'. Once the

deconfliction is completed a manual check is required to further check for edge matching and topological errors. Accordingly the requirement exists for techniques for surface generation that negate the need for this human interaction by being more 'smart' about how the surface is realized.

At Ifremer any overlapping data sets are processed interactively. Statistics are computed on pairs of data sets and the one with the best confidence kept. In order to avoid artifacts from mixing different origins and sensors, the selection of data to be retained for interpolation is performed after analyzing the coherence between data sources. To this end the correlation between the different data sources is calculated, after performing a "migration" of the first file values onto the second file values. The maximum migration distance is defined by the user. The following figure shows examples of obtained correlations.

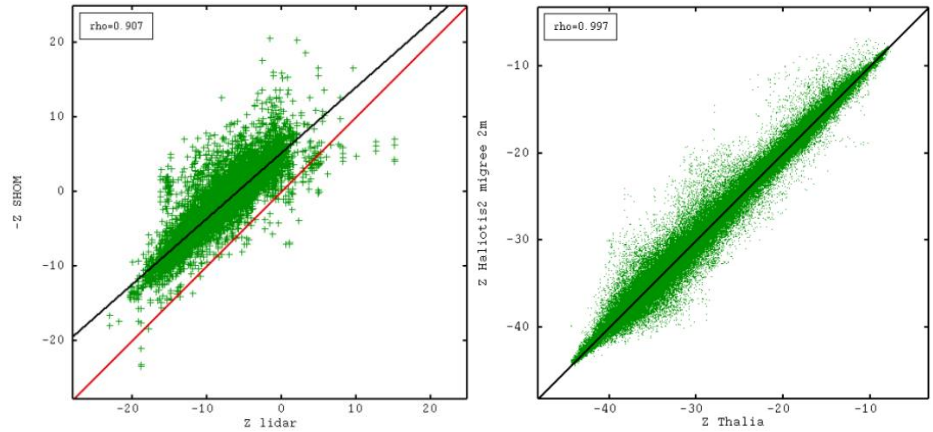


Figure 14. *Left: Example of scatterplot between hydrographic and Lidar soundings showing an overestimation by classical hydrography, Right: MBES and interferometric soundings. The first bisector is in red and the regression line in black.*

According to the result, the user decides how each data source is used for the interpolation: (i) maintain all sources as they are, (ii) correct one source relative to another if the bias is constant (in the case of a malfunction of the tide gauge used for acquisition) or (iii) remove one dataset.

Process: Surface Generation

At HR Wallingford, surface generation currently uses standard triangulaton methods to create a surface from the deconflicted point data set, using tools available within FME for generating surfaces from a point cloud data set (Figure 15). A land mask is also applied to this surface such that land elements are not included. The triangulated surface is then mapped onto a regular grid.

Currently, as mentioned already, the process for generating the marine DEM is performed using the COTS software FME (Feature Manipulaton Engine) developed by Safe Software. FME provides a library of geospatial data processing tools and an environment in which they can be scripted as workflows.

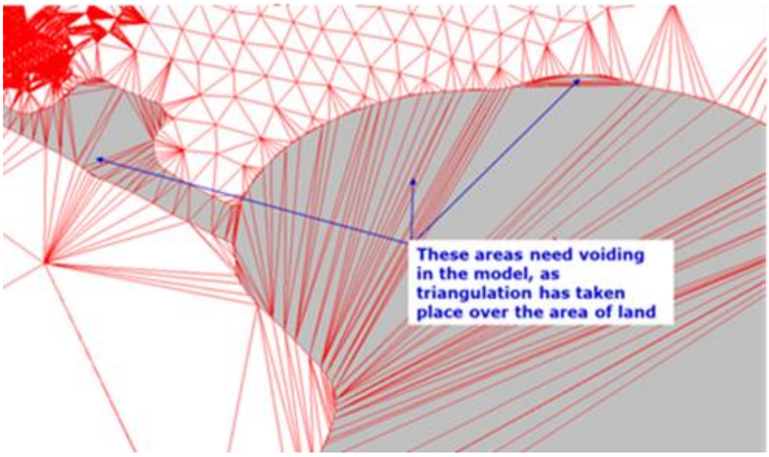


Figure 15. *Triangulation method used to create surface from data points*

These workflows can call external libraries or user-generated code. Each of the three steps in the process have a manual checkpoint on completion before the next step.

For large areas, the time taken to create a surface is far from instant. For example creating a surface for the UK takes several minutes.

At Ifremer, the interpolation is carried out by kriging. An analysis of the variogram is performed to apply the mathematical model that best fits the data while taking anisotropy into account. The neighborhood can be finely tuned by selecting distance, sectors and number of points required. Masks or "barriers" can be set to avoid interpolating over land or across known faults or objects that are not part of the regional structure.

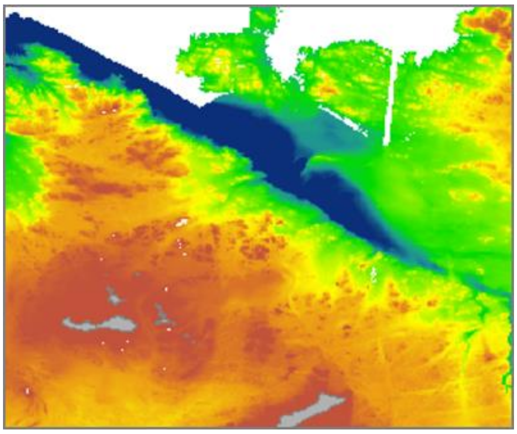


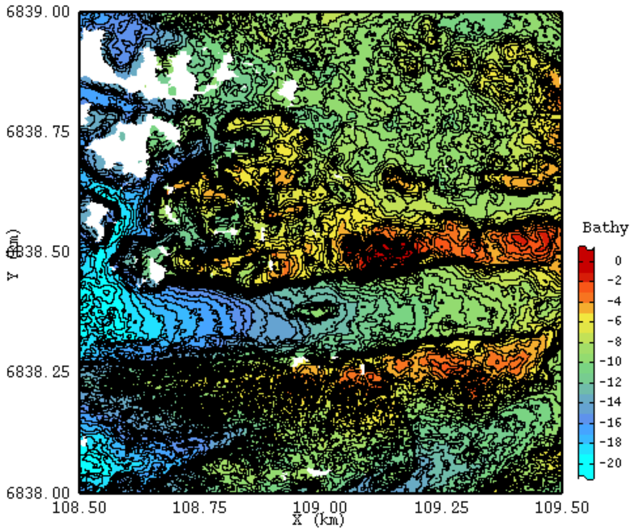
Figure 16. *Triangulation method used to create a surface from data points*

With the tools used by Ifremer, a selection of one Lidar point every 50cm is performed in order to reduce computation time. With IQmulus's ability to handle large data sets we expect to simplify this process by removing this preselection step.

Process: Quality control

The quality of the resulting model is assessed by:

- 3D visualization by zooming in on small geographic areas,
- Graphical visualization of the results in the form of contours (Figure 17&18), the smoothness and stability of which help identify potential interpolation artefacts.



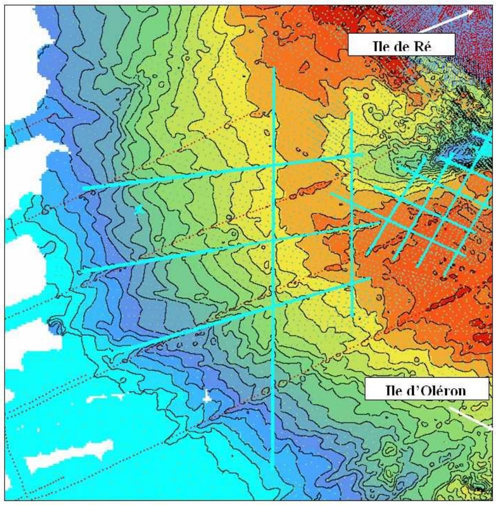


Figure 17. *Figure. An Example of isolines display (every 50cm) calculated from obtained DEM.*

Figure 18. *Artifacts resulting from the coincidence of old (red dots) and recent (blue dots) data sets.*

About visualization aspects

Visualization is not an inherent part of the current DEM process. GIS is used to visualize intermediate steps in the production process, but this is an offline activity. Likewise once the surface is generated tools such as Fledermaus are used to view the surface in 2D or 3D.

Figure 19 illustrates the current process for deconfliction and surface creation at HR Wallingford:

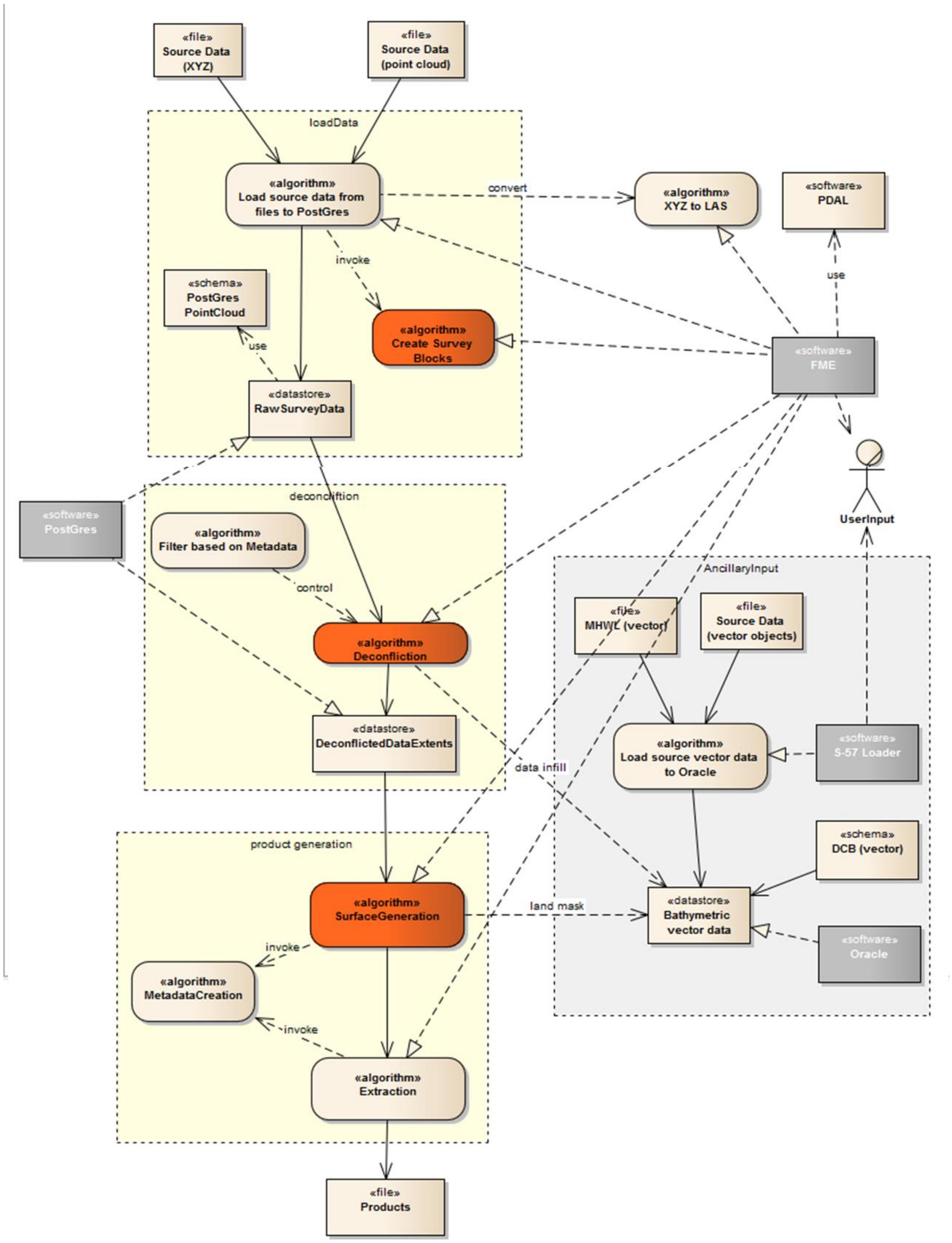


Figure 19. *The current process for deconfliction and surface creation at HRW*

User perspective and expectations, innovation aspects

Within this showcase IQmulus is demonstrating two key aspects of functionality that can be used in the marine test beds. The first is a reduction in the time it takes to generate a surface so surfaces can be generated (and subsequently visualized) 'on demand'. The second is an

improved method for deconfliction that negates the need for human interaction. It should be remembered that deconfliction is fundamentally about selecting data to be used to generate a surface. How this is actually implemented is not critical to the user, as long as the resultant DEM is 'correct'. The current method is just one solution to selecting and merging multi-source data.

The relevance of the workflow to "big data" issues is synthesized in Table 6.

Table 6. Big Data issues for the MS1 workflow

Indicator	Relevance	Comment
Volume	High	In many cases surfaces need to be generated over large areas which contain 10^3 datasets with each data set being 10^4 MB.
Variation	Medium	Data from different sensors, process models and charts, quality varies as well as the number of surveys in different regions.
Velocity	Medium	The temporal frequency of bathymetric surveys is currently rather low and real-time processing at sea is typical not considered as an option. However, project specific surface generation on demand should be possible and renders the processing speed a critical issue.
Analytics	High	Need for efficient tools to extract information from data for visualization and other uses

Proposed IQmulus solutions

As LR-splines will reduce the data representation and thus provide a potential of much faster visualization, interactive deconfliction will be better supported by LR-splines than the current raster based representations. So the strategy is to exploit this as faster access and visualization opens new possibilities.

We can proceed with novel approaches for interactive deconfliction when we have demonstrated and discussed the results and the potential of the three planned LR-spline workflows: MS1, MS2, MS3. Seeing is believing, and currently when we discuss the potential of LR-splines for interactive deconfliction, LR-splines is just "promiseware", not demonstrated software.

Discussion of LR-spline representation

While (piecewise polynomial) splines are the numerical tools of choice for surface representations in wide application areas, one drawback has been the lack of local refinement that traditional tensor product B-splines have. The novel approach of LR-splines [Dokken, Lyche, Pettersen 2013[13]]overcomes this problem. From the theoretical aspects of approximation of smooth functions it is known that:

1) Degree 1 polynomial/spline approximation is $O(h^2)$. In the 1-variate case this would be straight lines, in the 2-variate case this would be triangulations.
2) Degree n polynomial/spline approximation is $O(h^{n+1})$. In the 1-variate case this would be a degree n spline curve, in the 2-variate case structures of triangular Bezier surfaces of degree n, tensor product B-splines or LR B-spline surfaces of bi-degree (n,n).

[13] Dokken,T.; Lyche, T.; & Pettersen, K. F. (2013). Polynomial splines over locally refined box-partitions. *Computer Aided Geometric Design*. ISSN 0167-8396. 30(3), s 331- 356 . doi: 10.1016/j.cagd.2012.12.005

a) The Bezier triangles have the same topological complexity at triangulations and many constraints have to be imposed to ensure proper continuity.

b) Tensor product B-splines are not locally refinable and will not give a reduction of the data volume corresponding to the $O(h^{n+1})$ approximation order.

c) LR-splines are locally refinable and can thus achieve a reduction of the data volume corresponding to the $O(h^{n+1})$ approximation order.

Consequently LR-splines have the approximation power and structure necessary; the tests will show what we can achieve, especially with a hybrid approach. However, we currently do not have good estimates of the computation times used for such approximations as this is ongoing research and the current version of LR-splines is in debug mode. The final version can be parallelized (MapReduce). The computation time will to a great extent depend on how much refinement is necessary. The thresholds and minimal element size can be used to avoid going too deep, and to split the data set into points approximated with the required tolerance (the smooth areas) and points outside the tolerance (areas with possible features). We should consider including a track in the year 3 processing contest that allows comparison of LR-splines with triangulations with respect to accuracy, computation time and data volume for the representation.

Experiments at the start of IQmulus have shown the feasibility of very compact LR-spline representations for the smooth sea bottom. In addition, such a spline representation is potentially also a tool to be used for feature detection:

1) Features that are non-smooth will not be accurately represented by splines. Such features are located in areas where the approximation has the largest error. This might, for example, be rocky outcrops, rocks, human-made structures or wrecks. Consequently making a spline approximation with a not too fine tolerance will help to identify such features, as the features will be excluded. When the points are compared to the surface the points representing features will clearly stand out.

2) Features that are smooth, such as moving sandbars, drag marks from glaciers on the sea bottom, or craters from gas emerging from the bottom will have a variation pattern different from the waviness of the global sea bottom. Consequently producing a spline approximation of the sea bottom not allowing detailed waviness will help to identify such features.

In the implementation of the Marine Showcase, the idea is to address first simple workflows that illustrate the potential of the hybrid representation using a combination of point sets, triangulations and splines. Thus in the second project year, the focus is to implement sufficient functionality for LR-splines to demonstrate the spline approach, then in the following year to focus more on the hybrid representation including (simple and hierarchical) triangulations.

As stated before, the ambition is to perform deconfliction and change detection in multi-temporal sea bottom data sets.

First, services for representing the information of the point cloud in the new and compact hybrid format have to be implemented. Consequently in year 2 of IQmulus the focus will be on simple workflows using (from WP4) services implemented for shape representation and (from WP5) first functionalities for visualizing the LR-spline surfaces, also comparing the LR-spline surfaces with point clouds and visualizing the results of those comparisons. More advanced workflows for deconfliction and change detection will be introduced in year 3 of IQmulus.

Workflow components

The components of the workflow and the corresponding services are described below, and illustrated with Figure 19.

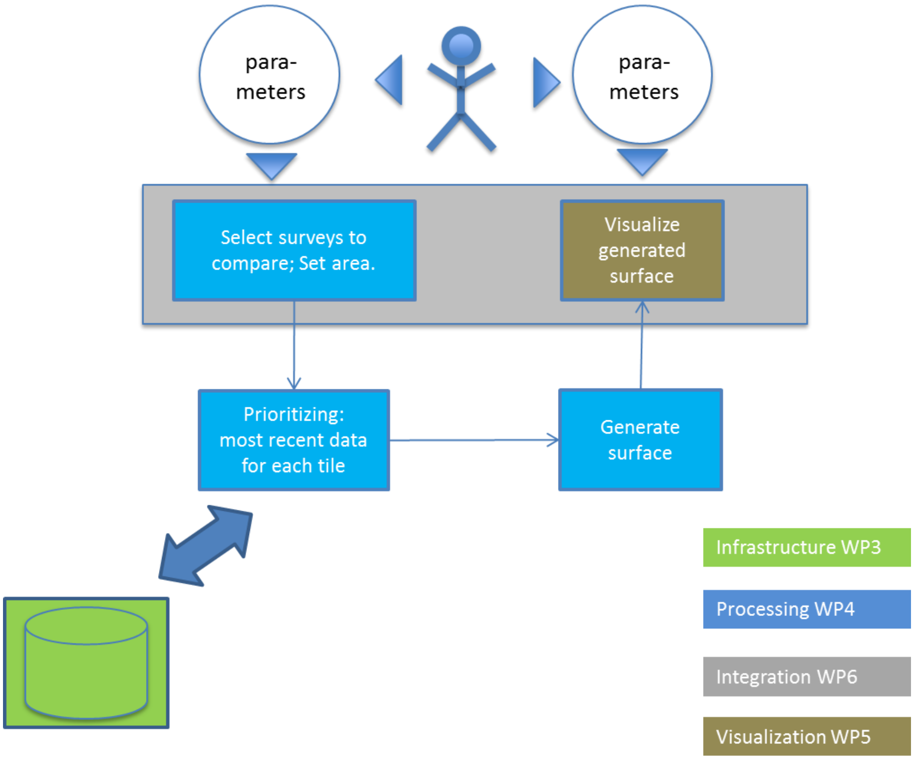


Figure 20. *Workflow components for MS1*

Preparation component

The workflow assumes that the deconfliction - prioritizing for tiles where multiple surveys exist - is provided. A first very simple support service for this component, selecting the most recent survey, is to be implemented for this purpose, to be replaced by more advanced interactive deconfliction in year 3.

The workflow is focused on the generation of the LR-spline surface and its visualization. It should be considered to allow accessing data from adjacent tiles/surveys to control the boundary behaviour (this can possibly be done by the user selecting extra surveys just outside the boundary):

- To create an initial surface from many surveys;
- To create a surface just based on one survey, but using adjacent surveys to help control boundary behaviour.

When creating the surface the uncertainty of the point clouds of the tiles used can be employed to direct the approximation tolerance used over the different tiles. The tolerances and uncertainties of the used raw data, the approximation error and possibly also the max error, as well as the standard deviation of the approximation of the surface across each tile is metadata that possibly should be attached to the LR-spline surface. The surface generation can be controlled by different parameters to be targeted at specific uses in more advanced workflows:

- A surface tolerance;
- The minimal polynomial cell width.

For feature detection, e.g., sand banks that often have a wave like pattern, it will simplify feature detection if a surface not able to reproduce the waviness at a certain scale is used. This can be done by restricting the minimal polynomial cell width of the spline. When later comparing the data set with the surface, the waves will stand out as a wavy pattern/periodic behaviour with regions above or below the surface. In a similar way boulders or rocky outcrops will stand out.

Generic system functionalities required:

Requirement name	*Functionality*
Import data	The user selects and imports data.
Definition of area of interest (Select AOI)	The user has to input an area of interest.

WP4 services required

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Point cloud to LR-spline surface (Generates an LR-Spline surface from a point cloud)	MS1_1_1	Service #9	Released Some update needed

WP5 visualization functionalities required: N/A

Feature detection component: none

Change detection component: none

Presentation/Visualization component

Generic system functionalities required: N/A

WP4 services required: N/A

WP5 visualization functionalities required:

For the presentation, a spline surface visualization can be used, possibly also adding contour lines.

Requirement name	*Functionality*
Visualization of an LR-spline surface	**Data to be visualized:** - Generated LR-splines representing the surface **Output:** 3D visualization of the LR-spline surface on a fat client

Input and reference data provided for workflow development and testing

HR Wallingford will provide a database of raw survey data for use and testing in the project. The trial area will cover the south coast of England and contains 10^3 surveys. This is shown in Figure 21.

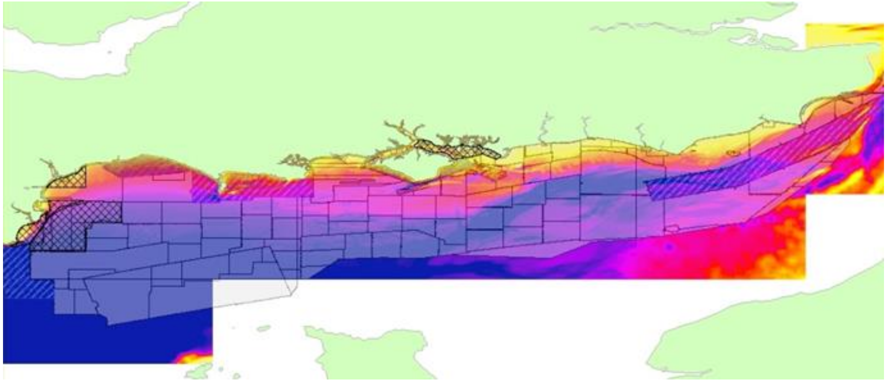


Figure 21. *Trial area for Marine workflow MS1*

The surveys are loaded as tiles in a point cloud data base. The data is commercially licensed and a copy of the licence has been forwarded to SINTEF for use on the project. A subset of this data (10^2 surveys) can also be provided under an open data licence.

Dataset ID	Dataset title	Dataset short description	Spatial resolution: Resolution distance (m) or scale	Spectral resolution: Number of bands (for RS data)	Spatial resolution: Point density (points/m2)	Data characteristics Dimensionality
To be uploaded	Bathymetry point cloud of south-east England offshore area	Commercially licensed, license can be provided to partners for project use	Multiple point clouds typically 1m-100m resolution depending on acquisition type	NA	10^9 points over 20^3 km2 (50^3/km2)	Multiple (10^3) 3D point clouds with different point density. Overlapping in time and space.

MS2 - Marine Workflow 2: Error checking - distance of survey data from a surface

Goal: comparison of a new survey point cloud with existing LR-spline approximation.

Description and As-Is Analysis

This workflow is a new workflow introduced for the LR-spline surface representation of the DEM. No as-is workflow exists.

Proposed IQmulus solutions

Workflow components

In the marine scenario in the second year surveys will have two representations:

- The point cloud
- The LR-spline approximation.

Comparing a new survey with a previous survey can thus be performed in two ways:

- Compare the existing LR-spline approximation with the point cloud of the new survey.
- Approximate the new survey with an LR-spline and compare the old and new LR-spline surfaces

We have chosen here, as shown in Figure 22, to address the first alternative as this only involves the approximation error of the LR-spline approximation to one point cloud. However, also the second alternative should be considered for implementation but this will require more discussions on how to handle approximation tolerances before it can be implemented. So this is not prioritized for year 2.

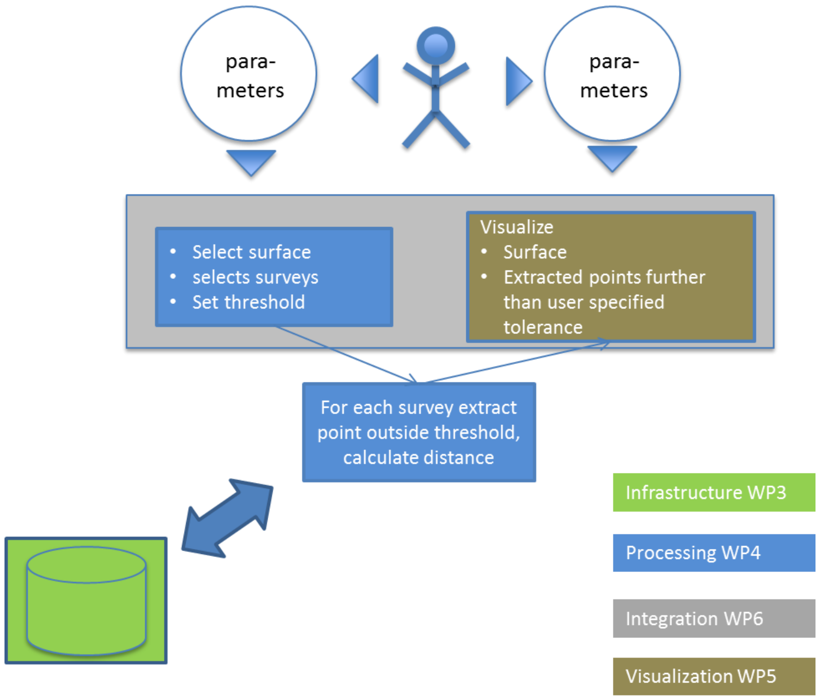


Figure 22. *Workflow for comparing an LR-spline surface and a point cloud*

The relevance of the workflow to "big data" issues is synthesized in Table 7.

Table 7. Big Data issues for MS2 workflow

Indicator	Relevance	Comment
Volume	High	Surfaces typically are generated over large areas and will be compared with large point clouds
Variation	Medium	Data from different sensors, process models and charts, quality varies and the number of surveys in different regions.
Velocity	Medium	Large velocity during surveys, however, most often months or years between surveys
Analytics	High	Need for efficient tools to extract information from data for visualization and other uses

Preparation component

The workflow assumes that an LR-spline surface representation exists and that the data sets to be compared are available for use.

Generic system functionalities required:

Requirement name	*Functionality*
Import data	The user selects and imports an LR-spline surface and the surveys to compare to
Definition of area of interest (Select AOI)	The user has to input an area of interest.

WP4 services required: N/A

WP5 visualization functionalities required: N/A

Feature detection component

This workflow can be used for comparing one or more, possibly overlapping surveys with an LR-spline surface. The idea is to produce a set of points that are outside a user specified threshold. If the threshold is set to the approximation tolerance used the points from the survey are split into two sets:

- Points within the threshold are represented by the LR-spline surface.
- Points outside the threshold. These points show where the approximation is not within the tolerances. This set of points will possibly contain the features, and will hopefully be much smaller than the original dataset of the survey. Thus we have generated from the given big data small data that can be more easily handled by the visualization.

We assume that the surface is smooth. Thus points outside the threshold can represent different aspects of the data, such as features (stones, wrecks, pipelines, etc.), changes in elevation between survey and moving sand dunes.

If the threshold is set too low and too many points are selected to achieve proper graphic performance then the user will be asked to address a smaller part of the surface or increase the threshold such that the number of points produced is smaller.

Generic system functionalities required: N/A

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Compare point cloud to LR-spline surface (Computes the distance between a point cloud and an LR-spline surface)	MS2_2_1	Service #86	Second year PM24

WP5 visualization functionalities required: N/A

Change detection component: N/A

Presentation/Visualization component:

By allowing the user to select the distance of points to be visualized (e.g., by slider bars) and 3D navigation, the user can visually inspect the structure of points outside the tolerance to better understand the surveys.

Generic system functionalities required: N/A

WP4 services required: N/A

WP5 visualization functionalities required:

Requirement name	*Functionality*
Outlier visualization	**Data to be visualized:** - LR-spline surface (from MS1_1_1) - original point cloud with computed distance to LR-spline surface **Output:** interactive 3D visualization of the points from the point cloud lying outside a user-given threshold of the LR-spline surface (from MS2_2_1).

Input and reference data provided for workflow development and testing

Same data as for MS1

M3 - Marine Workflow 3: Simple feature extraction

Description and As-Is Analysis

Introduction and main steps of the current solution

The current process consists of 3 processes, which are the following:

Process: Data Load

Existing DEM (Grid ascii or Geotiff) source data file is loaded into a database.

Process: Geomorphologic parameter derivation

It consists in the creation of several DEM-derived rasters which allow significantly improving data visualization and interpretation, and facilitating rocky areas identification:

- hillshade is obtained by taking into account a defined angle of a light source;
- roughness is considered as a descriptor of topographic complexity;
- slope of a relief, angle of inclination relative to a planar surface at sea level, and its orientation.

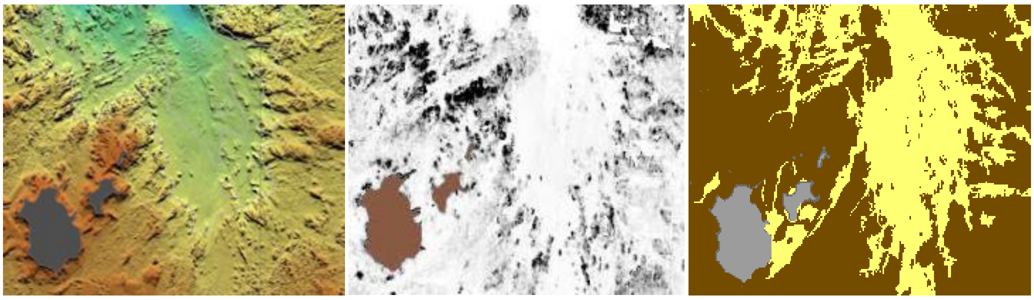


Figure 23. *Examples of DEM-derived geomorphologic parameters (Hill shade, roughness) and its interpretation (brown : rock; yellow : sediment)*

Process: Rocky area delineation

Bedrock is defined on the basis of manual interpretation of geomorphological parameters calculated from the digital terrain models. The result is a layer that distinguishes rocky areas from soft substrate. This interpretation is performed using ArcGis tools.

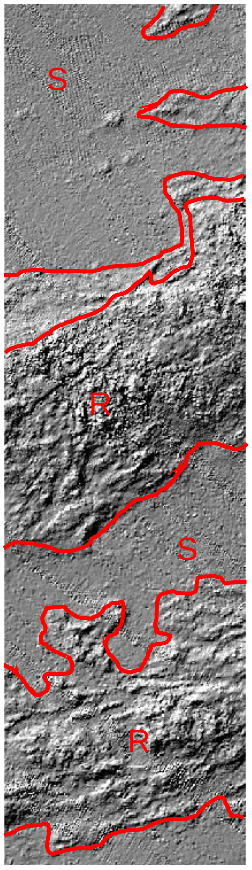


Figure 24. *Manual delineation of rocky areas*

User perspective and expectations, innovation aspects

The year 2 marine showcase is going to implement the possibility to generate accurate surface representations from heterogeneous measurements, which is an important pre-requisite to subsequently detect features of interest and detect changes among multiple time steps. In year three it will therefore be possible to further extend the functionalities of the prototype allowing further analysis of multi-date surface models. In particular this will allow the user to access services for the detection of significant changes through stochastic and topological change detection as well as the quantification of sediment fluxes using digital image correlation. Figure 25 provides an overview of the envisaged components for processing, visualization, infrastructure and integration, their links and the channels for user interaction.

The relevance of the workflow to "big data" issues is synthesized in Table 8.

Table 8. Big Data issues for MS3 workflow

Indicator	Relevance	Comment
Volume	High	The data from sea bottom surveys is most often big as large areas are covered in most such surveys. In some cases surfaces need to be analysed over large areas comprising data sets being 10^4 MB.
Variation	Medium	Data from different sensors, process models and charts, quality varies as well as the number of surveys in different regions.

Velocity	Medium	The temporal frequency of bathymetric surveys is currently rather low and real-time processing at sea is typically not considered as an option. However, project specific surface analysis on demand should be possible and renders the processing speed a critical issue.
Analytics	High	Efficient tools are needed that identify automatically non-smooth areas and help direct users to these areas that may need their attention.

Proposed IQmulus solutions

This workflow, shown in Figure 25, will in its first implementation have very simple feature extraction mechanisms. Most of the components for this first simple feature extraction are already available from the work on LR B-splines. For year 3 more advanced feature extraction should be realized.

Workflow components

The workflow will address clusters of points that are above or below the surface (rough surface approximation is probably best) and cluster these into groups. Geometrically these groups can possibly be classified according to the shape of their outline. The procedure is illustrated in Figure 25.

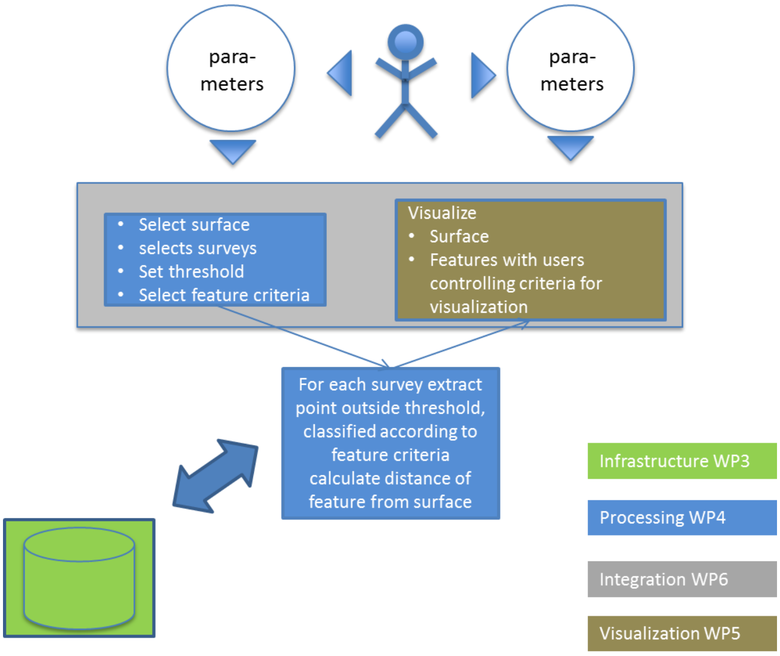


Figure 25. *Simple feature extraction.*

We should possibly also have a minimum number of points within a feature and a minimal feature extent so that the part of the local area covered by the feature is not too small. These properties can also be calculated for the user to control which type of features are to be visualized.

Examples of feature criteria in the first simple version are:

- Feature with a long and narrow outline, possibly with min/max parameters for extent. Could possibly identify drag marks on the bottom, e.g., for icebergs, pipelines, ...
- Features with an outline with approximately the same extent in most directions, possibly with min/max parameters for extent. Could possibly identify wrecks, boulders.

Based on discussions we will probably find a number of more feature criteria, e.g., related to the waviness of a sequence of adjacent features (sand dunes), also here with possible geometric min/max parameters.

Preparation component

Generic system functionalities required: N/A

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Select feature criteria (Choose the criteria to identify a feature or a set of features)	MS3_1_1	New service to be proposed	Second year PM24

WP5 visualization functionalities required: N/A

Feature detection component

Generic system functionalities required:

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Detection of features by given criteria (Detection of features in the LR-spline surface)	MS3_2_1	New service to be proposed	Second year PM24

WP5 visualization functionalities required: N/A

Change detection component: N/A

Presentation/Visualization component

Generic system functionalities required:

WP4 services required: N/A

WP5 visualization functionalities required:

Requirement name	*Functionality*
Feature visualization	**Data to be visualized:** - LR-spline surface (from MS1_1_1) - detected features (from MS3_2_1) **Output:** Visualization of the features detected in an LR-spline surface according to given feature criteria (from MS3_2_1).

Input and reference data provided for workflow development and testing

Same input data as for MS1

MS4 – Marine Workflow 4: Measuring submarine dune migration

Description and As-Is Analysis

Introduction and main steps of the current solution

Submarine dunes and sand banks are among the most dynamic geomorphologic formations in coastal waters. They play an important role in the sediment transfer and their movement can have significant impacts on the benthic ecosystem, marine transport in coastal waters, and infrastructure such as pipelines and communication cables. While in the past it has been difficult to observe the dune dynamics directly, recent advances in multi-beam echo sounding (MBES) now enable increasingly frequent observations of the sea floor morphology. Currently, the prevailing standard for the analysis of multi-temporal MBES surveys is visual analysis by trained experts delineating the crest line of the dunes at several time steps using commercial or open-source GIS software.

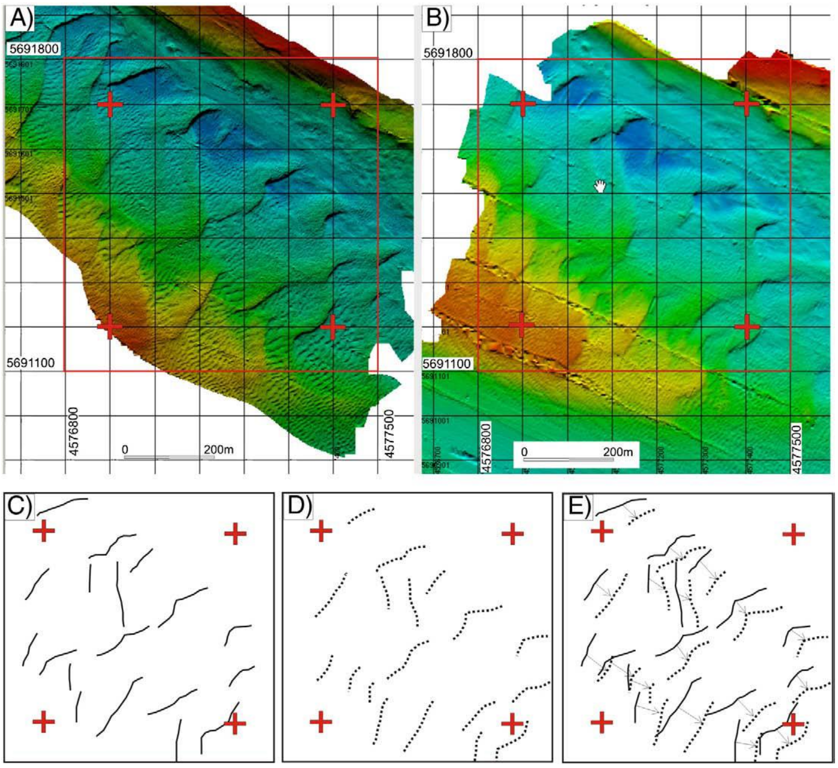


Figure 26. *Illustration of a manual analysis of two bathymetric surveys (Gómez et al., 2010) recorded in September 2007 (A) and May 2008 (B). The crestlines of the main dunes are delineated for both time steps (C, D) and their center points are matched manually to derive the displacement vectors (E).*

The main steps of the currently existing solution are:

1) Raster surface interpolation from MBES point clouds for two or more time steps;
2) Derivation of terrain parameters such as hillshade maps, aspect and slope;
3) Visualization of the derived terrain parameters in a standard GIS software and manual tracing of the dune crest lines with polyline features;
4) Manual matching of the crest lines delineated on two or more time steps and calculation of a mean displacement vector for each polyline.

Raster surface interpolation:

To derive a continuous surface from raw bathymetric data typical comprises two steps. In the first step the raw data is combined with navigational data, corrections of the sound velocity and tidal corrections. Furthermore, outliers are removed manually or based on statistical methods. A 3D point cloud is derived which is interpolated to a continuous raster in a subsequent second step. The most commonly used interpolation methods comprise simple binning, weighted moving averages, triangulation, or splines with tensions. CARIS HIPS, Fledermaus and the Generic Mapping Tools (GMT) are examples for software packages which are frequently used in this context.

Derivation of terrain parameters

The derivation of terrain parameters is a standard operation that can be realized in almost any GIS environment. Often several trials are necessary to find a visual representation of the surface that sufficiently highlights the crest lines of the dunes.

Tracing of the dune crest lines

The derived terrain parameters are visualized and the expert interpreter traces the crest lines with polylines. Depending on the size of the study area and the number of surveys this can be very time consuming and comprises a high degree of subjective judgement.

Matching of the traced crest lines

Following the previous step the polylines from two or more dates are aligned manually through simple shift of corresponding lines. The transformation does not take into account the commonly encountered deformation of the crest lines and, therefore, only provides a rough approximation of the motion at the centre point of the polyline.

User perspective and expectations, innovation aspects

The outlined processing chain comprises a large degree of manual intervention and the results will depend strongly on the skill and the judgement of the respective expert analyst. Since the matching of the crest lines does not take into account deformation, it provides only a rough approximation of the real displacement and a very sparse displacement field with only one vector per crest line. Furthermore, the derived vectors capture only the horizontal component of the three-dimensional motion. To fully exploit the bathymetric surveys it would be desirable to provide a tool that enables to automatically and accurately recover dense displacement fields, measure all three components of the displacement, calculate volume budgets of the sediment transfer and process multiple time steps efficiently in a batch.

The relevance of the workflow to "big data" issues is synthesized in Table 9.

Table 9. Big Data issues for the MS4 workflow

Indicator	Relevance	Comment
Volume	Medium	Multi-beam surveys can cover several hundred square kilometres and multiple time steps may amount to several GB.
Variation	High	The use of different instruments, vessels and survey protocols among multi-temporal data acquisitions is rather the rule than the exception making an initial pre-processing step to harmonise the input data indispensable.
Velocity	Low	The temporal frequency of bathymetric surveys is currently rather low (typically > 1 month) and the analysis is consequently not time critical.
Analytics	High	Automatic analysis of changes and motion require tools that go beyond a simple

		data transformation to reveal higher level information about the seafloor dynamics.

Proposed IQmulus solutions

Workflow components

This workflow targets to combine several components for the detection and quantification of changes among two or more bathymetric surveys of the same study area. The employed tools comprise algorithms for harmonization of the input data, detection of significant vertical and topological changes and the derivation of a dense horizontal displacement field representing the dune migration. The components of the workflow are the following (Figure 27):

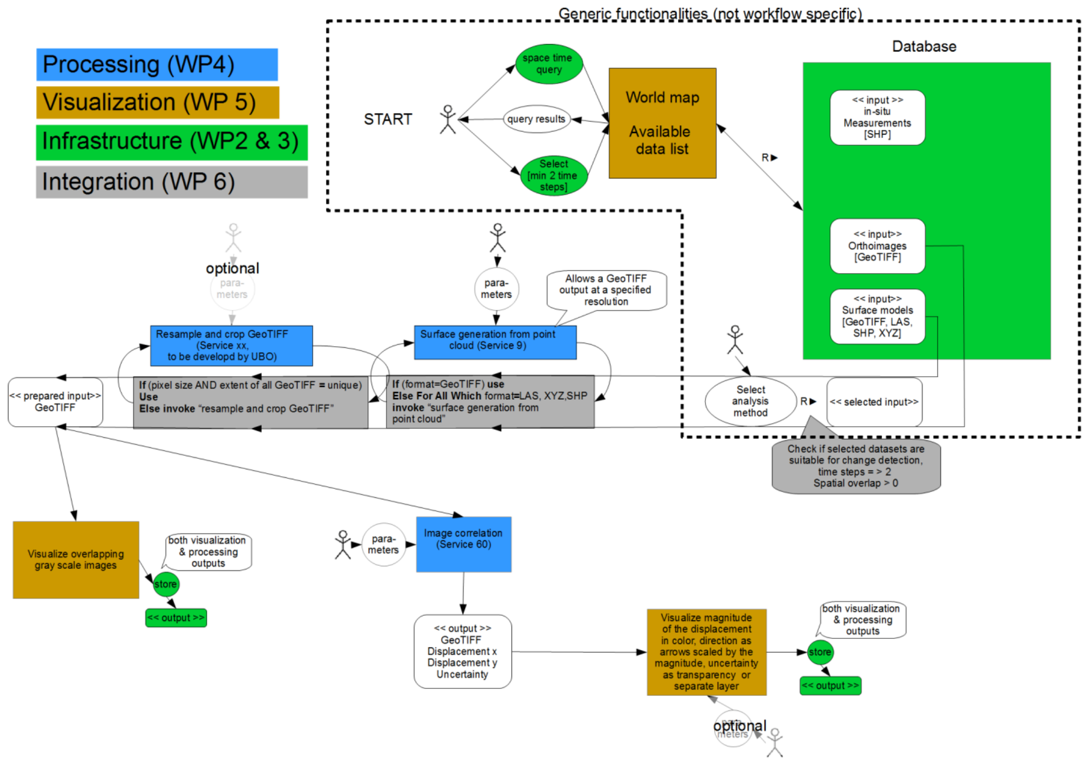


Figure 27. *Overview for the marine workflow MS4 showing the interactions of the user, infrastructure, processing services and visualization.*

Preparation component

Two types of input data can be expected as a starting point for this workflow. The first option is the provision of two overlapping 3D bathymetric point clouds which will require the interpolation of two raster grids with an identical user-defined resolution. The second option is to provide two or more already gridded raster datasets. An initial script checks if all input rasters have the same resolution and extent and if they contain any missing values. If necessary, missing values will be interpolated automatically and a resampling scheme is invoked to upsample the raster with the lower resolution matching the raster with the highest resolution; all rasters are cropped according to their overlapping area.

Generic system functionalities required:

Requirement name	*Functionality*
Storage of resulting raster	Store the resampled raster as GeoTIFF along with the original data. Temporal storage is mandatory and permanent storage will be possible if desired by the user.

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Approximate surface from bathymetric point cloud(LR-splines is used to generate a gridded raster representing the seafloor from bathymetric point clouds)	MS4_1_1	New service to be proposed	Under discussion
Generate spline surface from parameterized point cloud	MS4_1_2	Service #9	Released
Spline surface to grid	MS4_1_3	Service #91	Under discussion
Resampling and cropping of raster **(**A sinc kernel interpolation is used to harmonize the resolution of the input images)	MS4_1_4	New service to be proposed	Under discussion

WP5 visualization functionalities required:

Requirement name	*Functionality*
Visualization of the resulting surface raster as greyscale images	**Data to be visualized:** - raster image of resampled spline surface **Output:** Visualization of the resulting surface raster as greyscale images

Feature detection component: none

Change detection component

Given the input raster hierarchical cross-correlation with spatial regularization will be invoked to measure the horizontal component of the dune motion at sub-pixel accuracy. At this step several parameters can be adjusted by the user or a predefined parameter set adapted to measurements on bathymetric data can be selected.

Generic system functionalities required:

Requirement name	*Functionality*
Storage of resulting raster	Store the resampled raster as GeoTIFF along with the original data. Temporal storage is mandatory and permanent storage will be possible if desired by the user.

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Change detection/Measuring surface displacement (Hierarchical cross-correlation is used to find homologous points in two images)	MS4_3_1	Service #60	Released

WP5 visualization functionalities required: N/A

Presentation/visualization component

Two different outputs should be visualized. First, the derived displacement components (given as two raster representing x and y components) should be converted into directed vectors. The vectors representing the direction of displacement should be displayed e.g. as on overlay on a colour coded image representing the magnitude of the displacement. Optionally it should be possible to display the vectors as an overlay on the input data from the earlier time step. Second, the correlation coefficient will be displayed for example with a colour scale reaching from green (strong correlation) to red (weak correlation).

Generic system functionalities required:

Requirement name	*Functionality*
Storage of resulting raster	Store the resampled raster as GeoTIFF along with the original data. Temporal storage is mandatory and permanent storage will be possible if desired by the user.

WP4 services required: N/A

WP5 visualization functionalities required:

Requirement name	*Functionality*
Visualization of displacement field	**Data to be visualized:** - derived displacement (GeoTIFF) - derived correlation coefficient (GeoTIFF) **Output:** (Combined) visualization of the calculated displacement field and the correlation coefficient.

Input and reference data provided for workflow development and testing

Input data

Dataset ID	Dataset title	Dataset short description	Spatial resolution: Resolution distance (m)	Spectral resolution: Number of bands (for RS data)	Spatial resolution: Point density (points/m2)	Data characteristics: Dimensionality
#12	Digital Terrain Model (DTM)	Bathymetry of submarine sand dunes in Brittany (Four)	2 m	N/A	N/A	2.5D (2D + height/elevation)
#13	Digital Terrain Model (DTM)	Bathymetry of submarine sand dunes in Brittany (Four)	2 m	N/A	N/A	2.5D (2D + height/elevation)

Reference data

Dataset ID	Dataset name	Dataset short description	Format
To be uploaded	Crest line polylines	A multi-temporal set of manually delineated crest lines	SHP

Workflows of the Urban Showcase

For the Urban Showcase setting, we propose two workflows that tackle the issue of updating existing 3D catalogues of urban topographic objects from very different angles:

- Urban Showcase Workflow 1 (US1): Detection of buildings for monitoring and cadastral updating,
- Urban Showcase Workflow 2 (US2): Individual tree extraction from urban LMMS data.

Workflow relations within the showcase

The first workflow concerns a frequent challenge of mapping authorities. It typically involves the integration and extensive use of existing commercial software components and thus also gives the opportunity to investigate how commercial software from the current "ecosystem" of mapping authorities relates to and can be incorporated in IQmulus.

The second workflow concerns Laser Mobile Mapping with its ability to capture large data volumes in a short time. For this currently developing important new technology, the main challenge is how to efficiently extract meaningful metric information on various types of 3D objects in an automated way from the captured huge point clouds sampling urban areas. The actual workflow concentrates on solutions for specific object types, with the ambition of future extensions.

US1: Detection of buildings for monitoring and cadastral updating

Description and As-Is Analysis

Introduction and main steps of the current solution

All over Europe national surveying authorities generate nationwide 3D databases for decision making and reporting requirements (e.g. for calculating noise pollution maps for the European Environmental Noise Directive (2002/49/EC)). Approaches for generating 3D surface models include a number of different remote sensing data acquisition and processing techniques such as aerial laser scanning, stereo-photogrammetric processing or radar interferometry. When it comes to buildings, usually aerial laser scans deliver the best results and also huge point data clouds which have to be processed. Due to the ever-changing nature of man-made structures on the earth's surface, the creation of the 3D building models has to take place in regular intervals. The same applies to tracking changes or comparing the “detected” to the “official” situation in 2D cadastral maps, though in this case the update process is usually performed in situ by surveyors. The surveying process is hereby triggered organisationally and as a consequence varies from country to country in accuracy, actuality and methodology. It is an even costlier process, though also the significance is greater, since this 2D data is relevant for taxing, city planning, and maintenance of infrastructure as well as deriving more generalized map products.

The main steps for the 2D surveying in most European countries usually include informing and getting a permission of an authority and then hiring a surveyor to measure the changes which are afterwards transferred to the national surveying authorities and integrated into the cadastral map. Sometimes there are also aerial images involved to manually or semi-automatically compare them with “official” 2D building footprints and look for changes. In some countries this is a quite accurate and reliable process, in others it is less so.

On the application side, buildings delineated in two dimensions are frequently used to compare “detected” and “official” situations by comparing observed/detected building objects to those present in official (cadastral) records. After proper verification, this information supports decision making at authorities dealing with the administration and authorization of building and

demolition activities. However, there are still some partly or non-automated steps in the process where IQmulus could provide improved solutions. FOMI's near-future plan includes the possibility of building change detection between regular interval surveys and determining the differences to the cadastral data, too. In the long run, the possibility of change detection in 3D is also considered when the input data has higher accuracy and fits better for this application.

When it comes to 3D data the national surveying authorities were and are occupied with the initial creation of the 3D building models till now. For change detection and updating the 3D models, two main strategies exist:

- Transfer changes from cadastral surveying (2D) together with height information to generate simple 2,5D/LOD1 buildings as a starting point to perform an update survey;
- Regenerate the 3D city model from remote surveying sources and compare it with the prior one and/or building footprints.

The first approach fits nicely to existing organisational procedures of 2D surveying, the second one has the benefit that changes which occurred are already provided in 3D and therefore enable a faster and more accurate (LOD2) update which in the long term could prove more cost effective. It is also feasible in countries where 2D cadastral maps are not available nationwide or not up-to-date.

Current solution 1: Operational Building Monitoring at FOMI

A current operational example developed at FOMI is based on satellite image analysis, and is implemented mainly in the eCognition software environment to detect building boundaries. A process chain is elaborated by using the software's own rule-set builder to support easy handling of data and processing, described in detail below.

Input data of the OBIA methods:

- Nation-wide mosaic of colour-infrared orthophotos (8-bit quantized, spatial resolution of 0.4 m) and for the major cities "quasi-true" ortho images are used. In the latter case, orthocorrection is carried out by using the photogrammetry-derived digital surface model (DSM) instead of the DTM. This leads to a better orthogonal representation of buildings, though has its caveats due to data gaps resulting from not-ideal acquisition conditions).

Figure 28. *Colour-infrared orthophoto*

- The aerial photo survey provides the digital surface model (DSM), and contour lines of the topographic maps covering the country (DEM). Both models are in raster format and pixel values contain the height in meters. The DSM has 1 m; the DEM has 5 m spatial

resolution. The difference of DSM and DEM per pixel gives the heights of the individual objects in the surface.
- Polygons of the building contours from the national building cadastre.

Main processing steps of the current solution – overview:

1) Preprocessing of CIR orthophotos and creation of "quasi-true" orthophotos
2) Preprocessing of DEM
3) Creation of DSM
4) Calculation of the DSM-DEM difference
5) Creation of the mosaic for 1:100.000 scale sections for the difference raster layer and the orthophotos.
6) Preprocessing of cadastral dataset.
7) Processing – segmentation, where several input data are processed, namely:
    a) creation of 5000x5000 px tiles for the processing (automatic)
    b) segmentation (can be processed in parallel)
    c) stitching of the segmented tiles
8) Processing – classification, OBIA filtering
9) The result is a vector map with the classified building boundaries and islands with a unique identifier.
    a) Merged and dissolved vector data
10) Postprocessing

Preprocessing – details:

The image processing and data storage are managed in the Hungarian Projection System (HD72) with sections of the 1:100.000 scales in the whole workflow. The preprocessing of input raster data is operated by the photogrammetric working group (5 people). The input vector datasets are pre-processed by the geoinformatic working group (3 people). In the implementation of the current workflow it was an important consideration that the processing units can be processed independently. The current processing cluster contains four workstations.

Preprocessing of raster data (we do not consider the photogrammetric workflow as a part of the preprocessing of the remote sensing image processing):

- Mosaic of the DSM-DEM difference for 1:100.000 sections (Erdas Imagine)
- Mosaic of traditional orthophotos and "quasi true'' orthophotos for 1:100.000 sections (Erdas Imagine)

Preprocessing of vector data:

- Exporting cadastral building boundaries and parcel boundaries with attributes needed for post-processing from the country-wide database to ESRI shp files for each settlement of the country stored in individual files (PostgreSQL).
- Collecting the above-mentioned country-wide input vector datasets for post processing

Processing - details:

The image processing is carried out with the object-based image processing method (OBIA). In this method the main idea is that instead of using only the spectral properties for each individual pixel the method is treating as a visual object a homogeneous set of pixels by their similar spectral properties. In contrast to the pixel-based processing, in the classification besides the spectral characteristics also geometric properties can be used. The buildings - unlike vegetation and water surfaces - cannot be recognized with certain accuracy using only spectral properties.

The two main steps of the OBIA method are the *segmentation* and the *classification*.

During segmentation, image objects (segments) are delineated. Each segment is homogeneous according to a selected set of properties (metrics) and consists of a connected set of neighbouring pixels. The result of the segmentation is the segment map, which provides the information for each image point which is in the enclosing segment. The process consists of iterative steps in two different approaches to create the segment map. The 'Bottom-up', where the homogeneous groups of pixels are merged, and the 'Top-down', where the inhomogeneous groups of pixels are cut down.

Considered properties during segmentation:

- spectral values (intensity values per bands)
- NDVI (Normalized Difference Vegetation Index) NDVI=(RNIR-RRED)/(RNIR+RRED)
- height (DSM and DEM difference)
- cadastral building dataset.

At this level, image objects do not (necessarily) correspond to buildings. Generally, one building consists of several objects, but in the case of buildings standing side by side the inverse situation may also occur.

Figure 29. *Segmentation result: image object boundaries shown with black lines.*

During classification, for each object it is clearly decided whether it is a building, a potential building or cannot be a building. The decision is based on the NDVI and the average height of the object. The potential building category contains the objects included in the building cadastre but which cannot be detected on the orthophoto, where it may occur that the entire building is obscured by high trees. Such cases cannot be solved using the available input dataset. After evaluating the first results an additional class implementation was introduced: the class of "high buildings". Based on tests the limit of the high buildings is 13.5 m.

Further clarification is carried out by using different condition systems and filtering. Geometric filtering is also used. After the classification, generalization is also needed in the results, applying morphological operations (erosion, dilatation) to the "imagined" building boundary and simpler building polygons are obtained. These polygons are further processed via GIS processes in vector format.

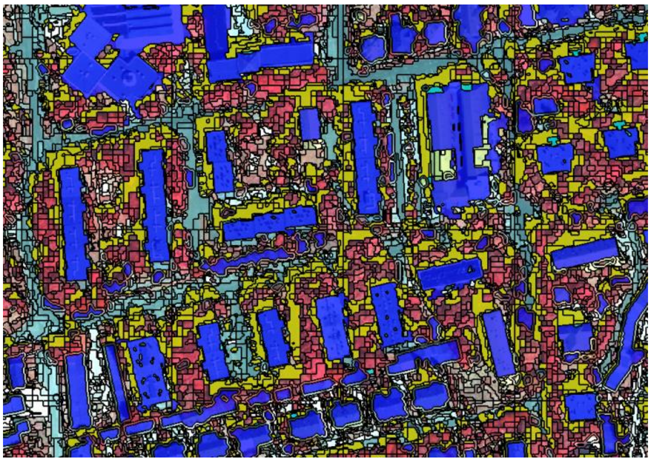


Figure 30. *Classification result. Detected buildings shown in blue.*

Performance:

One section of the 1:100.000 raster layer takes 16-24 hours (processing the segmentation and the classification) depending on the number and the geometry of the resulting objects.

The processing environment for remote sensing image processing is eCognition.

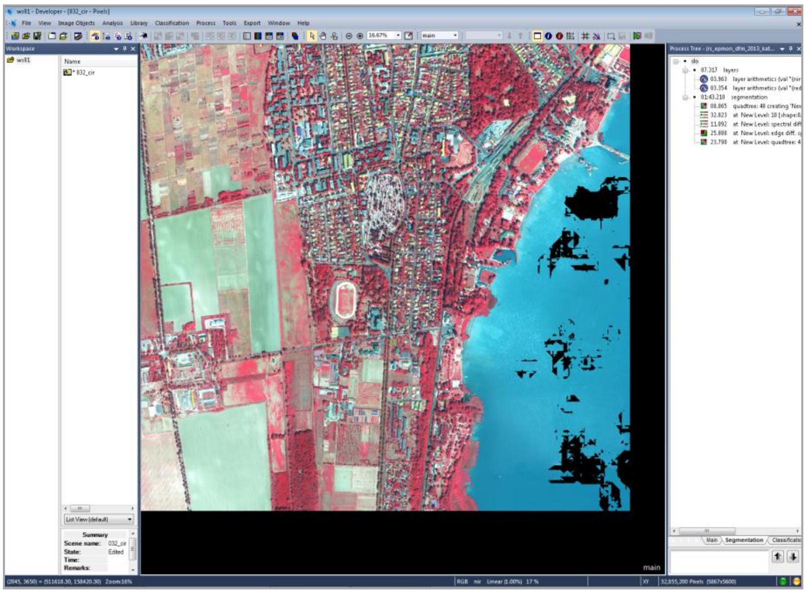


Figure 31. *eCognition software GUI – an example*

The results of the segmentation and classification are stored in vector format (ESRI shp file) for post-processing in a GIS environment.

Main steps of the current solution of the processing (eCognition)

1) Define area of interest (AOI)
2) Creating tiles (5000x5000px) automatically inside eCognition: needed because of the software performance limit
3) Segmentation:
    - NDVI calculation
    - Brightness filtering
    - Quad-tree segmentation
    - Multi-resolution segmentation
    - Spectral Difference calculation
    - Contrast Split
    - Quad-tree segmentation (force cadastral boundaries)
4) Classification:
    - NDVI calculation
    - High building classification
    - Shadow 1-2 classification
    - Building classification
    - Possible building classification
5) Filtering
    - Geometric filtering
    - Generalization

Postprocessing (ArcGIS environment)

The aim of post-processing of the detected building boundaries is to detect the changes. The result of the change detection is a classified thematic map with four thematic classes for each cadastral parcel: No difference, enlargement, reduction, and not decidable.

Figure 32. *Post-processing: rule-based filtering of false detections by using external data sources. Objects removed in this step are shown in red.*

Input data for post-processing:

- public route dataset
- railway dataset
- water bodies dataset

- land cover dataset
- cadastral building boundaries
- cadastral parcel boundaries

Main steps of the postprocessing:

1) Creation of buffered and clipped route dataset
   - buffering by 8m
   - clipping by the cadastral parcels
2) Land cover dataset preparation
   - selected land cover classes converted to clipping layer
3) Water bodies dataset clipping
   - selected water bodies classes merged to one nation-wide clipping layer
4) Merging all the input datasets
5) Intersecting and erasing the merged layer elements from the detected building boundaries layer.

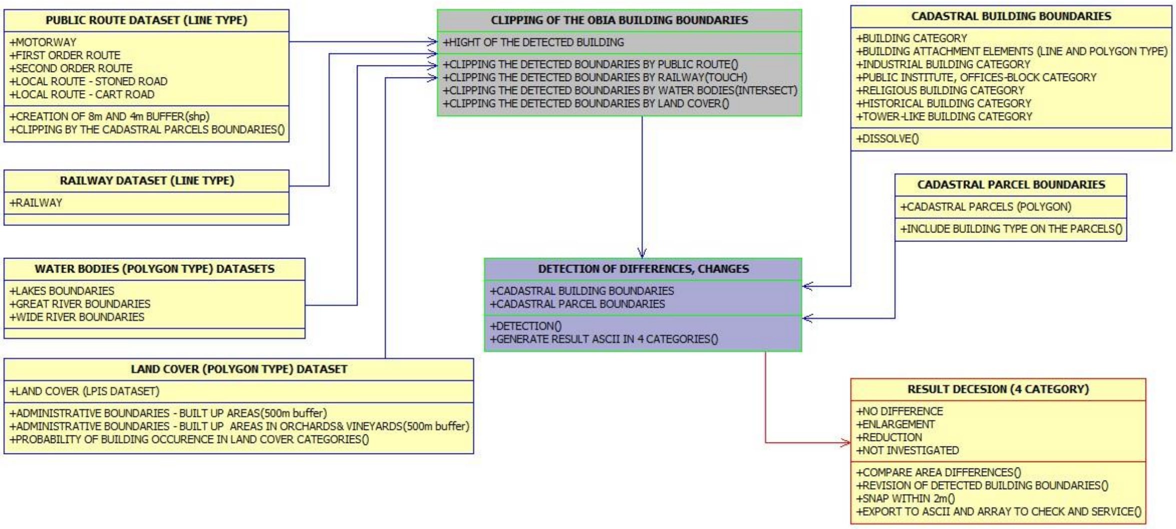


Figure 33. *Workflow of the post-processing and the detection of changes and differences*

To illustrate the current workflow for building detection, a Use Case diagram is presented below:

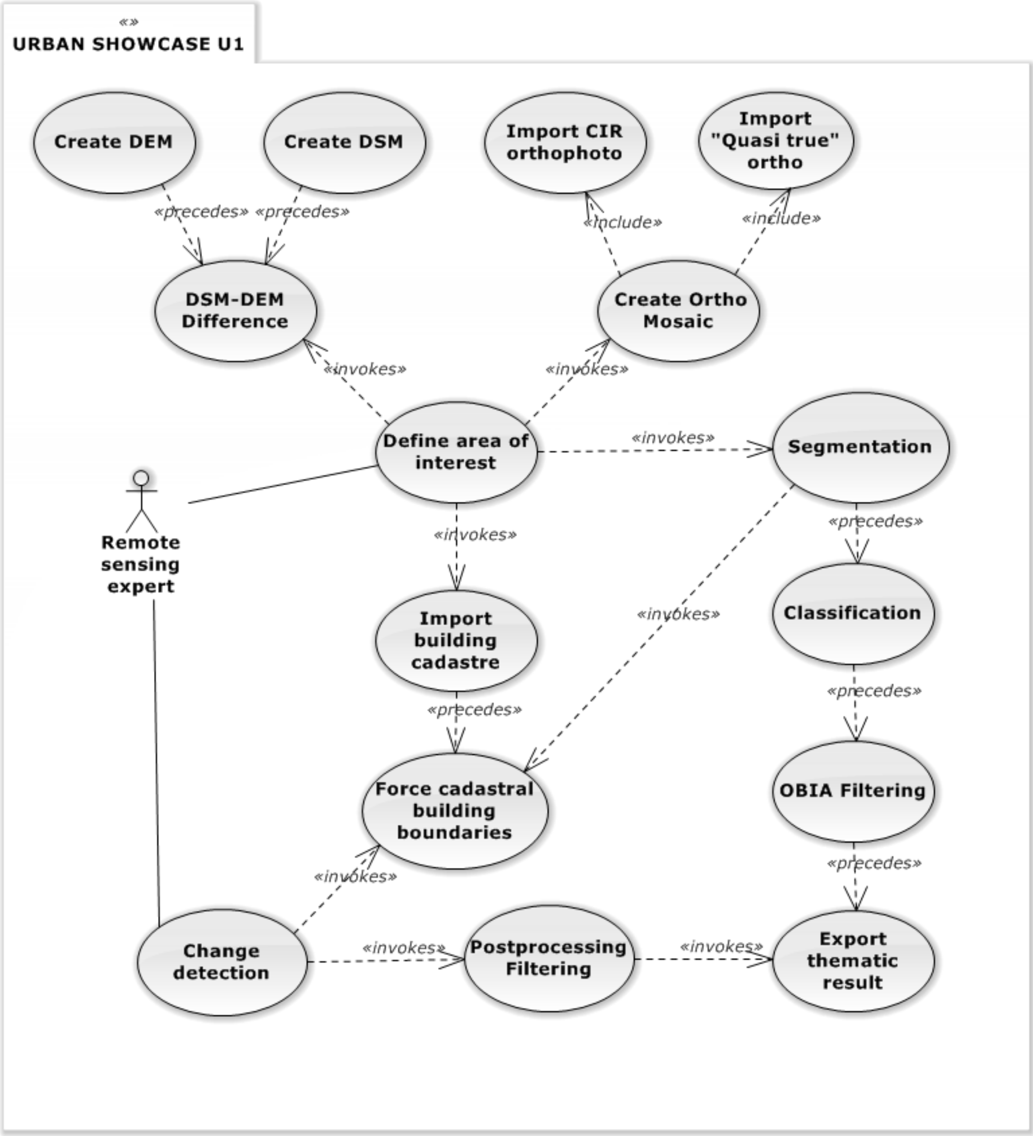


Figure 34. *Use Case Diagram for Building Detection (current workflow)*

Current solution 2: M.O.S.S. novaFACTORY with Tridicon CityMapper

The following workflow is used by several of the federal states in Germany for producing and maintaining state wide building models. The data is managed in novaFACTORY (DTM, orthoimages, building footprints). For building generation this data is exported for a given region. The following figure shows the GUI approach, though for production for a bigger region an automated export process generation is used.

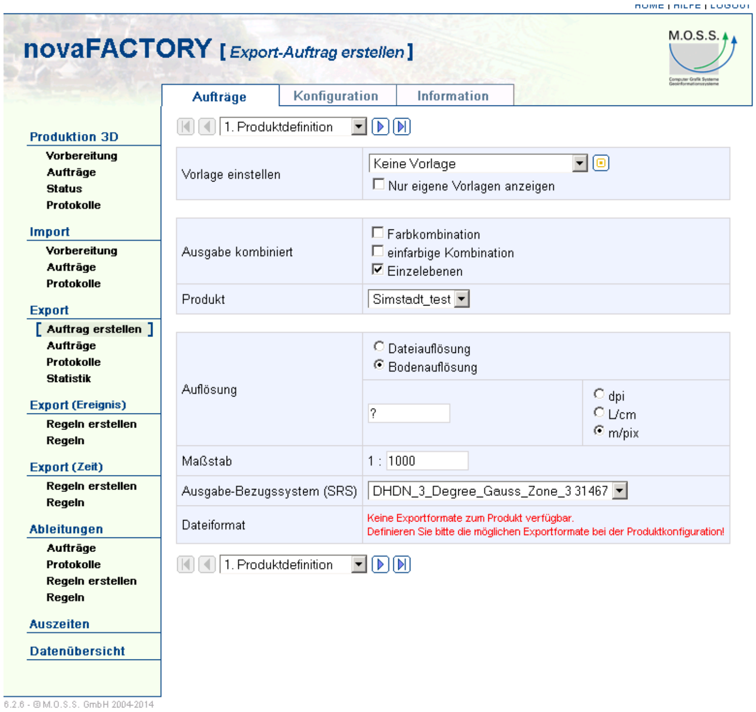


Figure 35. *novaFACTORY GUI*

The building generation with 3D CityModeller can take place with the help of a graphical user interface or as a batched process for mass production. In batched mode the novaFACTORY export process is extended by a script which restructures and moves the necessary input data for building generation to the 3D CityModeller import folder(s). In the following the GUI process is described but its steps are essentially the same ones as in batched mode.

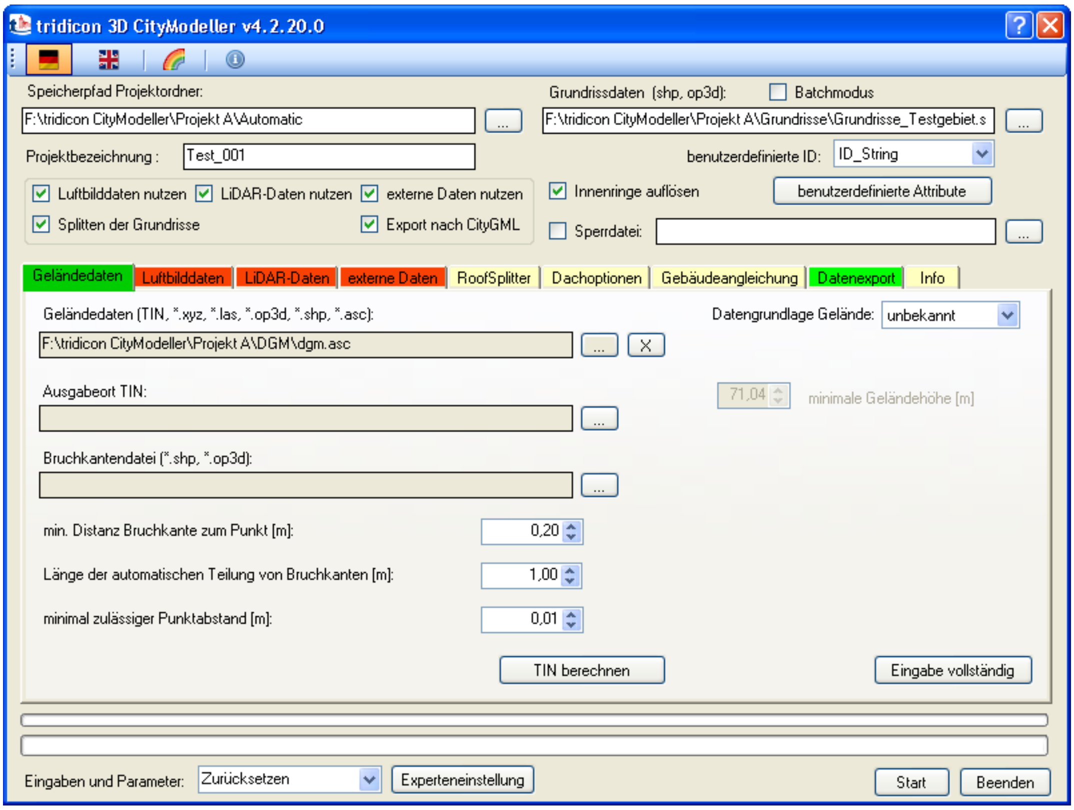


Figure 36. *CityModeller GUI*

Input data

Building Footprints: A .shp or .op3d file containing the building footprints, e.g. from cadastral surveying.

DTM (optionally): In the beginning a DTM is provided as input in formats TIN, *.xyz, *.las, *.op3d or *.shp. It can be used for calculating the height of the terrain for building footprints and for generating a DTM for CityGML city models. A DTM could also be generated automatically from the other data.

Apart from building footprints remote surveying data containing 3D measurements is necessary for building generation. This can be provided either as stereoscopic orthoimages, as a LIDAR dataset or as a point cloud from ground-based laser scans.

Stereoscopic images: They can be provided in .vit or .tif format. Also an orientation file in .gori format is necessary.

The orientation files can also be generated manually by comparison with the building footprints as shown in the following figure.

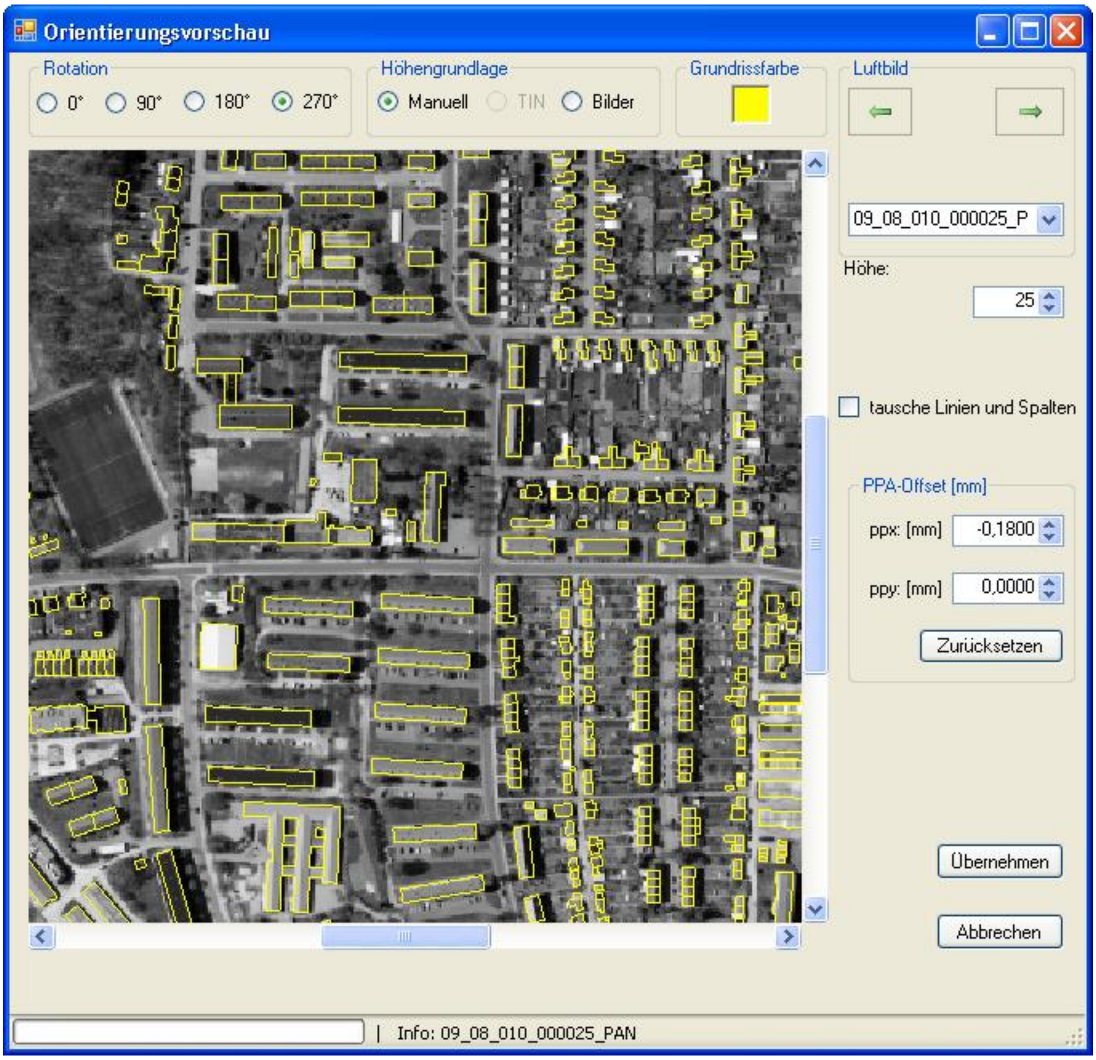


Figure 37. *Manual updating of building footprints*

LiDAR: Airborne laser scan data can be provided in the following formats *.txt,*.dat,*.las, *.vef, and *.lpl. They can be used categorized, filtered by intensity or pulse or raw.

External data / Point clouds: As mentioned before point clouds do not necessarily have to be LIDAR data but can also be other data, e.g. from ground-based laser scans. Formats are also *.txt,*.dat,*.las, *.vef, and *.lpl.

With the given input data the building generation can take place by analyzing the 3D elevation at the given building footprints.

The next step in the processing usually is the **RoofSplitter**, a tool which can be used automatically, interactively or preprocessed. Buildings till now might consist of plenty of polygons. The RoofSplitter detects the most likely roof structure of the building based on the 3D point cloud. Thereby the building can be separated into different building parts which can be further simplified. Parameters for the RoofSplitter include tolerance values for distance, area, point- and polygon number and roof structures which should be considered.

The following figure shows some of the options for roof structure detection.

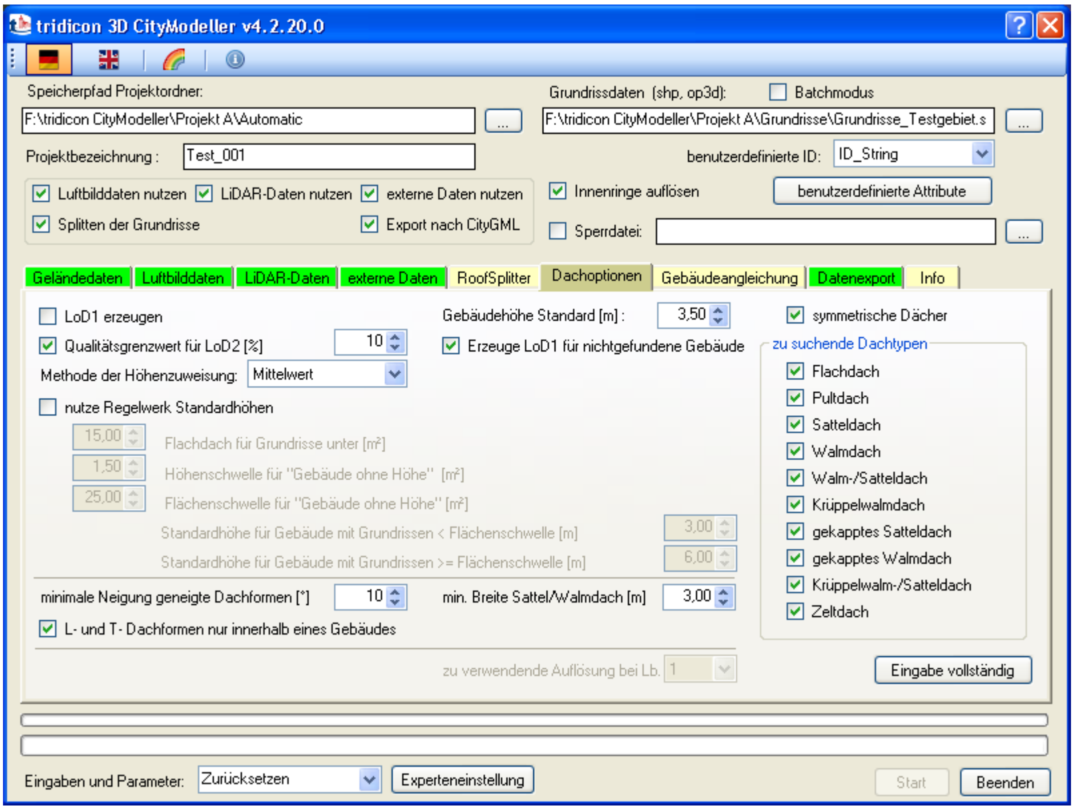


Figure 38. *Options for roof recognition*

It should also be mentioned that minimal detection accuracy values can be set for the building generation process. If the CityModeller cannot detect the correct roof structure with - for example - more than 80% probability it will simply generate a LOD1 geometry and skip the roof detection for LOD2 buildings. But usually the roof generation can be performed successfully for up to 90% of the buildings with sufficient quality.

CityGML generation

There are several further options for post processing and CityGML generation:

- averaging heights
- filling certain attributes (like building height)
- coloring of roof and wall areas
- LOD1 and 2 export
- CRS

Import in novaFACTORY

The generated CityGML files can afterwards be imported (also automated) into novaFACTORY for warehousing and management of the 3D data. Coherent management of the data is important since often the given building asset for a certain time period is of interest. Data access via web service (e.g. custom http, wfs or wps) is also made available through novaFACTORY, whereas visualization currently is mainly handled on the client side based on specific solutions using e.g. CityGML visualizers or SketchUp.

User perspective and expectations, innovation aspects

Over time updates of existing data sources become more time and cost intensive than the original acquisition. Therefore it is of great interest for the users to automate and simplify the process of updating as far as possible while maintaining the actuality and quality of the data.

For the update of 2D cadastral data itself we cannot offer any improvements, at least in terms of accuracy. But the point cloud from remote surveying can nonetheless produce 2D footprints of buildings. These could be used for example to detect discrepancies between official cadastral data and remote surveyed data.

For the update of 3D building models we suspect the reason that till now few national surveying authorities use the comparison of existing data with a newly generated 3D city model as the means for updates, is that the creation of a 3D city model is a costly and time consuming process. Now if we manage to provide the means to effectively handle huge amounts of data automatically, accurately detect buildings in LOD2 and reliably detect changes to prior data, we can make this approach more appealing and effective for the national surveying authorities.

The generation of buildings in 2D and 3D from point clouds can be interesting in countries where cadastral surveying is not available. This may prove especially interesting when it comes to disaster management, e.g. to quickly assess building damage during a natural disaster like flooding or earthquake events where ordered bureaucracy like cadastral surveying is not an option.

Expectations for improvement over the current operational solutions

The expectations are mainly focused on automation and improvement of several steps of the existing workflows to make it more efficient in the long run. Integration of third-party software components (such as eCognition or Tridicon) from the current operational environments of Mapping and Cadastral Agencies is to be tested as this would also allow reusing existing classification rule sets.

More specifically, we expect improvements in the following specific aspects:

- Automation of workflows to be able to run country-wide projects with minimal supervision.
- Detection of buildings in 2D and 3D without the dependency on available 2D cadastral data
- Integrating photogrammetric processing of stereoscopic aerial imagery to extract surface model or even 3D building objects directly.
- Distributed processing to decrease processing times for country-wide mapping projects.
- Handling of really large datasets to avoid the need for data partitioning during pre-processing.
- Increase detection accuracy by minimizing both omission and commission errors.
- Publish data and services in a web-based environment to facilitate interaction and data sharing.

The relevance of the workflow to "big data" issues is synthesized in Table 10.

Table 10. Big Data issues for US1 workflow

Indicator	Relevance	Comment
Volume	High	Point clouds for large areas (up to countries), Existing vector data for buildings 2D / 3D, optionally raster aerial and satellite images for manual verification.
Variation	Medium	Original image data (optical); Point clouds and/or raster-based elevation models (from laser scanning or stereo-photogrammetry); Calculated spectral indices; Vector data (cadastral and supplementary data sources).
Velocity	Low / High	For continued governmental surveying tied to update cycle (e.g. yearly) without tight deadline the velocity is low. In case of disaster management the velocity is high.
Analytics	High	2D and 3D building detection from point cloud data, roof type detection for 3D buildings, change detection for 2D and 3D vector data. Nation-wide cadastral dataset also has high volume and needs extremely high processing performance for spatial queries and geoprocessing. High importance of the accuracy of the results.

Proposed IQmulus solutions

The challenge in this use case is not about exploring new scientific boundaries. The theories and technologies as well as commercially available software already exist. This use case is more about combining them to a mostly automatic, user-friendly workflow which can deal with huge amounts of data, also fulfilling other requirements raised by the users. Moreover, integration of such solutions within the IQmulus platform will allow testing of the IQmulus platform, services and process chains. Integrated commercial solutions can also be useful in benchmarking new developments and algorithms.

Workflow components

The automatic change detection can be divided in five main steps:

- generation of buildings based on the new point cloud / surface data
- selection of area of interest
- comparison of the existing buildings with the newly generated ones
- presentation of the results to a human decision maker
- updating the 3D building model
- preparation component

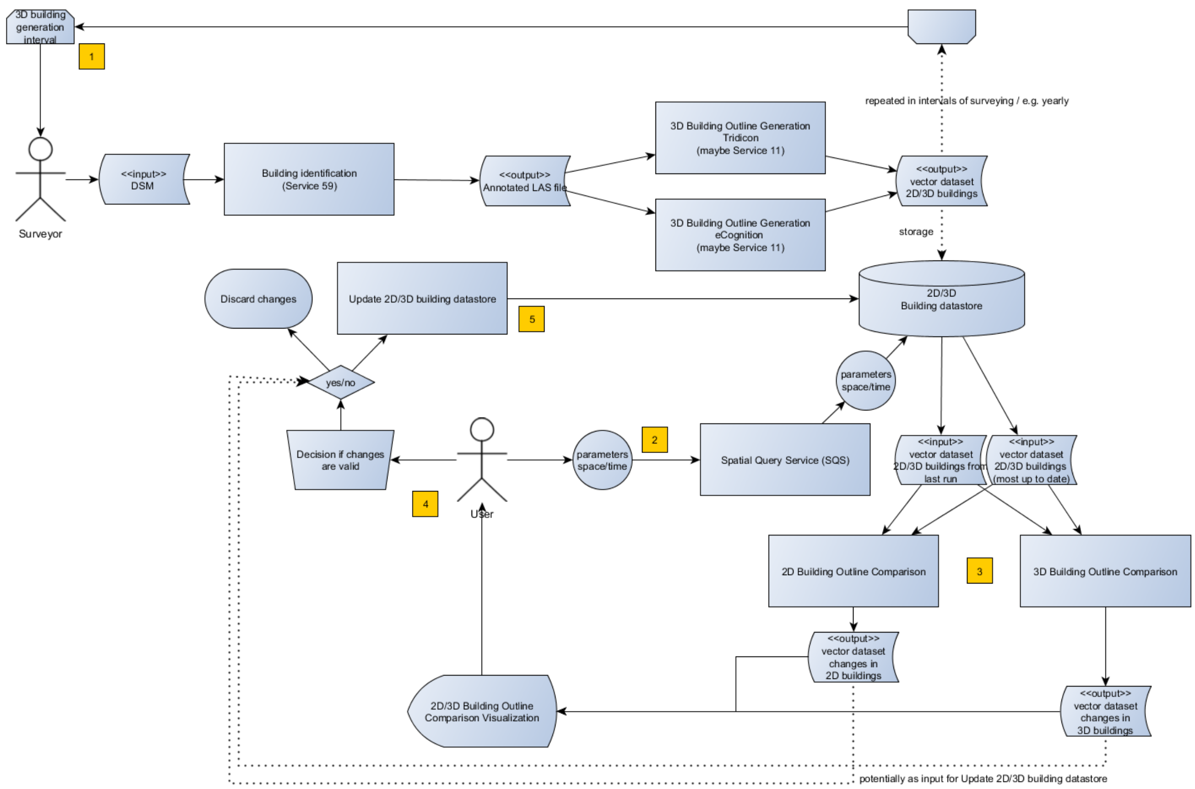


Figure 39. *IQmulus workflow for building detection and update*

Though we will focus in this workflow on change detection of buildings it can be extended to other objects of the architectural environment and the landscape. As a result we also want to evaluate the possibilities to transfer the workflow and the methods developed here to other environmental objects in a later phase of the IQmulus project.

Preparation component

Since the user has to validate the detected changes in the building asset the change detection is restricted to an Area of Interest. The user starts the process by selecting this area.

Generic system functionalities required:

Requirement name	***Functionality***
AOI selection	Selection of an area of interest for further processing

WP4 services required: -

WP5 visualization functionalities required: -

Feature extraction component

In this step data from remote surveying – in our case a new point cloud of LIDAR-data or stereoscopic aerial imagery in the given area of interests – is used to extract buildings. Professional and commercially already available software is used to generate the buildings based on the point cloud. Two software suites will be compared to evaluate which is better suited for an automated process with large data volumes and delivering accurate results. The

result of this step 1 will be outlines (2D) or solids (3D) of buildings based on the newly collected DSM data.

eCognition (FÖMI)

An online description can be found here: http://www.ecognition.com/

In D2.2 the software was already described:

"Object-based image analysis (OBIA) provided by eCognition is an intelligent way for geospatial information extraction and as such relevant for the project. It can provide help for the following several steps:

- Segmentation of raster and point cloud data
- Hierarchical organisation of spatial objects
- Calculation of diverse metrics including shape, colour, texture, neighbourhood, distance, etc.
- Customized classification and feature extraction by utilising fuzzy logic based on membership functions
- Communication with the GIS world in terms of importing and exporting standard raster, vector and point cloud formats

eCognition is a complete environment for importing data and carrying out analysis / feature extraction / classification tasks in a workflow-based manner. The suite consists of three products:

eCognition Developer for the development of custom rule sets for specific feature extraction and classification tasks.

eCognition Architect to configure, calibrate and execute image analysis workflows based on rule sets developed in eCognition Developer.

eCognition Server to automatically process thousands of images and perform detailed analysis in a single, fully automated run, based on rule sets developed in eCognition Developer."

FOMI has a lot of experience in the usage of eCognition and will introduce it in this project. eCognition processing functions have already been used within novaFACTORY, and can be exposed as services within the IQmulus platform. Rule sets of current solution developed by FOMI will be made available for testing this approach.

Tridicon BuildingFinder (MOSS)

An online description of this software can be found at the following URL: http://www.tridicon.de/software/tridicon-buildingfinder/?L=1

As it is stated on the homepage: "In a fully automatic process the tridicon BuildingFinder generates building footprints, 3D buildings and even corresponding textures from the used imagery.

Compared to the tridicon CityModeller the tridicon BuildingFinder does not require the input of existing building footprints for detecting the buildings. This makes the tridicon BuildingFinder very useful for the generation of 3D building models in areas for which no building footprints are available."

MOSS has plenty of experience using the tridicon CityModeller which needs 2D building footprints as input. Since the software proved very accurate in the process of generating 3D buildings for several German federal states it is of interest to evaluate how well the BuildingFinder performs, especially when it comes to generating building models for countries with less reliable or up-to-date cadastral surveying (e.g. for disaster management).

Generic system functionalities required: -

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Building identification (Identify ALS data that is sampling buildings	US1_2_1	#59, Multi-object classification of 3D Point Clouds,	Proposed
2D Building Outline Generation (Use commercial eCognition software to generate 2D outlines for the ROI)	US1_2_2	[Commercial third-party component: eCognition]	Proposed
3D Building Outline Generation (Use commercial Tridicon software to generate 3D outlines for the ROI)	US1_2_3	[Commercial third-party component: Tridicon BuildingFinder]	Proposed

WP5 visualization functionalities required: -

Change detection component

2D comparison:

The existing building dataset includes polygons of the building footprints either from cadastral surveying, from derivation of point cloud data or segmentation from orthoimagery and DSM. These footprints are geometrically compared with the areas of the newly generated buildings from step 1 to detect changes which occurred since the last surveying cycle.

Once again the building solids of the prior surveying cycle are compared with the newly generated solids of step 1.

Detected changes result from the following reasons:

- missing buildings of the former data pool in the new collection:
  - select all existing buildings without an intersect with areas of buildings of step 1 based on 2D comparison;
  - these buildings are not deleted but marked as missing.
- new buildings:
  - select all areas of buildings of step 1 without an intersect with existing buildings based on 2D comparison;
  - These areas are marked as new.

- changed buildings:
    - select all existing buildings which intersect with footprints of buildings from step 1 based on 2D comparison
    - analysis of this selection:
        - 2D comparison based on calculated differences between the areas A1 (existing buildings) and A2 (new buildings of step 1)
            - The following examples show varieties of this calculation:
                - area (A1 – A2) / area (A1) > Δ
                - area (A2 – A1) / area (A2) > Δ
                - area (A1 intersect A2) / area (A1 + A2) < Δ
                - area (A1 xor A2) / area (A1 + A2) > Δ
                - Δ describes predefined variances for the different compares.
        - 3D comparison, based on calculated value of the similarity of the 2 solids
            - open issue for research: How to identify these similarities?

The areas selected in one of the above examples are marked as "changed" and delivered as a vector dataset with 2D/3D data.

Generic system functionalities required: -

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
(2D Building Outline Comparison Project 3D building outlines to obtain 2D building outlines, compare result to existing 2D building outlines, and distinguish between a) no change, b) growing or new building, and c) shrinking or disappeared building	US1_3_1	proposed service – to be added	Proposed
2D/3D Building Outline Comparison (New 3D building outlines are compared to existing 3D building outlines)	US1_3_2	proposed service – to be added	Proposed

WP5 visualization functionalities required: -

Presentation/Visualization component

1) As result the user gets a map of the selected area where changes in the building asset are highlighted. Following categories are displayed:
    a. removed buildings
    b. new buildings
    c. changed buildings with an estimated detection accuracy.
2) All detected changes are based on statistical and geometrical analysis and will only help in the process of the decision making. In the end the user has to decide how the automatic calculated changes are processed. That is why the user has to define in a following external step which calculated changes should be transferred into the data pool.

Generic system functionalities required: -

WP4 services required: -

WP5 visualization functionalities required: -

Requirement name	*Functionality*
2D/3D Building Outline Comparison	**Data to be visualized:** - Cadastral records (map) - detected changed building outlines/shapes (2D/3D) **Output:** 3D visualization of base data and detected changes (2D/3D)

Update of building database

Approved changes in the building asset are transferred into the building database as the most up-to-data building version. It is worth to mention that old versions of the building asset usually are kept for monitoring changes over time and for keeping a fixed state for a certain period with judicial relevance.

Please note: This step is described for theoretical understanding of the showcase but is not implemented in the showcase.

Input and reference data provided for workflow development and testing

This workflow relates to a Hungarian project aiming at the country-wide monitoring of building activities. This project involves several terabytes of orthophotos plus DSMs so it is definitely "big data". The workflow will be implemented on a smaller area however. Sample data sets are already available for the city of Eger on the IQmulus SFTP server.

Input data

Dataset ID	Dataset title	Dataset short description	Owner
#5	Orthophotos, Eger, Hungary	Multichannel Imagery (RGB+NIR) Ortophotos produced from aerial photographs taken 2000, 2005 2007, 2008, 2009, 2010, 2011, 2012 with native resolution of 0,5m/px.	FOMI
#14	Existing Building Outlines, Eger, Hungary	• Contains building outlines as 2-D polygons for a study area in the town of Eger, Hungary. extracted from the official digital cadastral dataset and hence represents the "official" situation. It was created by using on-the-spot measurements by land surveyors. • Data volume is not substantial; this is reference data.	FOMI

#6-7	“New” stereophoto-grammetric data, Eger, Hungary	• Sample area over Eger (town in Hungary)
		• Bounding box (Latitude, Longitude): [47.95287, 20.41394], [47.86223, 20.33154]
		• Digital Surface Model (DSM):
		  o Created by multi-view photogrammetric processing of aerial imagery.
		  o Original resolution of images: 0,4 m
		  o DSM available both in a raster of 1 meter regular sampling and in point cloud format.
		• Digital Elevation Model (DEM)
		  o Created by interpolation from contour lines (scale: 1:10 000) and updated by stereo-photogrammetric processing
		  o Spatial resolution: Resolution distance (m): 5
			FOMI

US2: Individual tree extraction from urban LMMS data

As an introduction, let us consider a detailed "futuristic" fully interactive workflow, casually described as follows:

1. *I am working as a data provider in close cooperation with the city municipality of Paris. It is my task to extract and monitor a given category of objects based on LMMS data. I get an email from Alice: the hard-disks with the new LMMS data set of Paris have arrived. I collect the hard-disks and* ***upload the new data to IQmulus.***
2. ***The Metadata Browser shows me the data properties****. The data is in PLY format as expected. There are also oriented images along with the point cloud*
3. ***I ask IQmulus to show me the data extent of the new data.*** *I hope they didn't forget to scan the Obelisk this time.*
4. *The extent looks OK, but let's check if the data itself looks OK as well.* ***I ask IQmulus to load the cloud representing the Rue du Mont Thabor****.*
5. ***I ask IQmulus to visualize the cloud representing the Rue du Mont Thabor***
6. *The default colors look a bit funny.* ***Let's colorize by distance to the laser sensor.*** *Spatial Query based on the mouse click: return 5m around the click and visualize that. Possibly some default map on the background*
7. *Good, now apply RGB colors to the point cloud: Additional information: acquisition time + location of point cloud + images.*
8. *Then I can check the result.* ***Colorizing by the RGB data.***
9. *Fair enough, the point cloud looks even better than the previous one.* ***What is again the reported quality according to the Metadata Window?***
10. *Let's start working. First take all cars out. What approach should I take?,*
    - ***Remove cars, just based on this epoch of point cloud data?***
    - ***Remove cars, while in addition incorporating the RGB data?***
11. *Let's* ***visualize the result for the Rue du Mont Thabor, the green areas are the locations where the algorithm was sure it removed a car, while for the red areas it is more in doubt****. Indeed, that looks more like a phone booth.*
12. *According to the* ***Preview window,*** *5147 certain cars, and 430 uncertain cars were removed from 10km of LMS data!*
13. *Still, the result is acceptable.* ***I store the cleaned dataset but keep the uncertain cars.***
14. ***I make a note, later I should manually check the `uncertain cars`.***
15. *The botanic department asked me to include a tree inventory this time. Let's check this* ***new tree identification service for Rue du Mont Thabor.***
16. ***Results look good, Preview Window, colored by confidence,*** *although I think that's not a tree but a cherry-picker.*
17. *OK, hit it.* ***Give me all trees for the full data set!***
18. ***Preview Window/Process metadata window?? The system indicates that may take 40 minutes,*** *time to discuss football with Bob...*
19. ***Separate points classified as trees into individual trees****.*
20. *Ready in 20 minutes, cool.* ***Preview Window: 5230 certain trees,*** *more than certain car, that must be because the Tuileries Gardens are partly in this data set as well.*

21. *Let's check* ***how well these trees are actually separated.*** *OK, ..., maybe send a screenshot of that to* bugs@iqmulus.eu
22. *Anyway,* ***let's save all these trees for the moment.***
23. *One thing left, the facade check for Rue Cambon! Check what people changed on their facade.* ***I select this TLS data set we got 2 months ago from the IQmulus cloud using the #DataBrowser.***
24. ***Take the intersection of the LMMS and the TLS data set.***
25. ***Co-register the two datasets to avoid the detection of false changes.***
26. ***Determine cloud to cloud distances between TLS data and LMMS data.***
27. ***What did really change? Stochastic change detection between TLS data and LMMS data.***
28. *Optional: Characterize changes in terms of objects****,*** *Classification on the changes.*
29. ***Visualize changes.***
30. ***Save result,*** *Output: PLY with possibly additional attributes.*

While such an intricate and highly interactive workflow is clearly beyond current capabilities, a more limited but important piece of it has been extracted as the second workflow of the Urban Showcase US2 - Individual tree extraction from urban LMMS data.

Figure 40. *Individual tree outlines from boomregister.nl*

As-is analysis

Introduction

A Laser Mobile Mapping System (LMMS) typically consists of a car equipped with lasers, cameras, GNSS and IMU systems. The GNSS and IMU are used to infer the position and orientation of the car during operation, while the lasers are used to determine the range distances between car and its surroundings. Combination of GNSS, IMU and laser data results in a georeferenced point cloud. The photos from the cameras are calibrated and similarly georeferenced so to be used as inputs for, e.g., orthophotos or point cloud texturing.

Laser Mobile Mapping has been around for a few years but is still not in use at a large scale by end users of geodata. The main strong point of LMMS is at the same time the reason for this relatively slow development, namely its ability to capture large data volumes in a short time. The main overall challenge that is addressed in this workflow is how to extract meaningful metric information in an automated way from these captured huge point clouds sampling road vicinities.

Main steps of the current solution

Current algorithms aiming at the extraction of metric and quantitative information from point clouds were typically demonstrated on small point clouds (< 100.000.000 points) sampling a few adjacent facades at most. Therefore there is really no current solution.

User perspective and expectations, innovation aspects of IQmulus

The goal of this workflow is to extract meaningful information at the neighbourhood level, which is innovative by itself. On top of that methods for extracting e.g. trees from LMMS data are certainly under development at different research institutes, but cannot yet be viewed as well established. It is still notably considered challenging to successfully identify individual trees from part of a point cloud classified as sampling trees.

This workflow will make use of a point cloud of 10 km acquired by a Laser Mobile Mapping system (LMMS). Such a data set is a perfect example of a big dataset. So far, this type of data has been mainly used by research institutes and data providers. The popularity of Google StreetView data gives some indication of the potential of LMMS point clouds. The current challenge is to extract meaningful metric information from LMMS point clouds. If this turns out to be possible, and IQmulus is a perfect platform to demonstrate this, it is expected that city municipalities will integrate LMMS data in their urban management facilities.

Table 11 Big Data issues for the US2 workflow

Indicator	Relevance	Comment
Volume	High	The data volume is huge as up to 3 million points can be generated per second (depending on the system) for a single day's campaign of up to 8 hours.
Variation	Medium	Point cloud acquired by one Laser Mobile Mapping system (LMMS) sampling road vicinities.
Velocity	Medium	To maximize the system's utilization the data needs to be downloaded and processed overnight.
Analytics	High	Extraction of meaningful metric information from LMMS point clouds is an important challenge.

Proposed IQmulus solutions

Workflow components and diagram

This workflow focuses on the extraction and counting of individual trees from an urban LMMS data set sampling a city neighborhood. Below the workflow is decomposed into components belonging to the classes i) preparation, ii) feature detection, iii) classification and iv) presentation. Note that this workflow does currently not foresee a change detection component.

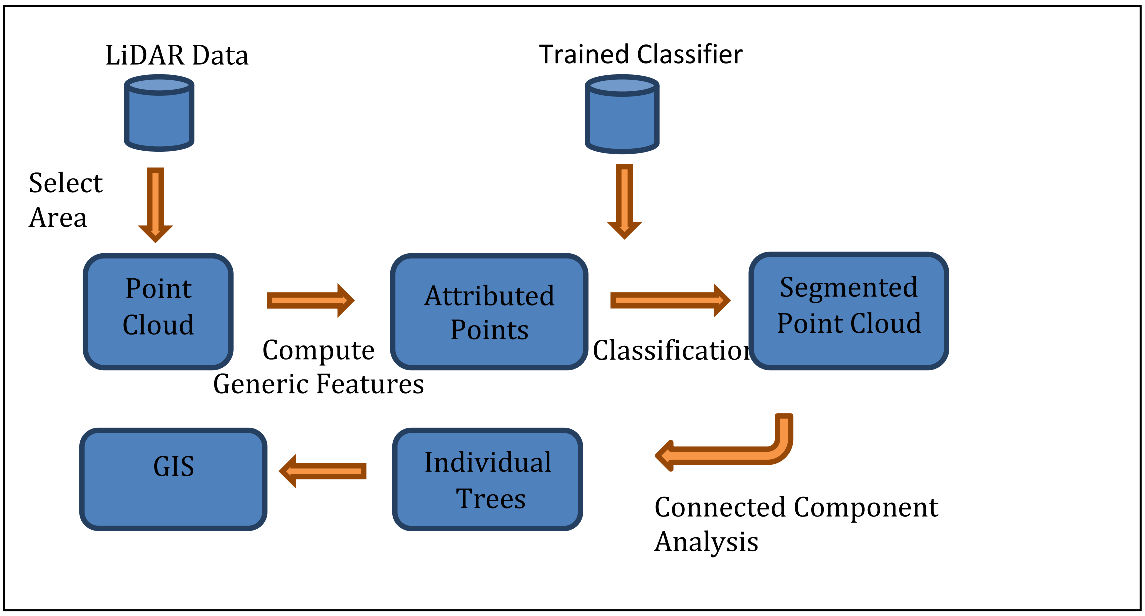


Figure 41. *US2 Workflow*

In short, the workflow consists of the following components:

- Selecting Area of Interest
- Computing generic features
- Classification
- Connected Component Analysis

The first step, selecting the area of interest, simply defines where the work will take place. Computing generic features is a step that takes the point cloud in the area of interest as input, and for each point determines local geometrical descriptors by analyzing points in a small neighborhood of the point at hand. Typically, local geometrical descriptors are derived from a Principal Component Analysis, which allows it to characterize the surface shape around a point in terms of e.g. its roughness and flatness and in addition returns a local surface normal.

Such geometrical features are the input for the Classification component, which classifies the points given their feature vector. The result of this step is a point cloud where each point is labelled by a class label, for example a class label "trees". In the final step, the connected component analysis, individual tree points are joined to form trees using e.g. a region growing procedure. The final results is a point cloud where each tree point has two additional labels, one identifying the tree it belongs to, the other a probability that this result is correct.

Generic system functionalities required:

Requirement name	***Functionality***
Setting Area of Interest	The user has to set the Region of Interest (ROI)

WP4 services required:

Service name	*Identifier according to workflow component*	*Identifier in eRoom WP4 service table*	*Status*
Multi-object classification of 3D point clouds (object type must be specified)), in this case: Extract all points sampling trees	US2_1	Service #59	Second year PM24
Individual Tree Identification (Decompose tree points as found in US2_1 into individual trees)	US2_2	New Service	Second year PM24

WP5 visualization techniques required:

The users will need to see the result of the classification step, selecting one area and viewing the segmented point cloud.

Requirement name	*Functionality*
Visualization of Individual tree component extraction	**Data to be visualized:** - List of identified trees - Associated Point Cloud parts **Output:** 3D visualization of Individual tree component extraction / segmented point cloud

Data provided for workflow development and testing

Name of data set	**Description**	**Source/Partner**
Mobile Mapping data Paris	Large LMMS data sets The Paris dataset will have the same data formats and specification as the following Toulouse dataset. Its production is planned for June 2014. Its volume and the number of successive acquisitions (e.g. for change detection applications) is subject to be extended, as required by the subsequent workflows or to demonstrate the scaling of the IQmulus framework.	IGN
Mobile Mapping data Toulouse	Medium LMMS data sets corresponding to a 10km acquisition in downtown Toulouse (~1 hour of acquisition) with IGN's STEREOPOLIS MMS. **1 012 160 278 Lidar points (83GB)** Acquired with a multi-return Riegl VQ-250 with the following attributes : - Absolute and relative measured 3D positions - Absolute and relative sensor 3D positions - GPS time - Multi-return echo number (i/N) - Measured range and angles - Measured quantities (amplitude, reflectance, deviation)	IGN

	Absolute positions are expressed in the Lambert93 system while relative positions are expressed relative to the mobile MMS frame system. **3603x5 images (227 GB, 75.5Gpix)** Each image node comprises a bundle of images acquired by a panoramic head with five 2k*2k cameras (63 MB per node). Images are delivered under RGB 8 bytes in uncompressed TIFF. Each image is accompanied with an xml file describing its intrinsic and extrinsic calibration.	

4 ANNEXES

4.1 SERVICES LIST FOR THE SHOWCASES

Integrated Land Showcase Workflow 1 (LS1): Terrain model preparation

Generic system functionalities (WP3 services):

Requirement name	***Functionality***
Import data	The user select and import data (point clouds and vector layer)
Definition of area of interest (Select AOI)	The user has to input an area of interests.
Area extraction	Specified area is extracted from the point cloud
Defining the level of detail	The user has to input the level of detail required

WP4 services required:

Component	Preparation
Name	**Vector layer partitioning of a point cloud**
Identifier	LS1_1_1
Description	Splits/order the available terrain data in areas defined by drainage boundaries
Input	Point cloud (LAS), Shapefile
Output	set of point clouds (LAS)
Responsible partner	IMATI
Needed Steering Parameters	
Status/Delivery Date	Service to be proposed (PM24)

Component	Preparation
Name	**Outlier Classification in Point Clouds (Service #10)**
Identifier	LS1_1_2
Description	Detect (classify) outlier points in point cloud
Input	Point cloud (LAS),
Output	Enriched clouds (LAS)
Responsible partner	UCL
Needed Steering Parameters	Threshold of distance to cluster centre as a multiple of standard deviation.
Status/Delivery Date	Released

Component	Preparation
Name	**Resampling of Point Clouds (Service #35)**
Identifier	LS1_1_3
Description	Resample an unordered point cloud to match a given point density
Input	Point cloud (LAS)
Output	Point cloud (LAS)
Responsible partner	UCL

Needed Steering Parameters	point density
Status/Delivery Date	Released

Component	Preparation
Name	**Sub-sampling (thinning) of Point Cloud (Service #36)**
Identifier	LS1_1_4
Description	Sub-sampling (thinning) of Point Cloud, preserving high-level features
Input	Point cloud (LAS)
Output	Point cloud (LAS)
Responsible partner	UCL
Needed Steering Parameters	point density, size of grid cells
Status/Delivery Date	Released

Component	Preparation
Name	**Constrained triangulation (Service #49)**
Identifier	LS1_1_5
Description	Create a triangulation of a point cloud preserving user-defined linear constraints (e.g. boundaries, feature lines).
Input	point cloud (LAS) lines, polygons (ShapeFile)
Output	triangulation (ply), polygons (ShapeFile)
Responsible partner	IMATI
Needed Steering Parameters	
Status/Delivery Date	Released

Component	Feature detection
Name	**Detection of flow lines and drainage basins from triangle meshes (Service #92)**
Identifier	LS1_2_1
Description	Detection of flow lines and drainage basins from triangle meshes
Input	triangulation (ply), polygons (ShapeFile)
Output	triangulation (ply), polygons (ShapeFile),
Responsible partner	IMATI
Needed Steering Parameters	size of the basins, degree of the network
Status/Delivery Date	Second year (PM24)

Component	Feature detection
Name	**Multi-resolution triangulation for hydrology (Service #48)**
Identifier	LS1_2_2
Description	Storage of the drainage basins hierarchy as a multi-res triangle mesh.
Input	triangulation (ply), polygons (ShapeFile)
Output	drainage mesh LOD (tbd)
Responsible partner	IMATI
Needed Steering Parameters	size of the basins, degree of the network

Status/Delivery Date	PM30

Component	Visualization
Name	**Extract the triangle mesh to be visualized at the given LOD**
Identifier	LS1_4_1
Description	Visualization of the terrain model at a user-selected level of detail, with colour coding of drainage basins.
Input	triangulation (ply)
Output	-
Responsible partner	IMATI
Needed Steering Parameters	level of detail, focus of the visualization
Status/Delivery Date	Service to be proposed (PM30)

WP5 visualization functionalities required:

Component	Preparation
Name	**Visualize Vector Layer partitioning of a Point Cloud**
Description	Visualization of the partitioned point clouds for the selected area/basin.
Input	Polygons (ShapeFile), set of point clouds (LAS)
Output	-
Responsible partner	SINTEF
Needed Steering Parameters	Selected area / basin
Status/Delivery Date	Year 2 Prototype

Component	Visualization
Name	**Visualize Triangle Mesh**
Description	Visualization of the partitioned triangle meshes for the selected area/basin with color coding / labelling with the selected LoD
Input	Polygons (ShapeFile), triangulation (ply)
Output	-
Responsible partner	SINTEF
Needed Steering Parameters	Selected area / basin
Status/Delivery Date	Year 2 Prototype

Integrated Land Showcase Workflow 2 (LS2): Precipitation analysis

Generic system functionalities (WP3 services):

Requirement name	***Functionality***
Selection of time interval	The user has to select the temporal interval for the analysis of the time series (this was the meaning of the service LS2_1_1 below)

WP4 services required:

Component	Preparation
Name	**Extraction of rain data**
Identifier	LS2_1_1
Description	Extract rain data for the time interval and prepare the cumulated rain values for the time interval required
Input	enriched point cloud (LAS), DEM (GeoTiff)
Output	Gridded point cloud (GeoTiff), enriched point cloud
Responsible partner	IMATI
Needed Steering Parameters	temporal scale for rain analysis (30min, 1h, 2h, 3h, 6h, 12h, 24h)
Status/Delivery Date	Service to be proposed (PM24)

Component	***Preparation***
Name	**Approximation rainfall data by Kriging (Service #40)**
Identifier	LS2_1_2
Description	Approximation of heterogeneous rainfall data (punctual gauge measurements or radar measures) with ordinary Kigring
Input	enriched point cloud (LAS), DEM (GeoTiff)
Output	Gridded point cloud (GeoTiff), enriched point cloud
Responsible partner	IMATI
Needed Steering Parameters	
Status/Delivery Date	Released (PM18)
Component	Preparation
Name	**Approximation of rainfall data by RBF (service #67)**
Identifier	LS2_1_3
Description	Approximation of heterogeneous rainfall data (punctual gauge measurements or radar measures) with radial basis functions
Input	enriched point cloud (LAS), DEM (GeoTiff)
Output	Gridded point cloud (GeoTiff), enriched point cloud
Responsible partner	IMATI
Needed Steering Parameters	
Status/Delivery Date	Released (PM18)

Component	Preparation
Name	**Approximation of rainfall data by LR-splines (Service #58)**
Identifier	LS2_1_4

Description	Approximation of heterogeneous rainfall data (punctual gauge measurements or radar measures) with LR-splines
Input	enriched point cloud (LAS), DEM (GeoTiff)
Output	Spline surface (g2 or iges)
Responsible partner	SINTEF
Needed Steering Parameters	
Status/Delivery Date	Year 2 (PM18)

Component	Feature detection
Name	**Critical Points (Service #44)**
Identifier	LS2_2_1
Description	Extraction of critical points (isolated and degenerate) from grids and triangulations.
Input	triangle mesh (PLY) or gridded point cloud (GeoTIFF)
Output	three Vector Data (Shapefile), one for each dimension of criticalities
Responsible partner	IMATI
Needed Steering Parameters	threshold for critical point filtering
Status/Delivery Date	Released

Component	Feature extraction
Name	**Tracking of critical points**
Identifier	LS2_2_2
Description	Given two rain field, with critical points extracted, match the corresponding critical points. It assumes that the two precipitation field are very close in time.
Input	triangle mesh (PLY)
Output	Shapefile (path of the critical points)
Responsible partner	IMATI
Needed Steering Parameters	Threshold (tbd)
Status/Delivery Date	Service to be proposed (PM24)

WP5 visualization functionalities required:

Component	Preparation
Name	**Visualize precipitation field over the multi-resolution model**
Description	3D visualization of approximated rain and the accuracy of the approximation
Input	Triangulation with rain data and approximation accuracy per vertex (ply)
Output	-
Responsible partner	SINTEF
Needed Steering Parameters	-
Status/Delivery Date	Year 2 Prototype

Component	Visualization
Name	**Visualize the movement of the maxima of precipitation over the**

	multi-resolution model
Description	(animated/interactive) 3D visualization of rain fields and critical points (maxima)
Input	Triangulation with rain data and approximation accuracy per vertex (ply) for each timestep, calculated path of maxima (ShapeFile)
Output	-
Responsible partner	SINTEF
Needed Steering Parameters	-
Status/Delivery Date	Year 2 Prototype

Integrated Land Showcase Workflow 3 (LS3): Flood and waterlogging detection

Generic system functionalities (WP3 services):

Requirement name	***Functionality***
Import data	The user select and import data (satellite image)
Definition of area of interest (Select AOI)	The user has to input an area of interests.
Area extraction	Specified area is extracted from the images (performed by applying a mask to the image, no image transformation is required)
Specify reference data	The user select and import reference data (e.g. classified image) for classification
Specify RGB colours	The user select and import a pre-prepared RGB colour table (colour table will be used at the visualization process)

WP4 services required:

Component	Preparation
Name	**Computation of Top Of the Atmosphere reflectance (Service #73)**
Identifier	LS3_1_1
Description	Part of the service "Preprocessing of raster data" eRoom (Service #73)
Input	Masked satellite image
Output	Raster image with ToA reflectance calculation
Responsible partner	FÖMI
Needed Steering Parameters	
Status/Delivery Date	Year 2 (PM22)

Component	Preparation
Name	**Computation of spectral indices (Service #76)**
Identifier	LS3_1_2
Description	Computation of spectral indices including: NDVI, NDSI, NDWI. NDxI=(NDVI, NDSI, NDWI)
Input	Raster image with ToA reflectance calculation
Output	Raster image with spectral indices
Responsible partner	FÖMI
Needed Steering Parameters	
Status/Delivery Date	Year 2 (PM19)

Component	Feature detection
Name	**Raster segmentation (Service #77)**
Identifier	LS3_2_1
Description	Raster segmentation, performed using a predefined or specified algorithm
Input	Raster image with spectral indices
Output	Segmented raster image
Responsible partner	FÖMI

Needed Steering Parameters	
Status/Delivery Date	Year2 (PM22)

Component	Feature detection
Name	**Clustering of raster (Service #78)**
Identifier	LS3_2_2
Description	Generating clustered raster image based on pre-defined algorithm
Input	Segmented raster image
Output	Clustered raster image
Responsible partner	FÖMI
Needed Steering Parameters	
Status/Delivery Date	Year 2 (PM22)

Component	Feature detection
Name	**Thematic classification of raster data (Service #79)**
Identifier	LS3_2_3
Description	Match regions of the images to preselected categories to produce water categories
Input	Clustered raster image, masked image with spectral indices, reference classification data
Output	Classified multipolygons (vector data) classes = water categories
Responsible partner	FÖMI
Needed Steering Parameters	
Status/Delivery Date	Year 2 (PM22)

WP5 visualization functionalities required:

Component	Visualization
Name	**Visualize classification results**
Description	False colour representation of the results using the specified category-colour mapping (RGB colour table is imported previously)
Input	-initial raster image (GeoTiff), extracted classification/categories (ShapeFile)
Output	-
Responsible partner	IGD
Needed Steering Parameters	-
Status/Delivery Date	Year 2 Prototype

Integrated Land Showcase Workflow 4 (LS4): Detection and characterization of landslides

Generic system functionalities (WP3 services):

Requirement name	***Functionality***
Import data	The user select and import data (satellite image)
Definition of area of interest (Select AOI)	The user has to input an area of interests.
Area extraction	Specified area is extracted from the images (performed by applying a mask to the image, no image transformation is required)
Storage of resulting rasters	Store the resampled raster as GeoTIFF along with the original data. Temporal storage is mandatory and permanent storage will be possible if desired by the user.

WP4 services required:

Component	Preparation
Name	**Raster Resampling (Service #93)**
Identifier	LS4_1_1
Description	A sinc kernel interpolation is used to harmonize the resolution of the input images.
Input	Single band GeoTIFF images with inhomogeneous resolution
Output	Single band GeoTIFF images with homogenous resolution
Responsible partner	UBO
Needed Steering Parameters	None, optional: target resolution
Status/Delivery Date	Year 2, PM 24

Component	Feature Detection
Name	**Computation of spectral indices (Service #76)**
Identifier	LS4_2_1
Description	Band math to compute the NDVI
Input	Red and near-infrared band of a multispectral GeoTIFF image
Output	Single band GeoTIFF representing the NDVI
Responsible partner	FOMI
Needed Steering Parameters	None
Status/Delivery Date	Year 2 (PM19)

Component	Feature Detection
Name	**Raster thresholding (Service #69)**
Identifier	LS4_2_2
Description	Generate a binary mask marking all areas where the NDVI is above a certain threshold OR the brightness is below a certain threshold
Input	NDVI GeoTIFF image and panchromatic GeoTIFF image
Output	Binary GeoTIFF image
Responsible partner	FOMI
Needed Steering Parameters	None
Status/Delivery Date	Year 2 (PM19)

Component	Feature Detection
Name	**Convolution filtering (Service #71)**
Identifier	LS4_2_3
Description	Combination of morphological closing and opening to eliminate isolated pixels
Input	Binary GeoTIFF image
Output	Binary GeoTIFF image
Responsible partner	FOMI
Needed Steering Parameters	None
Status/Delivery Date	Year 2 (PM22)

Component	Change detection
Name	**Measuring surface displacement (Service #60)**
Identifier	LS4_3_1
Description	Hierarchical cross-correlation is used to find homologous points in two images
Input	Multiple GeoTIFF raster with homogenous resolution and extent
Output	3 GeoTIFFs (x-component, y-component, correlation coefficient)
Responsible partner	UBO
Needed Steering Parameters	Choose pre-defined parameter set for landslide displacement measurements or optional: window size, correlation threshold, regularization strength, etc.
Status/Delivery Date	Released

WP5 visualization functionalities required:

Component	Visualization
Name	**Visualization of displacement fields**
Description	(Combined) visualization of the calculated displacement field and the correlation coefficient.
Input	3 GeoTIFFs (x-component, y-component, correlation coefficient)
Output	-
Responsible partner	IGD
Needed Steering Parameters	
Delivery Date	Year 2 Prototype

Integrated Land Showcase Workflow 5 (LS5): Simulation of floods with observed rain data

Generic system functionalities (WP2/WP3 services):

Requirement name	***Functionality***
Import data	The user select and import data (satellite image)
Definition of area of interest (Select AOI)	The user has to input an area of interests.
Area extraction	Specified area is extracted from the images (performed by applying a mask to the image, no image transformation is required)
Storage of resulting rasters and graphics	

WP4 services required:

Component	Preparation
Name	**Compute cumulated precipitation over multi-resolution for hydrology**
Identifier	LS5_1_1
Description	The quantity of cumulated observed precipitation is computed for the time interval and area selected
Input	multi-resolution for hydrology (format tbd)
Output	multi-resolution for hydrology with attributes (quantity of water), arrays and vector data (.mat)
Responsible partner	IMATI
Needed Steering Parameters	time interval within which to compute cumulated precipitation; basin of interest
Delivery Date	Service to be proposed (Year 3, PM30)

Component	Preparation
Name	**Convert multi-resolution for hydrology to regular grid**
Identifier	MS3_2_1
Description	The simulation model runs on a regular grid, the conversion prepares the terrain model so that it can be fed to the simulation
Input	multi-resolution for hydrology (format tbd)
Output	GeoTIFF

Responsible partner	IMATI
Needed Steering Parameters	grid size
Delivery Date	Service to be proposed (Year 3, PM30)

Component	Preparation
Name	**Convert cumulated precipitation to regular grid**
Identifier	LS5_1_3
Description	The cumulated observed precipitation, computed over the multi-resolution for hydrology, is mapped onto the regular grid, so that it can be giveb as inout to the simulation
Input	multi-resolution for hydrology with attributes (quantity of water), arrays and vector data (.mat)
Output	geoTIFF
Responsible partner	IMATI
Needed Steering Parameters	
Delivery Date	Released

WP5 visualization functionalities required:

Component	Preparation
Name	**Simulation of flood over terrain**
Description	The input parameters generated by LS5_1_3 are used to run a simulation using existing simulator developed by SINTEF
Input	GeoTIFF hydrology with attributes (quantity of water)
Output	NetCDF
Responsible partner	SINTEF
Needed Steering Parameters	
Delivery Date	Year 3

Component	Visualization
Name	**Visualization of flood over terrain**
Description	The simulation results are visualized in an interactive manner for user inspection.
Input	NetCDF or direct output from “Simulation of flood over terrain”
Output	
Responsible partner	SINTEF
Needed Steering Parameters	
Delivery Date	Year 3

Marine Showcase Workflow 1 (MS1): Generating elevation model from point cloud data

Generic system functionalities (WP2/WP3 services):

Requirement name	***Functionality***
Import data	The user select and import data
Definition of area of interest (Select AOI)	The user has to input an area of interests.

WP4 services required:

Component	Preparation
Name	**Point cloud to LR-spline surface (Service #9)**
Identifier	MS1_1_1
Description	Generates an LR-Spline surface from a point cloud, where the point cloud is created by just taking one data set (the most recent one for a start) and later as the result of deconfliction of several sets
Input	Point cloud(s) over a whole data set or a specified tile (and its neighbours) Boundary information for instance from an already created surface on an adjacent tile, if this exists
Output	LR-Spline surface
Responsible partner	SINTEF
Needed Steering Parameters	uncertainty measure related to the point cloud(s), surface tolerance, minimal polynomial cell width, more
Delivery Date	Released, some update needed
Notes	Service #9 (must be updated with information about minimal cell width) The service works on one tile or a reduced point cloud. This preprocessing is expected to be done. Service #9 may be used to create an intial surface for a reduced point cloud over the full area to be updated tile by tile by service #57.

WP5 visualization functionalities required:

Component	Visualization
Name	**Visualization of an LR-spline surface**
Description	3D visualization of the generated LR-spline surface
Input	LR-Spline surface
Output	-
Responsible partner	SINTEF
Needed Steering Parameters	-
Status/Delivery Date	Year 2 Prototype

Marine Showcase Workflow 2 (MS2): Error checking - distance of survey data from a surface

Generic system functionalities (WP2/WP3 services):

Requirement name	***Functionality***
Import data	The user select and imports LR-spline surface and the surveys to compare to
Definition of area of interest (Select AOI)	The user has to input an area of interests.

WP4 services required:

Component	Preparation
Name	**Compare point cloud to LR-spline surface (service #86)**
Identifier	MS2_2_1
Description	Computes the distance between a point cloud and an LR-spline surface
Input	Point cloud and LR-spline surface . The point cloud and the surface must correspond in the respect that they represent the same area.
Output	Set of outliers in the point cloud outside the LR-spline surface
Responsible partner	SINTEF
Needed Steering Parameters	Threshold for points considered as lying outside the surface, more than one level can be given to group points according to their distance to the surface more
Delivery Date	Year 2

WP5 visualization functionalities required:

Component	Visualization
Name	**Outlier visualization**
Description	Interactive 3D visualization of the points from the point cloud lying outside a user-given threshold of the LR-spline surface (from MS2_2_1).
Input	LR-Spline surface, point cloud with computed distances (LAS)
Output	-
Responsible partner	SINTEF
Needed Steering Parameters	-
Status/Delivery Date	Year 2 Prototype

Marine Showcase Workflow 3 (MS3): Simple feature extraction

Generic system functionalities (WP2/WP3 services): none

WP4 services required:

Component	Preparation
Name	**Select Feature criteria**
Identifier	MS3_1_1
Description	Choose the criteria to identify a feature or a set of features
Input	LR-spline surface
Output	LR-spline surface with feature criteria
Responsible partner	SINTEF

Needed Steering Parameters	-
Delivery Date	Service to be proposed (Year 2)

Component	Feature detection
Name	**Detection of features by given criteria**
Identifier	MS3_2_1
Description	Detection of features in the LR-spline surface
Input	LR-Spline surface, feature criteria
Output	List of the various features detected in the LR-spline surface
Responsible partner	SINTEF
Needed Steering Parameters	Feature parameters: extent, minimum of points, etc.
Delivery Date	Service to be proposed (Year 2)

WP5 visualization functionalities required:

Component	Visualization
Name	**Feature visualization**
Description	Visualization of the features detected in an LR-spline surface according to given feature criteria (from MS3_2_1).
Input	LR-Spline surface, detected features (ShapeFile)
Output	-
Responsible partner	SINTEF
Needed Steering Parameters	-
Status/Delivery Date	Year 2 Prototype

Marine Showcase Workflow 4 (MS4): Measuring submarine dune migration

Generic system functionalities (WP2/WP3 services):

Storage of resulting raster	Store the resampled raster as GeoTIFF along with the original data. Temporal storage is mandatory and permanent storage will be possible if desired by the user.

WP4 services required:

Component	Preparation
Name	**Approximate surface from bathymetric point cloud**
Identifier	MS4_1_1
Description	LR-splines is used to generate a gridded raster representing the seafloor from bathymetric point clouds.
Input	Bathymetric point cloud (LAS, SHP))
Output	Gridded surface in GeoTIFF format
Responsible partner	SINTEF
Needed Steering Parameters	uncertainty measure related to the point cloud(s), surface tolerance, minimal polynomial cell width, resolution of the output raster
Delivery Date	Year 2
Notes (WP4 Services used)	Service #9, service #91

Component	Preparation
Name	**Raster Resampling (Service #93)**
Identifier	MS4_1_2
Description	A sinc kernel interpolation is used to harmonize the resolution of the input images.
Input	Single band GeoTIFF images with inhomogeneous resolution
Output	Single band GeoTIFF images with homogenous resolution
Responsible partner	UBO
Needed Steering Parameters	None, optional: target resolution
Status/Delivery Date	(Year 2, PM 24)

Component	Change detection
Name	**Measuring surface displacement (Service #60)**
Identifier	M4_3_1
Description	Hierarchical cross-correlation is used to find homologous points in two images
Input	Multiple GeoTIFF raster with homogenous resolution and extent
Output	3 GeoTIFFs (x-component, y-component, correlation coefficient)
Responsible partner	UBO
Needed Steering Parameters	Choose parameter set for dune migration or optional: window size, correlation threshold, regularization strength, etc.
Delivery Date	Released

WP5 visualization functionalities required:

Component	Preparation
Name	**Visualization of raster images**
Description	Visualization of the resulting surface raster as greyscale images
Input	Raster image of resampled spline surface (GeoTIFF)
Output	-
Responsible partner	IGD
Needed Steering Parameters	-
Status/Delivery Date	Year 2 Prototype

Component	Visualization
Name	**Visualization of displacement fields**
Description	(Combined) visualization of the calculated displacement field and the correlation coefficient.
Input	3 GeoTIFFs (x-component, y-component, correlation coefficient)
Output	-
Responsible partner	IGD
Needed Steering Parameters	-
Delivery Date	Year 2 Prototype

Urban Showcase Workflow 1 (US1): Detection of buildings for monitoring and cadastral mapping

Generic system functionalities (WP3 services):

Requirement name	***Functionality***
AOI selection	Selection of an area of interest for further processing

Component Description	Preparation
Name	**Setting Area of Interest**
Identifier	US1_1_1
Description	The user has to set the Region of Interest (ROI)
Input	New ALS data set (LAS?)
Output	New ALS data set, constrained to ROI (LAS?)
Responsible Partner	MOSS
WP4 Service	Spatial Query (probably a generic service from WP2/3?)
Remarks:	

WP4 services required:

Component Description	Feature Extraction
Name	**Building identification**
Identifier	US1_2_1
Description	Identify ALS data that is sampling buildings
Input	ALS data set, constrained to ROI
Output	Annotated LAS file
Responsible Partner	UCL
Parameters:	
WP 4 Service	#59, Multi-object classification of 3D Point Clouds, Proposed
Remarks:	

Component Description	Feature Extraction
Name	**2D Building Outline Generation**
Identifier	US1_2_2
Description	Use commercial eCognition software to generate 2D outlines for the ROI
Input	Orthoimagery + DSM, constrained to ROI
Output	Vector data set
Responsible Partner	FÖMI
Parameters:	2D spatial resolution, e.g. 0.5 m
WP 4 Service	Not yet identified, one possibility would be to extract building blobs from the classification results(new service) and consecutively get the outline of each blob with Service #11, Spatial Extent, which works both in 2D and 3D
Remarks:	

Component Description	Feature Extraction
Name	**3D Building Outline Generation**
Identifier	US1_2_3
Description	Use commercial Tridicon software to generate 3D outlines for the ROI
Input	New ALS data set, constrained to ROI

Output	Vector data set
Responsible Partner	MOSS
Parameters:	3D Spatial Resolution
WP 4 Service	Not yet identified
Remarks:	Data set is not clear yet

Component Description	Change Detection 2D
Name	**2D Building Outline Comparison**
Identifier	US1_3_1
Description	• Project 3D building outlines (from US1_2_1 or US1_2_2) on ground surface to obtain 2D building outlines • Compare result to existing 2D building outlines, and distinguish between a) no change, b) growing or new building, and c) shrinking or disappeared building
Input	• New 3D building outlines (from US1_2_1 or US1_2_2) • Existing 2D building outlines
Output	Vector data set (+ quality of result)
Responsible Partner	MOSS
Parameters:	3D spatial resolution
WP 4 Service	Not yet identified
Remarks:	Dataset is not defined yet

Component Description	Change Detection 3D
Name	**2D/3D Building Outline Comparison**
Identifier	US1_3_2
Description	New 3D building outlines (from US1_2_1 or US1_2_2) are compared to existing 3D building outlines. For the eCognition process this takes place based on 2D data.
Input	2 times 3D Vector data
Output	Vector data set (+ quality of result)
Responsible Partner	MOSS
Parameters:	3D spatial resolution
WP 4 Service	Not yet identified
Remarks:	Data set is not clear yet

WP5 visualization functionalities required:

Component	Visualization
Name	**2D/3D Building Outline Comparison**
Description	3D visualization of base data and detected changes (2D/3D)
Input	Cadastral records (ShapeFile), detected changed building outlines (ShapeFile)
Output	-
Responsible partner	IGD
Needed Steering Parameters	-
Delivery Date	Year 2 Prototype

Urban Showcase Workflow 2 (US2): Individual tree extraction from urban LMMS data

Generic system functionalities (WP3 services):

Component Description	Preparation
Name	**Setting Area of Interest**
Identifier	US2_1
Description	The user has to set the Region of Interest (ROI)
Input	LMMS data set
Output	New LMMS data set, constrained to ROI
Responsible Partner	MOSS
WP4 Service	Spatial Query (not WP4?)
Remarks:	

WP4 services required:

Component Description	Feature Extraction
Name	**Tree classification**
Identifier	US2_1_1
Description	Extract all points sampling trees, using WP4, Service #59, General Feature Extraction
Input	LMMS data set, constrained to ROI
Output	Annotated LMMS data set
Responsible Partner	UCL (IGN, TU Delft)
Parameters:	Per pixel quality of classification
WP 4 Service	#59, proposed
Remarks:	

Component Description	Feature Extraction
Name	**Individual Tree Identification**
Identifier	US2_1_2
Description	Decompose tree points as found in U2Fa into individual trees
Input	Annotated LMMS data set, with label indicating tree class points
Output	• Annotated LMMS data set, with labels indicating individual trees • List of locations of individual trees
Responsible Partner	IGN or TU Delft (UCL)
Parameters:	Decomposition quality
WP 4 Service	New service: individual component extraction
Remarks:	

WP5 visualization functionalities required:

Component	Visualization
Name	**Visualization of Individual tree component extraction**
Description	3D visualization of Individual tree component extraction / segmented point cloud
Input	List of identified trees, associated point clouds (ply)
Output	-
Responsible partner	IGD
Needed Steering Parameters	-
Delivery Date	Year 2 Prototype