FINALIZED USER REQUIREMENTS

Deliverable D1.2.4

Circulation:	PU: Public
Lead partner:	FÖMI
Contributing partners:	All
Authors:	Mónika Antal, Touria Bajjouk, Márta Belényesi, Beril Sirmacek, Jan Boehm, Mathieu Brédif, Romain Cancouët, Alexander Cicimol, Christophe Delacourt, Tor Dokken, Roberto Giachetta, Attila Képes, Iván Kovács, Dániel Kristóf, Róbert Lehoczki, Roderik Lindenbergh, Gábor Mikus, Gizella Nádor, Thai Binh Nguyen, Angéla Olasz, Giuseppe Patanè, Simone Pittaluga, Jean-Yves Royer, Vibeke Skytt, Ottó Tancsik, Mickael Tulldahl, Michela Spagnuolo, Judit Szabó
Quality controller:	Anna Cristofol
Version:	1.0
Date:	19.11.2015

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	November 19, 2016	Final	Version submitted to the Project Officer

[1]Integers correspond to submitted versions

EXECUTIVE SUMMARY

This document summarizes the final analysis of user requirements, which was carried out taking into account the experience gathered during the third project year.

The previous WP1 deliverable on *Revised User Requirements* (D1.2.3) documented the procedure and results of filtering, categorisation and prioritisation of requirements consolidated in previous stages of the IQmulus project (published in D1.1[2] and D1.2.1[3] and D1.2.2[3]). So called 'Showcases' were defined (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. Later each Showcase was revised and decomposed into workflows, for which – beside As Is Analyses – user perspectives, expectations and the potential for innovation was provided.

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 third project year exploitation aspects (WP9) – as an important task for the fourth project year – were also taken into account during the final revision of the requirements.

For the beginning of the third project year the first system prototype of the IQmulus system was set, and its evaluation procedure has started (PM 24-30). The prototype already implements some key workflows previously identified by the IQmulus partners (according to the pre-defined Showcases in D1.2.3). Based on the developer's experiences, the feedback of the users and the testers (results of the evaluation) and taking into account the exploitation strategy, in Year 3 we arrived to the point when a final set of requirements had to be defined. Present document describes the methodology and the results of this procedure.

As a result of the on-going requirement consolidation process Showcases elaborated in the previous project years and described in D1.2.3 were all revised again. According to the experiences in the third project year, the workflows and the corresponding services (requirements) were re-prioritized, and their implementation status was clearly defined, providing help for keeping further development and refinement work, and dissemination and exploitation possibilities in focus this way. Thus, efforts from workflows which were strongly down-prioritized (three of the original twelve) will be moved to the ones which are kept with high priority.

Higher levels of details are provided in the document concerning the above mentioned requirements, giving short summary of the earlier proposed IQmulus solutions, and the final outcomes as well: requirement prioritization and their implementation status. For the shake of clarity, services already implemented, and will be most likely implemented by April 2016 for the Showcases are summarised in tables as Annexes also.

[2]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
[3] 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
[3] D1.2.2 "Consolidated User Requirements" Link on IQmulus website:
https://project.sintef.no/eRoom/math/IQmulus/0_2d41f

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 (Revised User Requirements, D.1.2.3) documented the filtering, categorisation and prioritisation of requirements consolidated in previous stages of the IQmulus project (published in D1.1[4] and D1.2.1[5] and D1.2.2[6]). Subsets of user requirements, so called Showcases were also defined in previous stage of the project to drive the early development of the infrastructure, prototypes and basic services in the first phase of the project. These showcases were also reviewed in the 2nd project year and were reported in D1.2.3.

The evaluation procedure of the first system prototype of the IQmulus system lasted 6 month (PM 24-30). The prototype already implements some key workflows previously identified by the IQmulus partners (according to the pre-defined Showcases – see D1.2.3). The results of the actual evaluation period were documented in D7.3.2.

As the development is an iterated procedure, based on the developer's experiences and the feedback of the users and the testers (results of the evaluation), in Year 3 we arrived to the point when a final set of requirements has to be defined. Present document describes the methodology and the results of this procedure.

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
- D.1.2.3 Revised User Requirements
  https://project.sintef.no/eRoomReq/Files/math/IQmulus/0_35c0f/D1.2.3_Revised_User_Requirements.pdf
- D7.3.2 Evaluation of first system prototype.
- https://project.sintef.no/eRoomReq/Files/math/IQmulus/0_3fe62/D7.3.2_Evaluation_of_first_system_prototype_final.pdf
- D9.1.3 Exploitation Plan
  https://project.sintef.no/eRoomReq/Files/math/IQmulus/0_40828/IQmulus_D9_1_3_Final.pdf
- D4.1.3 Amended/Updated Development Guidelines For Data integration and Processing – Version 3
  https://project.sintef.no/eRoomReq/Files/math/IQmulus/0_37096/D4.1.3-UPDATED-AMENDED-VERSION-FINAL.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
[6] D1.2.2 "Consolidated User Requirements" Link on IQmulus website:
https://project.sintef.no/eRoom/math/IQmulus/0_2d41f

2 METHODOLOGY

This document summarizes the final analysis of user requirements, which was carried out taking into account the experience gathered during the third project year. 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.

2.1 ASPECTS TAKEN INTO ACCOUNT DURING FINALIZATION

We defined four pillars final user requirements should be based on:

1. ***D1.2.3 Revised User requirements document:*** in the document User perspective and expectations as well as innovation aspects were summarized. As Is Analysis were created for each workflow and against this background proposed IQmulus solutions were set up by the internal User community. List of services necessary for implementation of proposed solutions was provided for each workflow by the Users (revised user requirements). By now – according to the developments and test results – the scope of services already implemented, and the scope of new services feasible during the fourth project year (until April, 2016) is clearly visible, so it is possible now to finalize the requirements. It means checking and filtering the service list in D1.2.3, and adding new services developed in the third project year to the list.
D1.2.3 document in eRoom: https://project.sintef.no/eRoom/math/IQmulus/0_2d422

2. ***Developer's feedback on feasibility concerning further services*** has to be taken into account. All internal users were asked to consult with their developers responsible for services of the workflows defined.

3. ***Conclusions drawn from test results (D7.3.2 Assessment of first system prototype)*** were used also as an additional aspect for finalizing the user requirements. Two main aspects were tested in the latest testing process: *functionality aspects* (test if the system is providing the functionalities written in user requirements documentation) and *usability aspects* (if the system is usable and how intuitive to use it). As a part of testing the first system prototype, test stories were defined as basic units for testing functions from the user side, and were categorized by their activity (Story) group, namely authentication, data management, GIS processing, and account management.
The D7.3.2 can be downloaded from:
https://project.sintef.no/eRoom/math/IQmulus/0_2d497

4. The results of discussions and analysis of each service developed or planned so far in ***D4.1.3 Amended/Updated development guidelines for data integration and processing (version 3)***
https://project.sintef.no/eRoomReq/Files/math/IQmulus/0_37096/D4.1.3-UPDATED-AMENDED-VERSION-FINAL.pdf

5. Data and information stored and managed in ***Artifactory*** repository manager and in ***Redmine*** project management system.

6. Last, but not least ***D9.1.3 Exploitation document*** was also used. Exploitation is an important aspect of the project and activities already started early in the first project year. The document among others summarizes the exploitable results, status and targeted user groups as well. Exploitation of the project outcomes commenced as a device

to help focus and shape the research to be accomplished within the project, so we think it is important to pay attention to it.
https://project.sintef.no/eRoom/math/IQmulus/0_2d4eb

2.2 INPUTS PROVIDED BY PARTNERS

Inputs provided by participants to this document are the next:

- experiences collected and conclusions drawn in the third project year regarding the workflow-development procedure, and
- the finalized list of services – in line with the conclusions drawn – for each workflow, with all textual descriptions necessary for better understanding.

To make inputs consistent, a pre-structured table was provided to the project partners responsible for each workflow. In the table the partners had to summarize all the services planned to be most likely implemented and integrated to the system. The result will be presented in the Annex. Explanations and reasons of changes occurred – compared to the D1.2.3 – were provided as textual descriptions by the partners.

2.3 DOCUMENTATION

The present Deliverable 1.2.4 contains the following results:

- The list of non workflow-specific general system requirements (prioritized);
- Updated list and description of Showcases and workflows,
- Updated list of services for each workflow (prioritized).

3 RESULTS

In this chapter we refer the User perspective and expectations, innovation aspects formulated in D1.2.3 according to the As Is Analysis, before we define the final list of user requirements. The proposed IQmulus solutions defined in D1.2.3 are also summarized, according to the preparation, feature detection, change detection and presentation/ visualization component of the workflows.

Assessment and evaluation phase is of major importance as it confirms whether requirements defined at an early stage of the project have been fulfilled by developments and integration. Therefore experiences formulated in D7.3.2 about overall (not workflow-specific) system requirements are presented in a separated chapter.

The main part of the chapter includes the workflow-specific final user requirements in the light of experiences collected in the third project year.

3.1 NON WORKFLOW-SPECIFIC GENERAL SYSTEM REQUIREMENTS

The effect of evaluation work package to the user requirements.

As a part of testing the first system prototype, test stories were defined as basic units for testing functions from the user side, and were categorized by their activity (Story) group, namely authentication, data management, GIS processing, and account management (D7.3.2). These test stories are generally formulated requirements on the system, and could be relevant for each workflow regardless of the workflow-purpose.

According to the test results documented in D7.3.2, table 1 summarizes these test stories which can be regarded as requirements, and shows the priority level defined by the users for each requirement. We concentrate on the “Data management” and “GIS processing” component of the IQmulus system.

Story name	***Long term priority***	***Short term priority***
Data Management Test Stories		
Data upload	essential	High
Data upload (Batch)	essential	High
Preview uploaded data	essential	Medium
Metadata upload	essential	High
Metadata upload (Batch)	essential	High
Editing existing metadata	essential	High
Metadata search (by attribute)	essential	High
Geographical search (by location)	essential	Medium
Dataset management (remove, select data from the list)	essential	High
Deleting data/metadata	essential	High
Metadata search for output data (processed results - by attribute)	essential	Medium
Geographical search for output data (processed results - by location)	essential	Medium
Output data (processed result) search	recommended	High
Output data (processed result) download	essential	High
User information	strongly	High

	recommended	
Help	strongly recommended	High
GIS processing Stories		
Create new workflow	essential	High
Select data for workflow	essential	High
Delete workflow	essential	High
Create simple workflow with one service in it	essential	High
Create simple workflow containing multiple services	essential	High
Create complex workflow with multiple data	essential	Medium
Set parameters for a workflow	essential	High
Edit existing workflow, change service	essential	High
Edit existing workflow, change dataset	essential	High
Edit existing workflow, change parameters	essential	High
Select workflows	essential	High
Reading CueCards	strongly recommended	Medium
Searching existing actions	essential	Medium
Run workflows	essential	High
Name workflow	essential	High
Cancel workflow	essential	High
Check workflow status	essential	High
Resume workflow	recommended	Low
Preview results	essential	High
Help	strongly recommended	High
User information	strongly recommended	High
Visualization		
Visualization of uploaded data, results of analysis, and legend	essential	High
Visualization of results of analysis, and legend	essential	High

Table 1. *Non workflow-specific general system requirements*

3.2 SHOWCASES

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, and User Stories related to them were presented in D1.2.3. In present document we provide just a short description as a reminder. Any developments occurred since the last project year will be indicated.

The Integrated Land Showcase (1.2.2_SC2)

As defined in D1.2.3, the showcase exemplifies the IQmulus approach to big and heterogeneous data processing for expert users supporting decision makers in civil protection in events related to flooding. The partners contributing to the showcase, from a user perspective, are mainly Regione Liguria, CNR-IMATI, FÖMI, and UBO. The corresponding usage scenario is the following:

> *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.

With respect to big data issues, the goals of the scenario are:

- to develop an *intelligent* approach to indexing large Lidar point clouds, and organize them in a multi-resolution fashion so that terrain models of a suitable size and resolution can be extracted and used to perform analysis on drainage basins, using high resolution (and therefore accuracy) only in user-selected areas;
- to provide a higher level of automation of flood and waterlogging detection via smarter algorithms, based on processing of multispectral imagery. Solution improves overall processing time and implies a better use of human resources.

Compared to the workflow and service list documented in D1.2.3, changes effecting the final user requirements have been made in the Land Showcase, according to the third year experiences:

- (LS1) Organization of the terrain data according to their membership to main alert area (terrain model preparation) [CNR-IMATI];
- (LS2) Analysis and modelling of the observed rain (precipitation analysis) [CNR-IMATI];
- (LS3) Detection and characterization of flood and waterlogging [FÖMI];
- (LS4) Detection and characterization of landslides [UBO];

Regione Liguria is the primary end user involved in the showcase, meaning that the as-is analysis was conducted on their current practices and their requirements guided the development of the showcase. With respect to the planned set of workflows, priority has been given to LS1-LS4.

Originally, in D1.2.3 a fifth workflow (LS5: Simulation of floods with observed rain data) was also presented, but it was decided not to proceed further with this workflow, due to lack of efforts to devote to its development.

The Marine Showcase (1.2.2_SC1)

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 ideal 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 ideal 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. The first result will be a spline representation where parts of the initial point clouds that are not well represented by the spline are identified. This representation is a good starting point towards a final result which will be a hybrid representation of point clouds, triangulations and splines, where the remaining point sets may be 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). Following the advises from the second year review with respect to prioritization, the focus of the Marine Scenario has been put on handling large data sets by tiling and stitching and deconfliction. This implies that the final hybrid representation is not expected to be realized within IQmulus.

Compared to the workflow and service list documented in D1.2.3, changes effecting the final user requirements have been made in the Marine Showcase, according to the third year experiences and advices from the second year review:

- MS3 workflow has been down prioritized and will not be implemented.
- A new workflow (MS5) is formulated, which demonstrates Lidar WaveForm coastal Feature Extraction. The new workflow is related to eRoom User stories 39 to 43 (see later).

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. The work flow includes a deconfliction step
- (MS2) Inspection of the quality of the representation. The work flow is tightly integrated with MS1 and can serve as a starting point for a more interactive deconfliction process than the one envisaged in MS1
- (MS4) Change detection by comparing DEMs from different dates.
- ***(MS5) is a NEW workflow.*** It demonstrates Lidar WaveForm coastal Feature Extraction. The workflow has a major role in production of a seafloor 2D map with labels corresponding to identified sea bottom types using Bathymetric Lidar Full WaveForm Data.

MS3 workflow has been down prioritized and will not be implemented. The workflow originally was about 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.

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 main advantage of the marine DEM generation method offered by the Showcase is the high accuracy control and a non-uniform, lean representation format. It can reduce the time taken to generate a surface. These are steps towards interactive and on-demand surface generation in response to user inputs.

The Urban Showcase (1.2.2_SC3)

For the Urban Showcase setting, two workflows were proposed in D1.2.3 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.

The first workflow (US1) concerned a frequent challenge of mapping authorities: the update of cadastral datasets with 2D and 3D building objects based on point clouds. Integration of third-party software components (such as eCognition or Tridicon) from the current operational environments of Mapping and Cadastral Agencies was introduced in D1.2.3, and gave the opportunity to investigate how commercial software from the current "ecosystem" of mapping authorities relates to and can be incorporated into the IQmulus infrastructure. Results have shown that Tridicon fits better the environment provided by IQmulus. After the third project year, the focus has been shifted towards the second urban workflow due to its relevance and importance in algorithm development (cf. WP4) and its strong links to scalable big data processing as the main driver in IQmulus during the last project period.

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.

3.3 WORKFLOWS

Workflows and final user requirements of 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

User perspective and expectations, innovation aspects formulated in D1.2.3 according to the As Is Analysis

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 technologies in use at the Regione Liguria. The same problem hold, in general, for users of such data, who are missing suitable indexing for point clouds and partitioning of this big data into smaller datasets that can be handled in a processing workflow. For LS1/LS2, this is the case of processing aiming at the analysis of drainage basins, which could be extracted at the high resolution provided by LiDar data if the terrain model could be built on this large data set. 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.

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

Table 2. *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 as presented in D1.2.3

The novelty proposed by IQmulus is to prepare *terrain data in a model that 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. WE will re-orginize the high-resolution elevation model, given as LAS point clouds, according to a semantic criteria. For environmental stuff we will arrange points according the membership of the points to watershed. Each watershed will be grouped in a macro-set according two different solution: the Catchment Area Plan that reflect homogeneous region respect environmental legislation or the Warning Areas that groups together area with similar meteo-hydrologically response.

These boundaries are available from the related cartography owned and maintained by the Civil Protection group (shape file). See Figure 1 for an example of such a map.

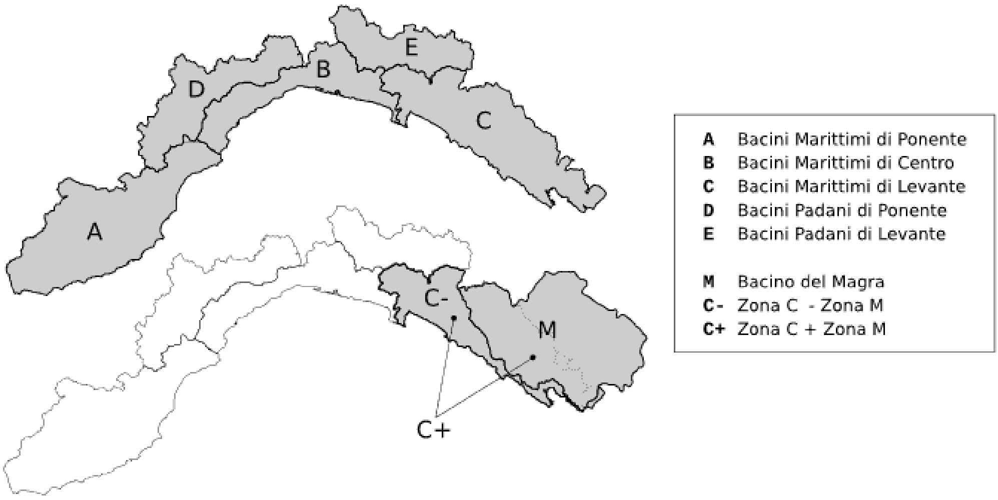


Figure 1. *Liguria Warning Area map*

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.

Implementation status and outcomes: Final User Requirements

The current state of the project shows that IQmulus developments targeted for LS1 will be implemented widely.

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	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Import data The users select and imports data belonging to the area of interest (point clouds and vector layer with semantic area boundaries)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI) The user has to input an area of interest.	High	Yes	Not implemented	Depends on GUI

WP4 services required:

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service #94 **Vector layer partitioning of a point cloud** (Splits/orders the available terrain data in areas defined by drainage boundaries) LS1_1_1	High	Yes	Implemented	Implemented and integrated
Service #10 **Outlier Classification in Point Cloud** (Detect (classify) outlier points in point cloud) LS1_1_2	High	Yes	Implemented,	Implemented and integrated
Service #35 **Resampling of Point Cloud** (Resample an unordered point cloud to match a given point density) LS1_1_3	High	Yes	Implemented	Implemented and integrated
Service #49 **Constrained triangulation** Create a triangulation of a point cloud preserving user-defined linear constraints (e.g. boundaries, feature lines) LS1_1_5	High	Yes	Implemented	Implemented and integrated

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. 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"

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Visualize Vector Layer partitioning of a Point Cloud	High	Yes	Not implemented	depends of visualization

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 and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service #92 **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	Low	Yes	Not implemented	Implemented and integrated
Service #48 **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	High	Yes	Implemented	Implemented and integrated

WP5 visualization functionalities required: see the visualization/ 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	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Multi-resolution model visualizer	High	Yes	Not implemented	-

WP4 services required:

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service #107 **Extract the triangle mesh to be visualized at the given LOD** (Extract the metadata related to the triangle mesh at a user-selected level of detail for visualization; the scale/level of detail is a parameter of this service). The mesh is then actually constructed by service #49 LS1_4_1	High	Yes	Implemented	Implemented and integrated
Service #49 **Constrained triangulation** Create a triangulation of a point cloud preserving user-defined linear constraints (e.g. boundaries, feature lines) LS1_1_5	High	Yes	Implemented	Implemented and integrated

WP5 visualization functionalities required:

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

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Visualize Triangle Mesh	High	Yes	Not implemented	Depends on visualization

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.

User perspective and expectations, innovation aspects formulated in D1.2.3 according to the As Is Analysis

The interests of Regione Liguria are concentrated on the possibility to fuse different sources of measured rain data in a single approximation model, while keeping explicit the contribution of each source of data to the approximation accuracy (see D1.2.3 for more details). Also, at the technological level, it is important to devise mechanisms to access the different repositories used, and select the time series needed for the analysis phase. Most of these issues have been addressed in the third year, and the plan for the finalization of the.

Table 3. *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) 3.4 Lidar, DTM, cartography 3.5 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 (see also publications on this subject, as reported in WP8);
- 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 (updated release of service #40);
- improvement of the precipitation field analysis by the development of a service for tracking the path of precipitation maxima (service #64);
- effective visualization of the precipitation field, including the rendering of the accuracy (to be fully developed).

In year 2 and 3 of IQmulus, we concentrated on services to support historical data analysis. During year 3, we have integrated the real-time monitoring settings and we have worked on the historical data analysis by taking into account the real-time perspective. To this end, developed services provide an output in a time frame suitable for the real-time monitoring. Moreover, we have already worked on a new release of service #40, able to merge rain data coming from radar measures, and we are now working on the development of a further refinement able to take into account the different contextual information and adopt corrective methods to properly model precipitations in urban areas as compared to rural ones.

Implementation status and outcomes: Final User Requirements

The workflow components and the corresponding services are listed below.

Preparation component

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. The service has been released as planned.

Generic system functionalities required:
The user has to select the temporal interval for the analysis of the time series (necessary for the service LS2_1_1 below)

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Selection of time interval	High	Yes	Not implemented	Depends on GUI

WP4 services required:

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service # 96 **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	High	Yes	Implemented	Implemented and integrated
Service #40 **Approximation by Kriging** (Approximation of heterogeneous rainfall data (punctual gauge measurements or radar measures) with ordinary Kriging) LS2_1_2	High	Yes	Implemented	Implemented and integrated
Service #67 **Approximation by RBF** (Approximation of heterogeneous rainfall data (punctual gauge measurements or radar measures) with radial basis functions) LS2_1_3	High	Yes	Implemented	Implemented and integrated
Service #58 **Approximation by LR-splines** (Approximation of heterogeneous rainfall data (punctual gauge measurements or radar measures) with LR-splines) LS2_1_4	High	Yes	Implemented	Implemented and integrated

WP5 visualization functionalities required:

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

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Visualize precipitation field over the multi-resolution model	High	Yes	Not implemented	Depends on visualization

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 and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service # 44 **Critical Points** Extraction of critical points (isolated and degenerate) from grids and triangulations LS2_2_1	High	Yes	Implemented	Implemented and integrated
Service # 64 **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	High	Yes	Implemented	Implemented and integrated

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:

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).

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Visualize the movement of the maxima of precipitation over the multi-resolution model	High	Yes	Not implemented	Depends on visualization

LS3 - Land Workflow 3: Flood and waterlogging detection

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

User perspective and expectations, innovation aspects formulated in D1.2.3 according to the As Is Analysis

The currently used flood detection method developed by FOMI is based on satellite image analysis, and is running in the ERDAS Imagine software environment. There are some partly-, or even non-automated steps in the process, where IQmulus could provide a solution beside other benefits.

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.

Expectations listed above are summarized in table 5 at the end of chapter LS3. Fulfilment of each expectations are indicated in this table, according to the final list of user requirements.

Table 4. *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 introduced in D1.2.3

The "flood and waterlogging detection" workflow is a complex workflow containing multiple algorithms. 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.

In D1.2.3 a Use Case diagram was showing the expected IQmulus solution for Flood and Waterlogging Detection workflow, including all the services planned to be implemented (see figure 2).

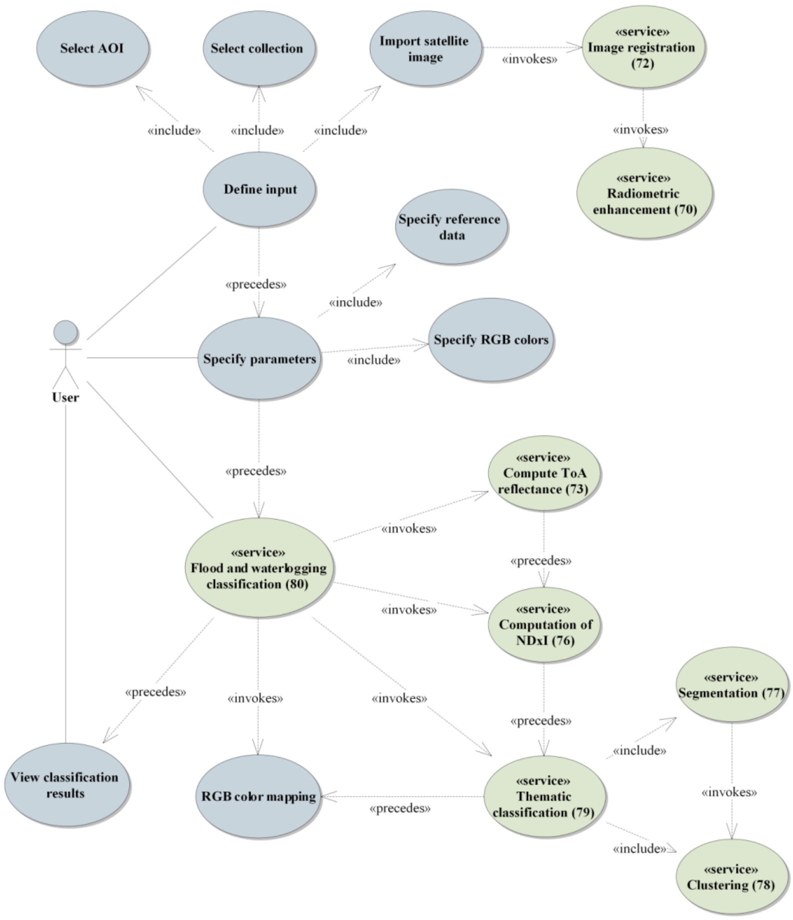


Figure 2. *Use Case Diagram for Flood and Waterlogging Detection, version 2, introduced in D1.2.3.*

Within the framework of the ***preparation component*** generic system functionalities and WP4 services were also defined. Importing data, defining and extracting Area of Interest (AOI) were determined as *generic system requirements*. Two WP4 services aiming pre-processing of satellite data were defined as “revised user requirement” in D1.2.3, because they are crucial steps before executing any kind of quantitative analysis:

- *Service #73: The Top of the Atmosphere (ToA) reflectance calculation* is necessary to complete, using satellite image metadata file. As hierarchical data storage formats with sub-datasets are available with all parameters necessary for calculation, automation of this step can be achieved.
- *Service #76: Spectral indices* are calculated as part of the preparation phase of the workflow. Version 3.0 will contain indices for soil (NDSI), vegetation (NDVI, SAVI, EVI, ARVI) and water (NDWI).The index values for the raster are computed from either the original image or the computed ToA reflectance data.

Two other WP4 services were not included among the revised user requirements, but were registered in eRoom Service Table, and planned to be implemented in the third project year:

- *Service #72*: *2D on 2D image registration*. Image registration is the process of transforming different sets of data into one reference system (by using control points or reference imagery. Registration is followed by resampling (scaling) with multiple supported interpolations. Version 2.0 contains control point based registration with nearest neighbour, bilinear, biqubic and Lanczos resampling techniques.
- *Service #70: Radiometric enhancement* transform image histogram as a preprocessing operation for further analysis.

In the ***feature detection component*** one generic system functionality was required, namely reference data specification, when the user selects and imports reference data (e.g. classified image) for classification. WP4 services as revised user requirements were the next:

- *Service #77: Raster segmentation.* The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse. The output of segmentation can be a segmented raster image, or a vector segment map, which contains various metadata on the individual segments.
- *Service #78: Clustering of raster data.* Clusters are formed by grouping the values in spectral space rather than image space as in segmentation. Clustering results a raster image categorized by the clusters or a vector cluster map containing various metadata on the individual clusters.
- *Service #79: Thematic classification of raster data,* match regions of the images to preselected categories to produce water categories.
- *Service #80: Flood and waterlogging detection:* Extracting flood/waterlogging information from multispectral remotely sensed images is a special form of thematic classification based primarily on several spectral indices.

There was no demand for functionalities in the ***change detection component***.

In the ***presentation/visualization component*** one WP5 visualization functionality (classification results need visualization), and one generic functionality was required (according to the plans, the user selects and imports a pre-prepared RGB colour table to visualize the results). WP4 services were not required.

Implementation status and outcomes: Final User Requirements

The current state of the project shows that IQmulus developments targeted for LS3 will be implemented widely. The most important services, which were included in the revised user requirements list in D1.2.3, and mainly are focusing on pre-processing and analysing of satellite images are implemented and integrated into the system already. Other services which were not included in D1.2.3, but listed in the WP4 service table are implemented also, and expected to be integrated into the IQmulus system until April, 2016. The chapter provides a summary of the final user requirements of LS3 workflow according to the component they belong to:

Preparation component

Generic system functionalities required:

The essential generic system functionality, namely "data import" is implemented and integrated to the IQmulus system, and is working well with satellite imagery. AOI selection and extraction could be a possible improvement until the end of the project, but we think, that further developments will not concentrate on this task, according to its low priority.

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Import data The user selects and imports data (satellite image)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI) The user has to input an area of interest.	Low	Yes	Not implemented	Depends on GUI
Area extraction Specified area is extracted from the images (performed by applying a mask to the image, no image transformation is required)	Low	Yes	Not implemented	Depends on GUI

WP4 services required:

All WP4 services proposed to be ready by the end of the third project year are already implemented. Some of them (TOA reflectance calculation and computation of spectral indices) are integrated to the IQmulus system, and is available for SPOT 5 and Landsat 8 Imagery. Until April, 2016 the rest of the services planned to be integrated to the system as well, because of their high priority.

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service #72 **2D on 2D image registration**	High	no	Implemented	Implemented and integrated
Service #70 **Radiometric enhancement**	Essential	no	Implemented, Hadoop enabled	Implemented and integrated # Service 70or 73
Service #73 **Computation of Top Of the Atmosphere reflectance**	High	Yes, as part of the Service #73	Implemented and integrated, Hadoop enabled	Implemented and integrated # Service 70or 73
Service #76 **Computation of spectral indices** Version 3.0 contains indices for soil (NDSI), vegetation (NDVI, SAVI, EVI, ARVI) and water (NDWI).	High	yes	Implemented and integrated, Hadoop enabled	Implemented and integrated

WP5 visualization functionalities required: N/A

Feature detection component

Generic system functionalities required:

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Specify reference data (Part of the Service #79: Thematic classification of raster data)	High	Yes	Not implemented	Implemented and integrated

WP4 services required:

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service #77 **Raster segmentation** performed using a predefined or specified algorithm	High	Yes	Not implemented	Implemented and integrated
Service #78 **Clustering of raster data** Generating clustered raster image based on pre-defined algorithm	High	Yes	Not implemented	Implemented and integrated
Service #79 **Thematic classification of raster data** Match regions of the images to preselected categories to produce water categories	High	Yes	Implemented and integrated (part of Service #80)	Implemented and integrated (Service #79)
Service #80 **Flood and waterlogging detection** special form of thematic classification based primarily on several spectral indices	High	Yes	Implemented and integrated, Hadoop enabled	Implemented and integrated
Service #74 **Topological Analysis of land 2D data**	High	No	Partly implemented	Implemented and integrated

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	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Specify RGB colours: The user selects and imports a pre-prepared RGB colour table	Low	Yes	Not implemented	Depends on GUI

WP4 services required: No

WP5 visualization functionalities required

Requirement name	***Exploitation priority***	***Included in D1.2.3***	***Status in 2015***	***Expected status in April 2016***
Thin client visualization: visualize classification results: - (1) initial raster image (GeoTiff); (2) extracted classification results/thematic rasters; + (3) legend	High	Yes	Not implemented	Depends on visualization

Table 5 is summarizing the fulfilment of User perspective and expectations, and innovation aspects according to the ones formulated in D1.2.3.

Expectation	Priority	Solution (2015)	Solution (April, 2016)
Pre-processing:			
Automation of ***cloud and cloud shadow masking***. 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	low	no	**probably**
Automation of ToA reflectance calculation	Essential	yes	yes
Automation of spectral index calculation	Essential	yes	yes
Processing:			
Threshold settings with manually digitized reference data	High	no	**yes**
Threshold settings with automated learning from the results	Medium	no	**yes**
Integrating terrain data (Digital Terrain Model) into the processing chain	Medium	no	**possibly**
Classification:			
Using decision tree learning classification methods	Medium	no	**possibly**

Table 5. *Expected fulfilment of User perspective and expectations by the end of the project*

LS4 – Land Workflow 4: Detection and characterization of landslides

User perspective and expectations, innovation aspects formulated in D1.2.3 according to the As Is Analysis

The currently used landslide detection method developed by FOMI & UBO is based on analysis of time-series of satellite images. The processing chain comprises several independent steps (resampling and cropping of input images, image correlation, filtering and visualization of the displacements fields) 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. 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 6.

Table 6. *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 introduced in D1.2.3

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

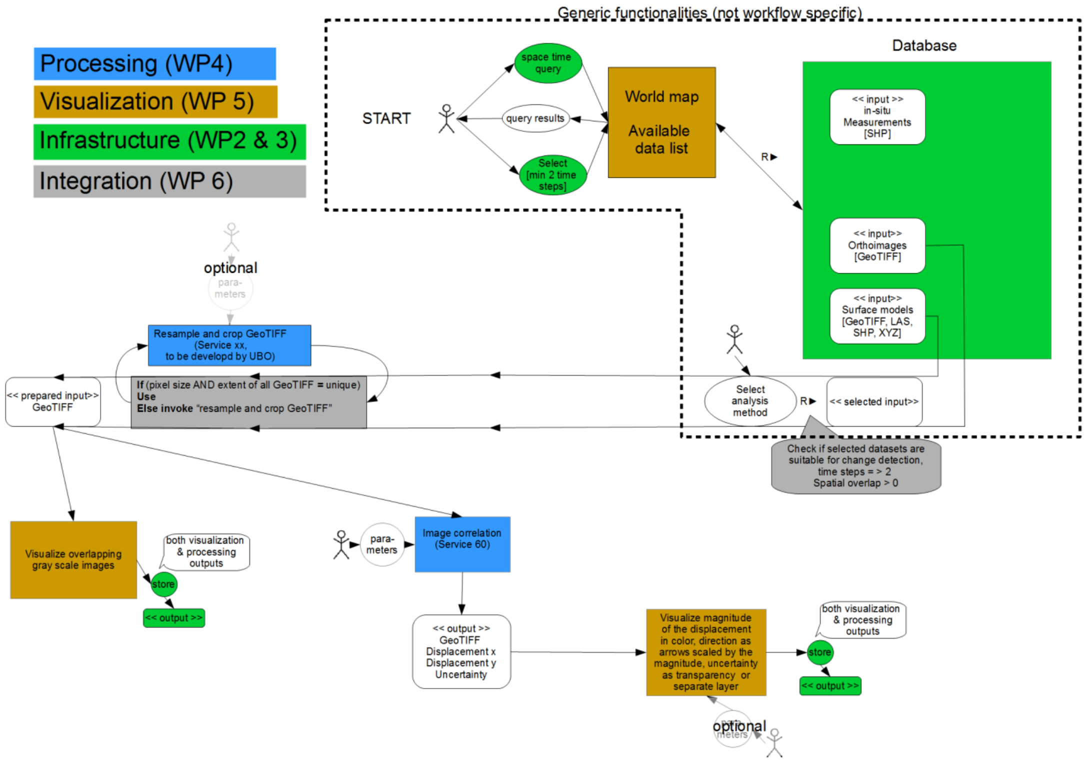


Figure 3. *Overview for the land workflow 4 showing the interactions of the user, infrastructure, processing services and visualization.*

The "detection and characterization of landslides" workflow presented above provides a higher level of automation compared to the current solution in use since it runs in the same infrastructure, and some of the services could be distributed by the Job Manager. It enables the processing of a large collection of multi-temporal VHR satellite images and is therefore addressing important user requirements.

A core part of the workflow is the image correlation service (#60) which has been demonstrated to excel over alternative solutions frequently used in the geoscience community in terms of robustness and scalability (see D4.5.1). The distribution of this multi-threaded algorithms service among multiple nodes with the Job Manager show the enhanced scalability of the workflow.

Implementation status and outcomes: Final User Requirements

The current state of the project shows that IQmulus developments targeted for LS4 will be implemented widely. The most important services, which were included in the revised user requirements list in D1.2.3, and mainly are focusing on pre-processing and analysing of satellite images are implemented and integrated into the system already. Another service (#103) which was not included in D1.2.3, but listed in the WP4 service table is expected to be integrated into the IQmulus system until April, 2016.

This is service #103, "spatio-temporal filtering of raster time series", which implements an algorithm for the analysis of time series of displacement fields as derived from service #60. Its main purpose is to find regions that show a consistent displacement over time and thereby eliminate false positive detections which are typically abundant in the displacement fields derived from sub-pixel correlation of optical orthoimages. It relies on multiple pair-wise matching to exploit the redundancy of deformation measurements recorded at different view

angles and over multiple time steps. This will offers new opportunities for the monitoring of surface deformation resulting from gravitational (e.g. glaciers, landslides) or tectonic forces (coseismic slip), and could exploit archived image time-series and the increasing flux of incoming data.

The chapter provides a summary of the final user requirements of LS4 workflow according to the component they belong to:

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 (or the other way around). The algorithm could use several interpolation methods such as sinc kernel interpolation to reduce aliasing effects and will automatically align the pixel grid of the input images.

Generic system functionalities required:

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Import data The user selects and imports data (satellite image)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI) The user has to input an area of interest.	Low	Yes	Not implemented	Depends on GUI
Area extraction Specified area is extracted from the images (performed by applying a mask to the image, no image transformation is required)	Low	Yes	Not implemented	Depends on GUI
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.	Low	Yes	Not implemented	Depends on GUI

WP4 services required:

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service #72 **2D on 2D image registration**	High	Yes	Implemented	Implemented and integrated

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 and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service #76 **Computation of spectral indices** Version 3.0 contains indices for soil (NDSI), vegetation (NDVI, SAVI, EVI, ARVI) and water (NDWI).	High	Yes	Implemented and integrated, Hadoop enabled	Implemented and integrated
Service #69 **Raster thresholding** Version 3.0 contains the following methods: constant based thresholding, balanced histogram based thresholding, Otsu's method, Bradley local thresholding, Bernsen local thresholding, Niblack local thresholding and Meantresh local thresholding.	High	Yes	Implemented	Implemented and integrated
Service #71 **Convolutional filtering** Combination of morphological closing and opening to eliminate isolated pixel	High	Yes	Implemented	Implemented and integrated

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 and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Service #60 **Measuring surface displacement** Hierarchical cross-correlation is used to find homologous points in two images	High	Yes	Implemented and integrated	Implemented and integrated
Service #103 **Time-series processing of VHR satellite images for surface deformation detection and measurements** Spatio-temporal filtering of raster time-series: relies on multiple pair-wise matching to exploit the redundancy of deformation measurements recorded at different view angles and over multiple time steps	High	No	Partially implemented	Implemented and integrated

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	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
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.	Low	Yes	Not implemented	Depends on GUI

WP4 services required: N/A

WP5 visualization functionalities required:

Data to be visualized:
- derived displacement (GeoTIFF)
- derived correlation coefficient (GeoTIFF)

Output: visualization of the calculated displacement field and the correlation coefficient.

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Visualization of displacement fields	High	Yes	Implemented in Fat Client	Implemented and intergration in Thin Client

LS5 - Comparison of simulated floods/landslides with observed data

Due to sharper priorities in the Land Workflow, LS5 will not be realized.

Workflows and final user requirements of the Marine Showcase

The Marine Showcase includes the following workflows:

1) Marine Workflow 1 (MS1): Marine DEM generation – deconflicted 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): This workflow will not be implemented.
4) Marine Workflow 4 (MS4): Change detection – measuring submarine dune migration.
5) Marine Workflow 5 (MS5): Lidar WaveForm costal Feature Extraction. NEW workflow.

LR B-spline surfaces are defined on regular domains. To handle large data sizes, regular tiling, approximation of each tile by one surface, and stitching of the resulting surfaces to produce a C1 continuous composite surface consisting of several patches, will be a part of the MS workflows.

The MS2 work flow assumes that a spline elevation model exists, either as one single spline surface or a composite surface. The model can be generated by the full MS1 work flow or a reduced service chain. In the MS2 description a reduced chain of services required to build the model will be described. Multiple surfaces generated according to the marine workflow MS1, with or without the deconfliction step, can serve as input for the displacement measurements in MS4.

MS1 – Marine Workflow 1: Deconflicted 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.

The 'As Is' analysis for MS1 workflow provided in D1.2.3 was based on current practices used within HR Wallingford and Ifremer. Summarizing presently used practices before listing User perspective and expectations we can state the follows:

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.

At Ifremer, deconfliction is presently overcome by manual intervention. Overlapping data sets are processed interactively, and according to the result of statistical and correlation analysis, the user decides how each data source is used for the interpolation.

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. Each of the steps in the process have a manual checkpoint on completion before the next step, so for large areas, the time taken to create a surface is far from instant.

At Ifremer the interpolation is carried out by kriging. 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.

Visualization is an offline activity of the current DEM process (Fledermaus is used to view the surface in 2D or 3D).

User perspective and expectations, innovation aspects formulated in D1.2.3 according to the As Is Analysis

Within this showcase IQmulus is demonstrating two key aspects of functionality that can be used in the marine test beds The first is the generation of a compact seamless surface with a representation which is adapting to varying levels of detail in the corresponding bathymetry. 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 7. *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.

Prior to year 3, the surface approximation part of MS1 is implemented for one tile or a point cloud of moderate size. The result is represented as one spline surface. Similarly, MS2 handles one tile or one moderately sized point cloud. Deconfliction and tiling and stitching as a strategy for handling large data sizes are the topics for year 3. One feature in this context is the possibility to alternately view point clouds originating from different sources as one combined data set or as separate point clouds.

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[7]] 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).
    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. Computation of the approximating surface is performed in an adaptive process where at each step in the iteration, the spline surface is refined where it is not accurate enough with respected to a specified tolerance and then the surface is recomputed given the refined spline space. This is a rather time consuming procedure, but the current experience indicates that if the sizes of the input data and the resulting spline surface are kept at a moderate level for each surface generation, an acceptable processing time can be achieved. Whether interactive surface generation can be obtained is still an open question.

Applying tiling and stitching in the surface approximation context enables processing with suitable data sizes and can be integrated with the IQmulus infrastructure. The approximation of tiles can be performed in parallel in the cloud. The computation time does 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). 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

[7] 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

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 was 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. In addition, as stated before, the ambition is to perform deconfliction and change detection in multi-temporal sea bottom data sets. As the work with the Marine Showcase has progressed, the focus has been dragged more towards large data sizes and deconfliction. This implies that the actual tiling and stitching process has been included in the work flow, but also that the hybrid representation will be limited to LR B-spline surfaces and point clouds which are not well represented by the spline surfaces. In year 3, the initial workflows have been extended with tiling, stitching and deconfliction.

Deconfliction: priority score for overlapping data surveys

In the following the creation of priority scores for overlapping data surveys will be explained in some detail.

Collections of bathymetric surveys are an example of a potentially ‘big data’ source structured as a point cloud. Individual surveys vary both spatially and temporally and can overlap with many other similar surveys. Where depth soundings differ greatly between surveys a strategy needs to be employed to determine how to create an optimal bathymetric surface using all of the relevant, available data, i.e. the *best* surface. As part of its SeaZone business, HR Wallingford employs the latest deconfliction techniques to make decisions on where to place extra priority on surveys or to remove surfaces entirely from those influencing the best, final surface. The output of this process is a bespoke set of scores attributed to the survey datasets that will allow for the creation of the best-possible surface for a given set of user specified criteria. The process allows the easy creation of a variety of surfaces by varying the criteria chosen.

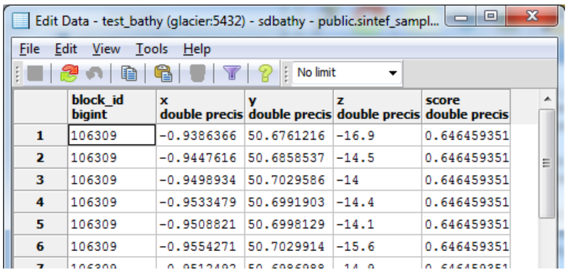
Bathymetry data is readily and freely available, however due to the high-skill levels required to process the data and the requirement to ‘filter’ to surveys to choose the most suitable, often bathymetry users do not have the resource to make best use of this data source. Working with a series of scenarios, HR Wallingford have developed a system with which to prioritise the surveys based on included metadata and generated statistics. The scenarios included criteria such as spatial coverage and time horizon. For example, for most use cases the most recent data would take a higher priority than an older one, although another client may have an interest in historic data. Other criteria included point density, survey technique, and even data type (digital survey data is often supplemented with charted data, but this may be deemed unsuitable for some uses due to shoal bias). It is possible to balance a user’s preference for different criteria to dominate and others to be subordinate.

Metadata including (but not limited to):

- Start date
- End date
- Technique of sounding
- Status of the survey
- Quality of data
- Quality of position
- Quality of sounding
- Type of survey

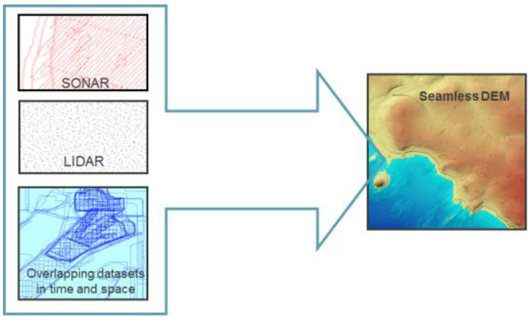
- Plus HR Wallingford's statistics based on survey data

The key result of this work is the creation of a workflow with which to automate the creation of *best-possible* marine bathymetry surfaces. The datasets are processed and stored within a database in a standard format and coordinate reference system. The criteria is then chosen and a score is assigned to each depth point - a score between 0-1 is given (at the time of writing, applicable to certain test datasets). The controlling scripts can be adapted to place a different priority on various metadata fields or evaluated statistics. A table is then created showing the location of every point and its associated score. Though the scores are the same for each survey block, this approach allows for outlying points to be assessed individually on their merit.



	block_id bigint	x double precis	y double precis	z double precis	score double precis
1	106309	-0.9386366	50.6761216	-16.9	0.646459351
2	106309	-0.9447616	50.6858537	-14.5	0.646459351
3	106309	-0.9498934	50.7029586	-14	0.646459351
4	106309	-0.9533479	50.6991903	-14.4	0.646459351
5	106309	-0.9508821	50.6998129	-14.1	0.646459351
6	106309	-0.9554271	50.7029914	-15.6	0.646459351

This method will play a key role in processing datasets for the IQmulus marine workflows, but will also go on to form an integral part of exploitable outcomes for the production of bathymetry products of both surface and point types. A simplified workflow is shown below.



Implementation status and outcomes: Final User requirements

In the third project year the initial workflows have been extended with tiling, stitching and deconfliction, because of the reasons set in the previous chapter. The components of the workflow and the corresponding services are described below, and illustrated with Figure 4 and 5.

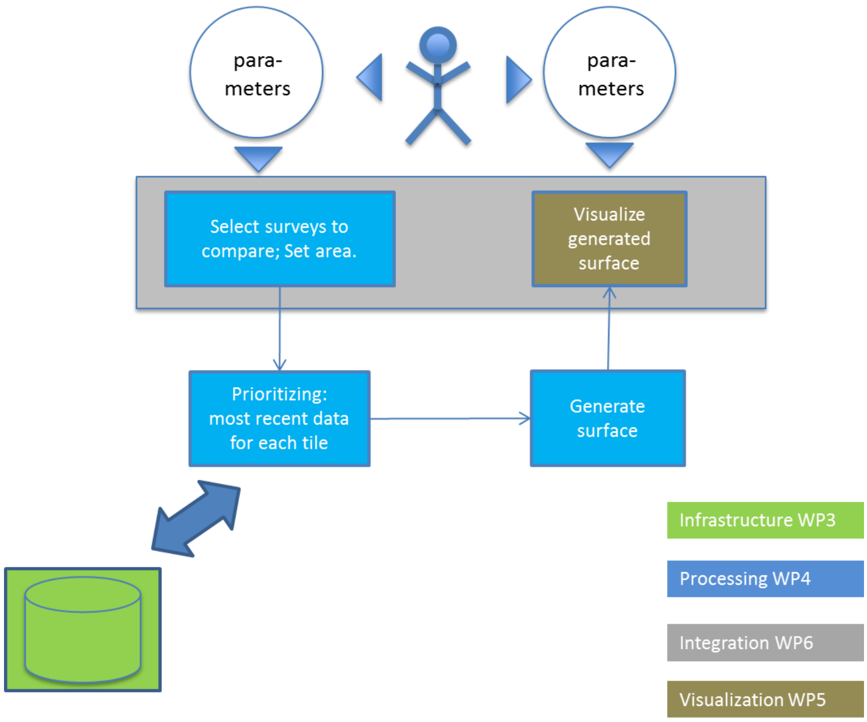


Figure 4. *Workflow components for MS1*

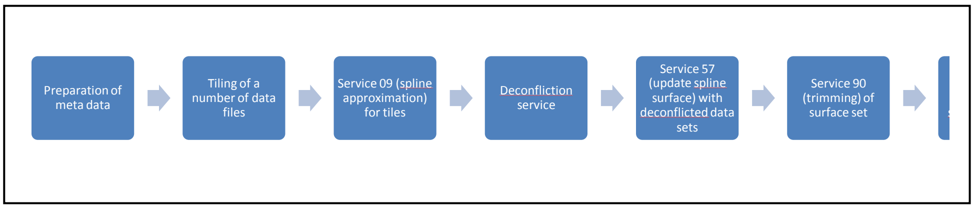


Figure 5. *Workflow components for MS1*

The workflow has a twofold aim: deconfliction and surface generation. The initial setting is a number of partially overlapping data surveys covering a selected area. The total data size is far above which is suitable for one spline surface approximation. The data surveys are likely to be split into blocks so each single entity has a moderate data size, but the domains of the data blocks are not necessarily adapted to the properties of the spline surfaces. The surfaces are defined on rectangular domains, while the domains corresponding to the data blocks can have any shape. The spline surfaces can be trimmed to fit a more arbitrary domain, but this makes the format more complex.

Preparation component

The first step in the work flow is to perform tiling, both to facilitate for multi node surface generation where each approximation process is exposed to data sizes appropriate for the functionality and to adapt the point cloud domains to the needs of the surface generation. Simultaneously, we may modify the scale of the point coordinates to get a good correspondence between the units used in the different coordinates. Information of any coordinate changes will be kept in metadata for the change to be reversed before completing the work flow.

Next, a first set of surfaces approximating the tiles are created. The approximation accuracy is not crucial at this stage, the surfaces will be at an intermediate position between trend surfaces

and a final surface representation. Data from all data surveys overlapping the surface domain are used in the approximation and only a few iterations in the adaptive surface approximation service are applied. The surfaces are created with an overlap, but pulled back to the tile boundaries. The surface generation can be done in parallel, tile by tile (figure 6).

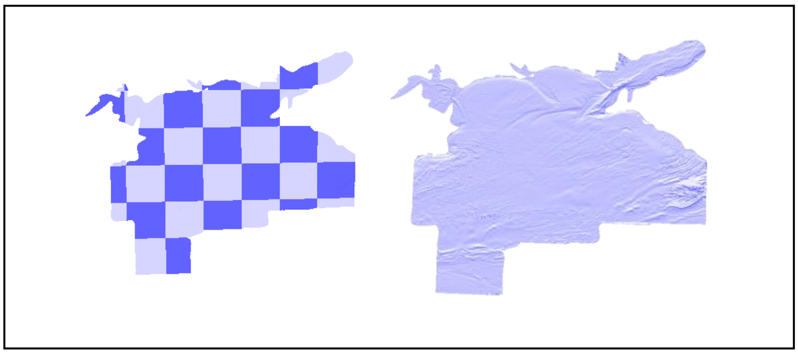


Figure 6. *surface generation can be done in parallel, tile by tile*

The surface set obtained from the initial surface generation serves as input to the deconfliction service together with the tiled data surveys. In addition, each survey is given a priority score related to confidence. In a situation where a conflict is detected between overlapping surveys, this score will be used to decide which points to keep for further surface generation and which points to remove. The dismissed points will be kept for quality control (MS2) and possible modifications of the final surfaces. The current surface set is used as a reference and distance fields are for each data survey computed with respect to the reference surfaces. This distance information is used to detect inconsistencies. The computation is done tile wise and allow for parallelization.

Having decided on which points to use, the current surfaces are updated to achieve a better accuracy. This involved further surface refinement and approximation of point clouds. The next step is to trim the surfaces to correspond to the point clouds. Also these two services can be applied in parallel. Finally, the surfaces are stitched to achieve C^1 continuity between adjacent surfaces and a seamless combined surface. The stitching operation is sequential, but applied to the surfaces and not the point cloud. The surfaces typically have only a fraction of the data size of the initial data.

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.
- The maximum allowed number of iterations in the adaptive approximation algorithm.

To indicate features, 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	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April, 2016*
Import data	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI)	Low	Yes	Not implemented	Depends on GUI

WP4 services required

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status I 2015*	*Expected status in April, 2016*
Service #128 **Define regular tiling (MS1_1_1)**	Low	No New service	Ongoing work	Implemented and integrated
Service #129 **Regular tiling (MS1_1_2)** Perform regular tiling of data surveys and prepare for approximation	Low	No New service	Ongoing work	Implemented and integrated
Service #122 **Tiled point cloud to LR-spline surface (MS1_1_3)** Tiled version of **Service #9**	Essential	Yes	Implemented and integrated Are being updated to a multi tile setting.	Implemented and integrated
Service #123 **Deconfliction (MS1_1_4)** Divides sub parts of the initial data surveys into two groups: keep and remove	Essential	No. New service	Ongoing work	Implemented and integrated
Service #124 **Update spline surface (MS1_1_5)** Tiled version of service #57	Essential	No	Implemented To be updated to a multi tile setting	Implemented and integrated
Service #125 **Trimming of LR B-spline surface (MS1_1_6)** Tiled version of service #90	Essential	No	Implemented and integrated. To be updated to a multi tile setting	Implemented and integrated
Service #85 **Stitching of LR B-spline surfaces (MS1_1_7)** Modify surfaces generated with respect to a regular tiling to achieve C^1 continuity between the surfaces	Essential	No	Ongoing work. Pre-released exists	Implemented and integrated

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:

The generated combined spline surface is visualized (3D) on a fat client. The boundaries of the polynomial patches from which the surface is constructed can also be shown. The selection of surface patches which actually is shown at each moment will depend on the view. Hardware constraints may imply a need for reducing the number of surfaces which are simultaneously kept in memory.

Requirement name	***Exploitation priority***	***Included in D1.2.3***	***Status in 2015***	***Expected status in April, 2016***
Fat client visualization: 3D visualization of a combined LR-spline surface consisting of a regular set of surface patches	Essential	Yes, single file version	Ongoing work	Implemented

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.

User perspective and expectations, innovation aspects formulated in D1.2.3 according to the As Is Analysis

This workflow was introduced in 2014 for the LR-spline surface representation of the DEM. No As-is workflow exists.

Proposed IQmulus solutions introduced in D1.2.3

Because of minor changes were made to finalize the user requirements compared to requirements documented in D1.2.3, we document the solutions in detail in the next chapter.

Implementation status and outcomes: Final User requirements

In the marine scenario 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.
- Compare the LR-spline approximation to the source point cloud to check suggested outliers or features
- Approximate the new survey with an LR-spline and compare the old and new LR-spline surfaces

Figure 7 addresses the two first alternatives as this only involves the approximation error of the LR-spline approximation to one point cloud. Note, however, that one might choose to set the threshold for whether or not a distance field is to be computed to zero. Then the entire point cloud will be included in the analysis.

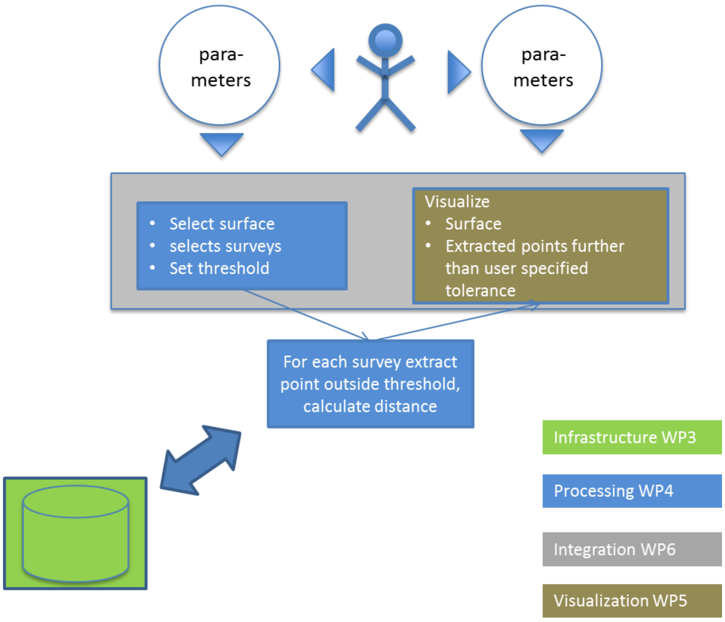


Figure 7. *Workflow for comparing an LR-spline surface and a point cloud*

Workflow components are illustrated in figure 8.

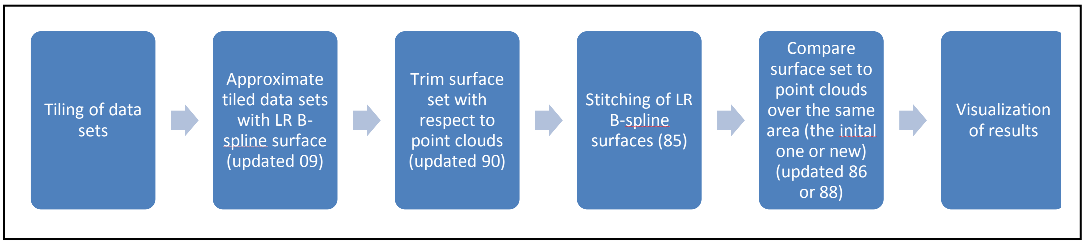


Figure 8. *Workflow components for MS2*

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

Table 8. *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 preparation component of MS2 can be identical to the preparation component of MS1 or a simplified version thereof (MS1_1_1, MS1_1_2, MS1_1_3 and MS1_1_7). The workflow can also be applied to one point cloud and a corresponding surface thereby omitting MS1_1_1, MS1_1_2 and MS1_1_7. Furthermore, the trimming service (MS1_1_5) is not required for the analysis, but improves the visualization experience.

Generic system functionalities required:

First, the user has to select and import an LR-spline surface and the surveys to compare to, and also has to define an area of interest to identify surfaces and point clouds. The comparison area is defined by the overlap between the surface domain and the domain corresponding to the point cloud.

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 or a surface combining a regular set of LR B-spline paths. 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.

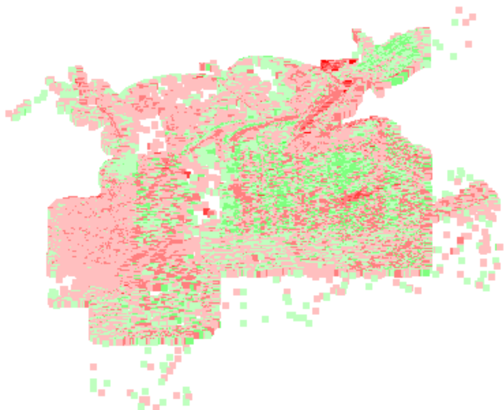


Figure 9. *A combination of several data surveys are coloured according to their difference field with respect to a set of LR B-spline surfaces. Red points lie above the surfaces and green points below. The colour gets more intense when the distance with increasing distance.*

Several data surveys may overlap an LR B-spline surface. This is the typical situation in a deconfliction setting (MS1). Analysing the distance field of the overlapping surveys, the distance pattern will differ. This feature may be a component in an interactive deconfliction approach.

Generic system functionalities required: N/A

WP4 services required:

Two alternative services are provided. One computes the distance between the surface and the point cloud for each point. The other organizes the points into groups depending on specified limits with respect to the distance between the point and the surface.

Service name and identifier in eRoom	***Exploitation priority***	***Included in D1.2.3***	***Status I 2015***	***Expected status in April, 2016***
Service #126 **Compare point cloud to LR-spline** **(**Computes the distance between a point cloud and an LR-spline surface) Tiled version of **Service #88**	Essential	Yes	Released. Are being updated to tiled data	Implemented and integrated
Service #127 **Compare point cloud to LR-spline surface and classify points with respect to distance (**Computes the distance between a point cloud and an LR-spline surface) Tiled version of **Service #86**	Medium	Yes	Released. To be updated to tiled data.	Implemented and integrated

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:

LR-spline surface and the original point cloud with computed distance to LR-spline surface has to be visualize. The points can interactively be included in or excluded from the visualization depending on the distance between the points and the surface.

Requirement name	***Exploitation priority***	***Included in D1.2.3***	***Status in 2015***	***Expected status in April, 2016***
Distance field visualization	Essential	Yes	Ongoing work to update previous version	Implemented

MS3 - Marine Workflow 3: Simple feature extraction

Due to sharper priorities in the Marine Workflow, MS3 will not be realized.

MS4 – Marine Workflow 4: Measuring submarine dune migration

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.

The main steps of the currently existing solution are (mainly rely on manual processing steps):

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.

User perspective and expectations, innovation aspects formulated in D1.2.3 according to the As Is Analysis

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.

The MS4 workflow proposed for IQmulus relies on a measurement of dunes displacements method based on image correlation of a pair of DEM generated from point clouds.

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 introduced in D1.2.3

Because of minor changes were made to finalize the user requirements compared to requirements documented in D1.2.3, we document the solutions in detail in the next chapter.

Implementation status and outcomes: Final User requirements

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 10):

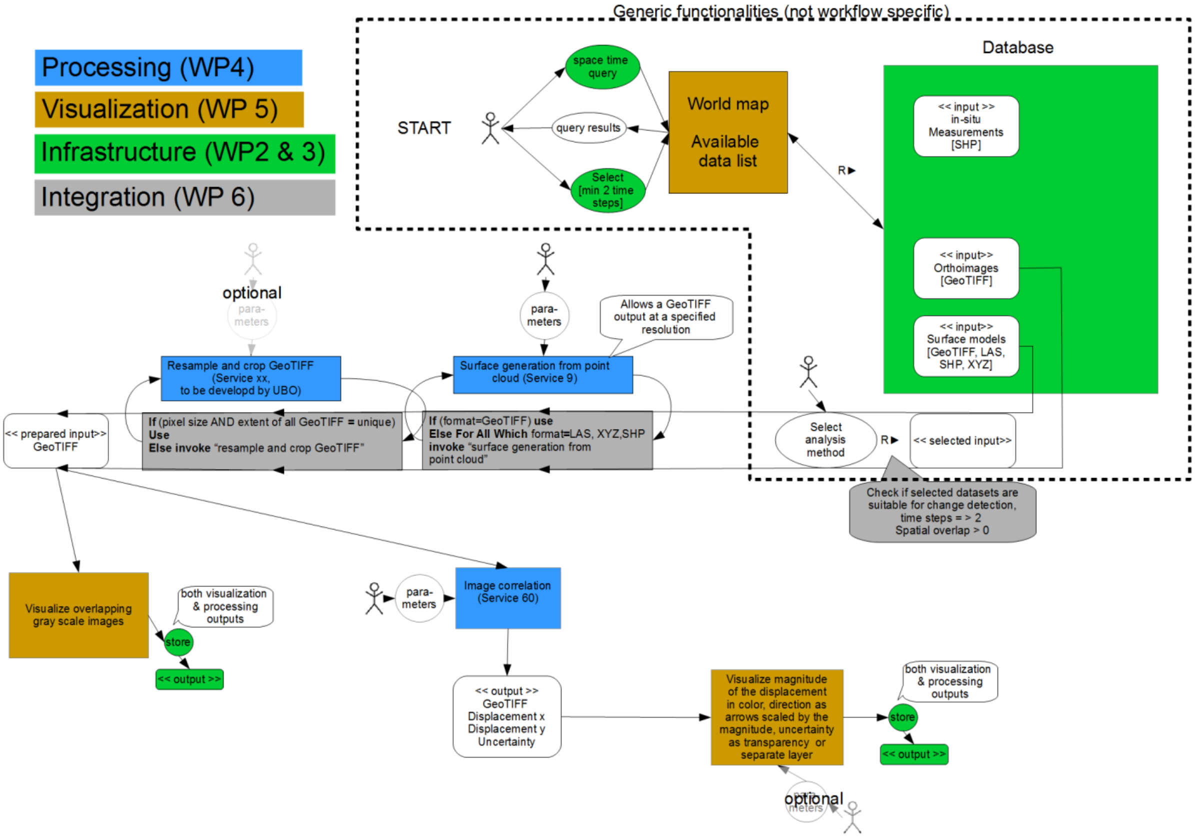


Figure 10. *Overview for the marine workflow MS4 showing the interactions of the user, infrastructure, processing services and visualization.*

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:

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

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April, 2016*
Import data (upload data and metadata)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI)	Low	Yes	Not implemented	Depends on GUI
Area extraction (AOI is extracted from the image)	Low	Yes	Not implemented	Depends on GUI
Storage of resulting raster	Low	Yes	Not implemented	Depends on GUI

WP4 services required:

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status I 2015*	*Expected status in April, 2016*
Service #9: Generate spline surface from parameterized point cloud ((MS1_1_2)	Essential	Yes	Implemented and integrated	Implemented and integrated
Service #90: Trimming of LR B-spline surface (MS1_1_5) Restrict the valid part of the surface to correspond to the extent of the point cloud by a trimming loop.	Essential	No	Implemented and integrated	Implemented and integrated
Service #91: Spline surface to raster	Essential	Yes	Implemented	Implemented and integrated
Service #72: Resampling and cropping of raster. A sinc kernel interpolation is used to harmonize the resolution of the input images.	High	Yes	Implemented	Implemented and integrated

WP5 visualization functionalities required:

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April, 2016*
Visualization of the resulting surface raster as greyscale images (raster image of resampled spline surface)	Low	Yes	Not implemented	Implemented and integrated

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:

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.

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April, 2016*
Storage of resulting raster	Low	Yes	Not implemented	Depends on GUI

WP4 services required:

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status I 2015*	*Expected status in April, 2016*
Service #60: Change detection/Measuring surface displacement. Hierarchical cross-correlation is used to find homologous points in two images.	High	Yes	Implemented and integrated	Implemented and integrated

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	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April, 2016*
Storage of resulting raster	Low	Yes	Not implemented	Depends on GUI

WP4 services required: N/A

WP5 visualization functionalities required:

The following data has to be visualized:
- derived displacement (GeoTIFF)
- derived correlation coefficient (GeoTIFF)

The output has to be: (Combined) visualization of the calculated displacement field and the correlation coefficient.

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April, 2016*
Visualization of displacement field	High	Yes	Implemented in Fat Client	Implemented and integratied in Thin Client

MS5 – Marine test bed: Lidar WaveForm coastal Feature Extraction

Goal: Production of a seafloor 2D map with labels corresponding to identified sea bottom types Using Bathymetric Lidar Full WaveForm Data

This workflow is a new one, i.e. was not introduced in D1.2.3. Therefore we provide here a full description, with an As-Is analysis. The workflow demonstrates Lidar WaveForm coastal Feature Extraction, and is related to eRoom User stories #39 to #43.

Description and As-Is Analysis

Introduction and main steps of the current solution

Airborne Lidar is a widely adopted technology for high-resolution shallow coastal waters applications. However, bathymetric lidar users generally have access only to the final product, a cloud of points with coordinates x, y, z to create a Digital Terrain Model (DTM). Although Lidar was developed for depth measurement, the recorded data (lidar backscatter intensity or waveform) contain much more information than simple bathymetry. As with the terrestrial applications (Mallet and Bretar, 2009; Puech et al., 2012; Duong, 2010), they consider more useful to study the signals themselves (Full waveform return), to develop their own algorithms for depth extraction as well as to get more information on the bottom types or the turbidity of the water contained in the signals (Tulldahl and Wikstrom, 2010). An example of such information is the presence of high dense vegetation on the sea floor. A schematic example is shown in Figure 11.

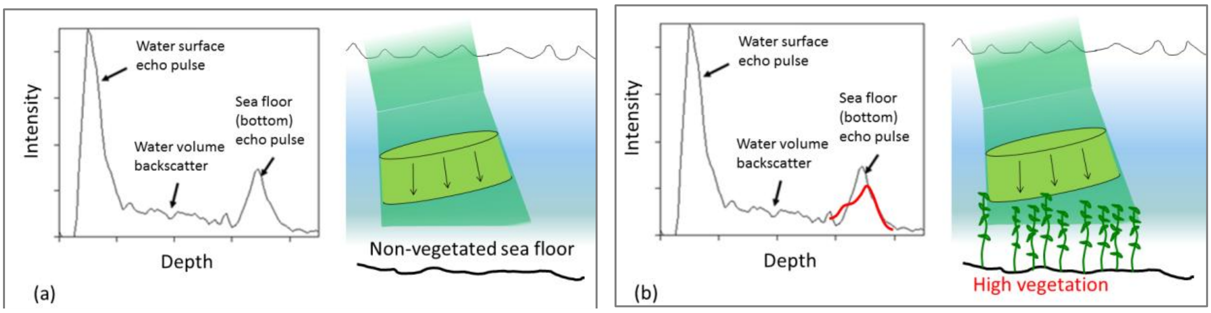


Figure 11. *Schematic example of how the lidar waveform can be used to detect presence of vegetation on the sea floor. (A) the bathymetry lidar waveform comprising the water surface echo, water volume backscatter, and the sea floor echo pulse from the bottom without vegetation. (B) Schematic view of the vegetation influence on the sea floor echo pulse.*

This 'as is' analysis is based on current practices used within Ifremer and his partner Swedish FOI (Defence Research Agency). It aims t describe the processing sequence for lidar waveform extraction and for looking at relevant lidar data structure.

There are four main stages of the process: (1) data import, (2) Wave form attribute calculation (3) Attribute classification and (4) Quality control (5). The following diagram summarizes the main processing step. The workflow can call external libraries or user-generated code. Each of the steps in the process could be based on external tool and have a manual checkpoint on completion before the next step.

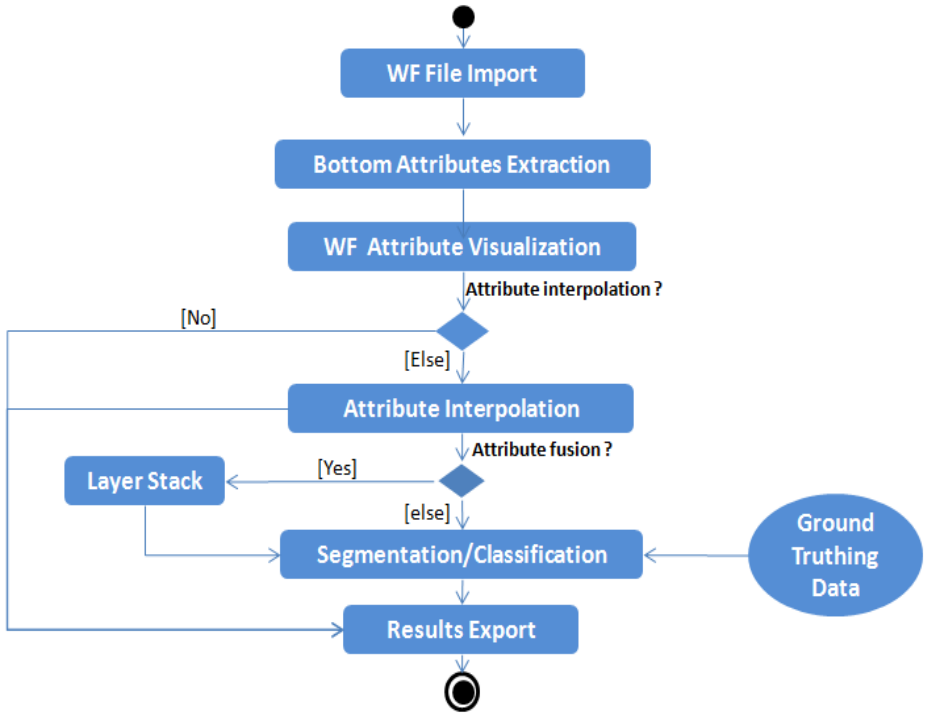


Figure 12. *Full wave form processing to produce seafloor 2D map*

Process: Data Load

Currently, the market for airborne bathymetric lidar is very limited. There are only four airborne bathymetric lidar manufacturers in the world (Table 10). The principles of operation are approximately the same, but the types of scanning transmission frequencies, heights and flight speeds may vary.

Table 10. *Airborne bathymetric lidar manufacturers*

Manufacturer		Lidar System
Optech Incorporated Canada	Canada	SHOALS 1000 et SHOALS 3000
Fugro Lads	Australie	LADS MkII
AHAB	Suède	HAWKEYE II
Optech International US	US	CZMIL

Raw data have also Sensor-specific structure and could be delivered in different file format types.

- One is the American Society for Photogrammetry and Remote Sensing (ASPRS) lidar laser (LAS) binary format in files with ".LAS" extension. Tool libraries to use this type of data are available on the Internet at www.liblas.org. LAS format version (1.3) is used to store the complete wave (LAS Specification version 1.3 – R10 ASPRS 2009).
- Waveform data (one file for each flight line) are also delivered in a modified proprietary binary format. So a set of programs to load, read, display, and work with the data (visualization, selection, subsetting, analysis) needs to be developed.

Process: Wave forms attribute extraction

Extraction of features from bathymetric lidar data is the first step in the classification of sea floor habitats. . Some information can be retrieved by analysing depth data (e.g. sea floor slopes) but more information can be extracted while analysing the shape of the received lidar waveform (Fig. 13).

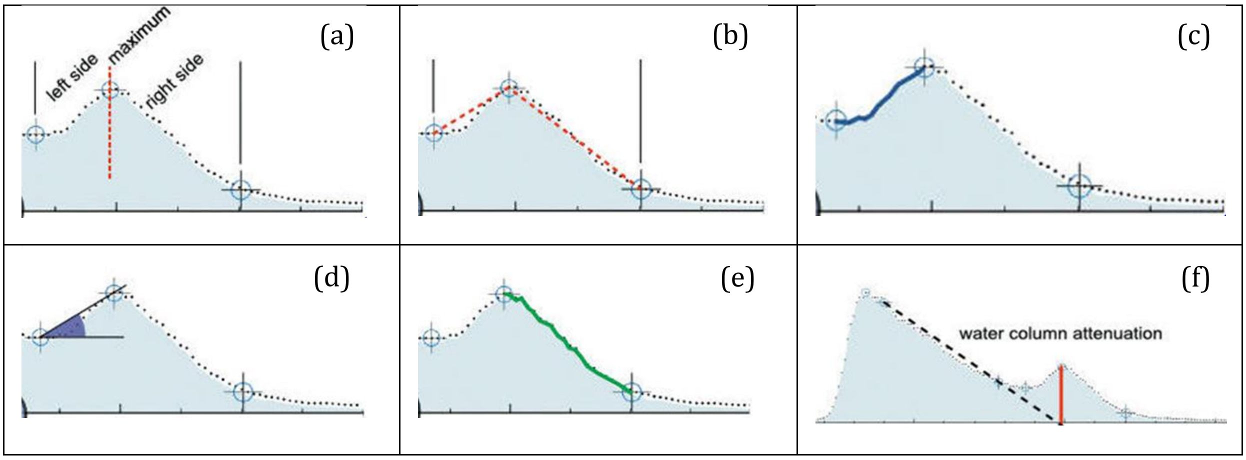


Figure 13. *Waveform and signal description for the deep green channel as used in the waveform analysis: (a) skewness; (b) kurtosis; (c) left-side length; (d) left-side angle; (e) right-side length; (f) amplitude (From Cottin et al. 2009)*

To the best of our knowledge, there is currently no off-the-shelf software for processing bathymetric lidar waveforms. Manufacturers have developed applications but are limited to their internal use. The scientific community tries therefore to develop by itself the tools to explore the potential waveform lidar to advance its research problematic.

Process: attribute Classification

Several GIS and image processing solution offer classification tools. Services were also developed by IQmulus could help to complete this step. Each parameter needs to be rasterized using an interpolation process. This process is repeated for each waveform parameter for which an image is created and witch can therefore be processed separately. However, staking the single attribute images to generate a multiband image may increases the classification capability. The following figure shows an example of the seabed map obtained by using decision Tree classification of Shoals Lidar waveform data (Cottin et *al.*, 2009).

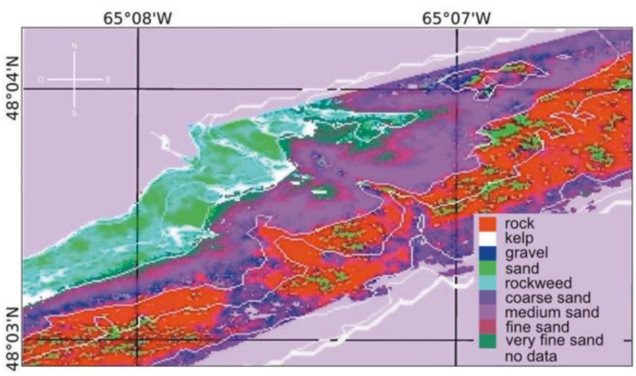


Figure 14. *Output map after application of decision tree, showing seabed classification over Saint Godefroi area using nine bottom classes. The white lines are the isobaths at 1 m intervals (from Cottin et al., 2009)*

About visualization aspects

To explore the raw data; a set of programs to load, read, display, and work with the data (visualization, selection, subsetting, analysis) needs to be developed, as the data formats are sensor specific and manufacturers' proprietary.

For the processed data, GIS is used to visualize intermediate steps in the production process. Likewise once attribute layers are generated tools such as ArcGIS are used to view in 2D results.

Visualization is not an inherent part of the current Lidar wave form process.

User perspective and expectations, innovation aspects

Today, there is no commercial tool to handle this data. Sensor manufacturers have developed applications for internal use and are not necessarily accessible to end-users. Also, raw data are provided by flight lines in separate full-waveform and navigation files. Either the signal from the shallow or the deep green receiver, are referred as Point Data Records (PDR), and the waveform data (WAV). Unlike topographic laser scanners, data are recorded in bathymetric laser scanners from more than one receiver; there are ten channels in the case of HawkEye II. All this make complex the input data that needs to be organized in optimal structure in order to make its processing more efficient.

The relevance of the workflow to "big data" issues is synthesized in Table 11.

Table 11. *Big Data issues for the MS5 workflow*

Indicator	Relevance	Comment
Volume	High	In many cases data need to be processed over entire survey (e.g. Penmarc'h site : 394 GO for a surface area of 124 km²)
Variation	Medium	Data from different sensors, quality varies as well as the number of surveys in different regions.
Velocity	Low	The temporal frequency of bathymetric surveys is currently rather low and real-time processing at sea is typical not considered as an option.
Analytics	High	Need for efficient tools to extract information from bathymetric Lidar data

Proposed IQmulus solutions, final user requirements

Within this workflow, IQmulus is focused on Lidar waveform processing. It's demonstrating two key aspects of functionality that can be used in the marine test beds.

- The first is a service to perform both sequential and spatial indexing and initial control of the data delivered from an airborne lidar bathymetry survey system. This allows for direct access to specific PDR or WF files given specific geographic coordinates is created based on the index results from the PDR data.
- The second is to contribute to the development of services for characterising coastal seabed and detecting vegetation using attributes extracted from bathymetric LiDAR waveforms.

Preparation component:

Data load

Currently, the data format is sensor-specific, and proprietary scripts are required for reading the waveforms. But we expect for the new sensor that data will be delivered with Free Reading Library as it's the case today for the Hawk Eye III. So the workflow assumes that datasets are available in ready format and structure. Flight lines are split in individual source data files containing 1000 shots records. Each shot corresponds to four return echoes from Infra red, deep and shallow green channels.

Reading raw data was performed using the LSS software from (AHAB *ReadWaveformsHE.mexw32)* and FOI scripts. These tools enable data extraction that was stored in a MATLAB structure array.

Data indexing

Service (#111) performs both sequential and spatial indexing and initial control of the data delivered from an airborne lidar bathymetry survey system. The input data to this service consists of waveform data files (WAV-files) and point data record files (PDR-files). The waveform data are the actual "raw" received time-signals recorded in the lidar system. The bathymetric receiver consists of a shallow and a deep channel each having a quad-pixel detector. Thus, from one single emitted laser pulse, four bathymetric waveforms are received in both the shallow and deep channels. From one lidar pulse, four point data records (PDR) are generated in the case of Hawk Eye II post-processor.

The input data to this service consists of WF (waveform) data files (WAV_HD files), point data record (PDR) files, and CH_MH files (Channel & Metadata Header files). The CH_MH files comprise information about the content in the data files. Indexing is necessary because the input survey data (PDR and WF) are stored in separate files. A spatial index file named "SpatIdx.mat" file is created based on the index results from the PDR data. The spatial index file allows for direct access to specific PDR or WF files given specific geographic coordinates. The connection between PDR and WF data is managed by use of the timestamp and pixel index which are stored both in PdrIdx.m and in WavIdx.m.

Generic system functionalities required:

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Data load (Import data: upload data and metadata)	Essential	No	Implemented and integrated	Implemented and integrated

WP4 services required: N/A

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status I 2015*	*Expected status in April, 2016*
Service #111: BWF1_indexing. Performs both sequential and spatial indexing and initial control of the lidar bathymetry survey system data. **MS5_1_1**	Low	No	Implemented	Implemented and integratied in Iqmulus platform

WP5 visualization functionalities required:

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April 2016*
Point cloud visualization	High	Yes	Implemented and integrated	Implemented and integrated

Feature detection:

Attribute calculation

This service allows the user to extract parameters from the lidar WF (waveform) useful for classification of Seabed bottom types (rocks, vegetation, soft bottom, etc.) for marine habitat mapping. Execution of this service requires that Service #111 was previously executed on the dataset.

As schematically shown in figure 15, the bathymetric lidar return waveform can be split into the three returns surface return echo, water volume (backscatter) return, and the sea floor echo return. The first, the surface return is generally the strongest. The amplitude and shape of the surface return depends on the ripples on the surface of the sea (on a completely flat surface, there would be no surface return because of specular reflection). The second one is the return of the water column that increases until the pulse is completely within water and decreases exponentially by a factor related to the depth and the attenuation coefficient of the water often denoted by K. This attenuation depends on the type of water and its turbidity. Finally the third part is the return from the sea floor. This last part is the most important part containing information for classification of the sea floor.

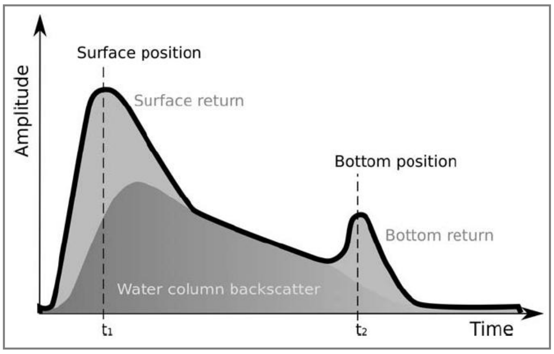


Figure 15. *Echoes from water in the green channel consists of three components: surface return, water volume backscatter and bottom; From Allouis et al., 2010*

The algorithm uses information both from the PDR and from the WF to extract information from the waveform. The main steps of the service are:

1. To classify land and water using parallel and perpendicular echoes from the NIR channels of the system.
2. The gain curves for shallow and deep channels used to linearize the shallow and deep green echoes respectively.
3. The reported depth in the PDR is used to identify the locations of the leading and falling edges in the original echo (Location of the maximum in the raw signal is usually the peak of the echo from the seabed in the echoes from deep water).

4. An interpolated water volume backscatter wave is subtracted from the sea echo, to obtain a difference wave (difference between sea floor and estimated backscatter level at the same depth as the sea floor).
5. Attributes (Pulse width, Pulse area, Peak amplitude...) of the sea floor echo are lastly extracted from the difference waveform, for both the shallow and the deep sensor channels.

The following figure shows an example of obtained attribute extraction results:

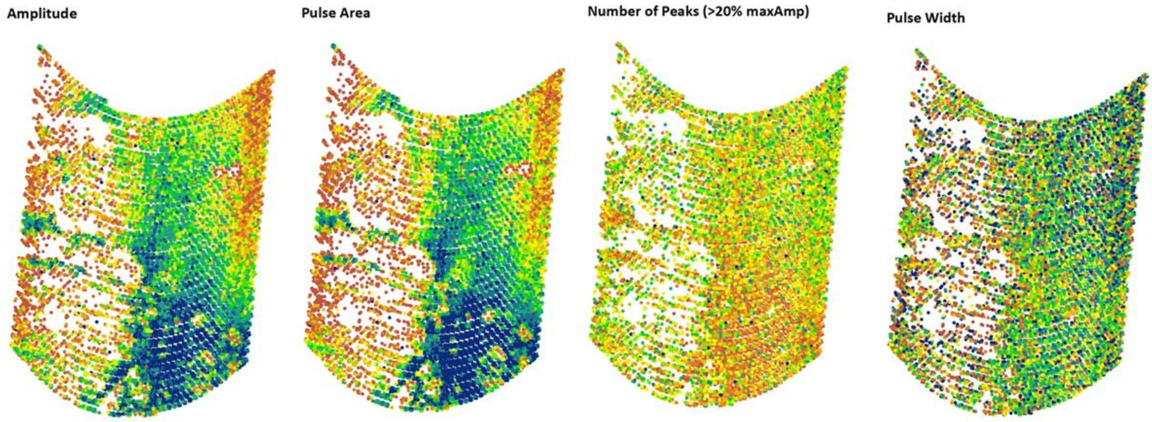


Figure 16. *Attributes extracted from the waveform: Amplitude (A), Pulse width (B), Pulse area (C), Corrected water depth (D), Number of peaks with amplitude greater than 10% of maximum amplitude (E) and Number of peaks with amplitude greater than 20% of maximum amplitude (F) for the shallow and deep channels.*

Attribute classification

This step is planned to be carried out using available GIS and image processing solution. Thereafter an example of obtained classification.

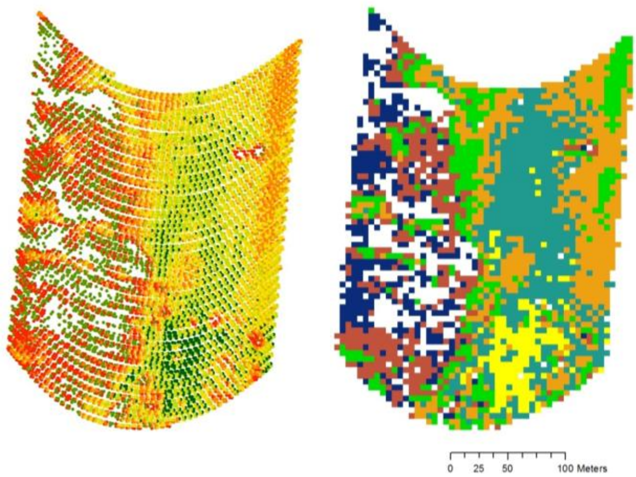


Figure 17. *Bathymetric LiDAR points classified into six classes using a K-means classifier (left); the points were converted to 5m × 5m pixels based on the majority class (right)*

Quality control

A prototype of this service has been used and tested on several small and larger scale survey data sets from the Hawk Eye II system. The largest those data sets were up to about 40 flight lines and 100 km^2. We also have ground truthing data (video and diving samples) concerning sea bottom types to check the ability to distinguish between vegetated and non-vegetated sea bottom.

Generic system functionalities required: N/A

WP4 services required: N/A

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status I 2015*	*Expected status in April, 2016*
Service #112: BWF2_WF_attribute_extraction extract parameters from the lidar waveform useful for classification of Seabed bottom types. **MS5_1_2**	Low	No	Pre-release exists Improvements are expected	Implemented and integratied in Iqmulus platform

WP5 visualization functionalities required: N/A

Change detection component: N/A

Presentation/visualization component

Generic system functionalities required: N/A

WP4 services required: N/A

WP5 visualization functionalities required:

The results of Lidar waveform processing need to be visualized as different extracted attribute layers. Basically it is a color scale visualisation of attribute layer according a point cloud values

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April, 2016*
Visualization of lidar wave form attributes	High	Yes		

Input and reference data provided for workflow development and testing

Ifremer will provide a survey data for use and testing in the project. The trial areas will cover 2 sites (Figure 20):

- Penmarc'h site : 394 GO for a surface area of 124 km^2
- Molène archipelago site : 113 GO

Subsets from Penmarc'h site are already available in IQmulus platform for scalability test.
http://146.140.214.126/fs/IFREMER/test_Penmarch

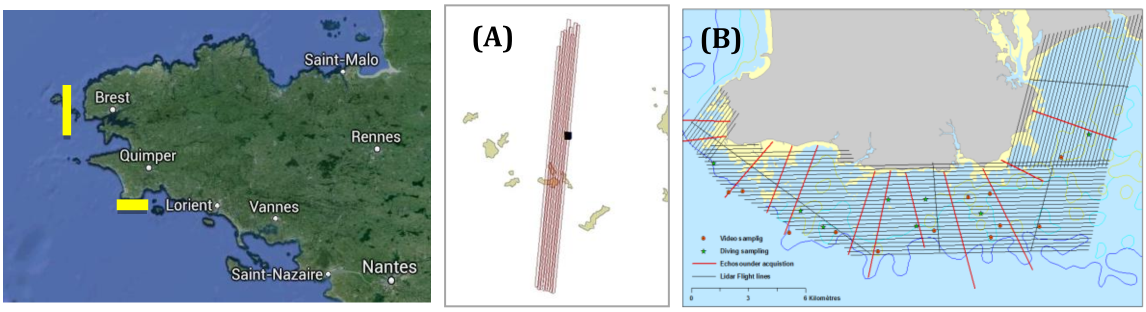


Figure 18. *Trial area for Marine workflow MS5*

References

Cottin A.G, Forbes D.L., Long B.F. 2009. Shallow seabed mapping and classification using waveform analysis and bathymetry from SHOALS lidar data. Can. J. Remote Sensing, Vol. 35, No.5, pp 422 – 434, 2009.

Duong H. V., 2010. Processing and application of ICESat Large Footprint Full Waveform Laser Range Data. Ph.D. Thesis. 193p.

Mallet, C. Bretar, F., 2009. Full-waveform topographic Lidar: State of the art. ISPRS Journal of photogrametry and remote sensing. 64(1), 1-16.

Puech C., Durrieu S. and Bailly J.S., 2012. Airborne lidar for natural environments: research and applications in France.

Tulldahl M., Wikstrom S. 2010. *Initial report on statistical separability of vegetation and bottom type in laser data*, Technical Report FOI-R--2961—SE, 51pp.

Workflows and final user requirements of the Urban Showcase

As it was written previously, for the Urban Showcase setting, two workflows were proposed in D1.2.3 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.

US1 – Detection of buildings for monitoring and cadastral updating

The first workflow (US1) gave the opportunity to investigate how commercial software from the current "ecosystem" of mapping authorities relates to and can be incorporated into the IQmulus infrastructure. Results have shown that among eCognition and Tridicon, Tridicon fits better the environment provided by IQmulus. After the third project year, the focus has been shifted towards the second urban workflow due to its relevance and importance in algorithm development (cf. WP4) and its strong links to scalable big data processing as the main driver in IQmulus during the last project period.

Therefor we won’t give further details about US1 in this document.

US2: Individual tree extraction from urban LMMS data

User perspective and expectations, innovation aspects formulated in D1.2.3 according to the As Is Analysis

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.

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.

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 12. *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.

Implementation status and outcomes: Final User requirements

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.

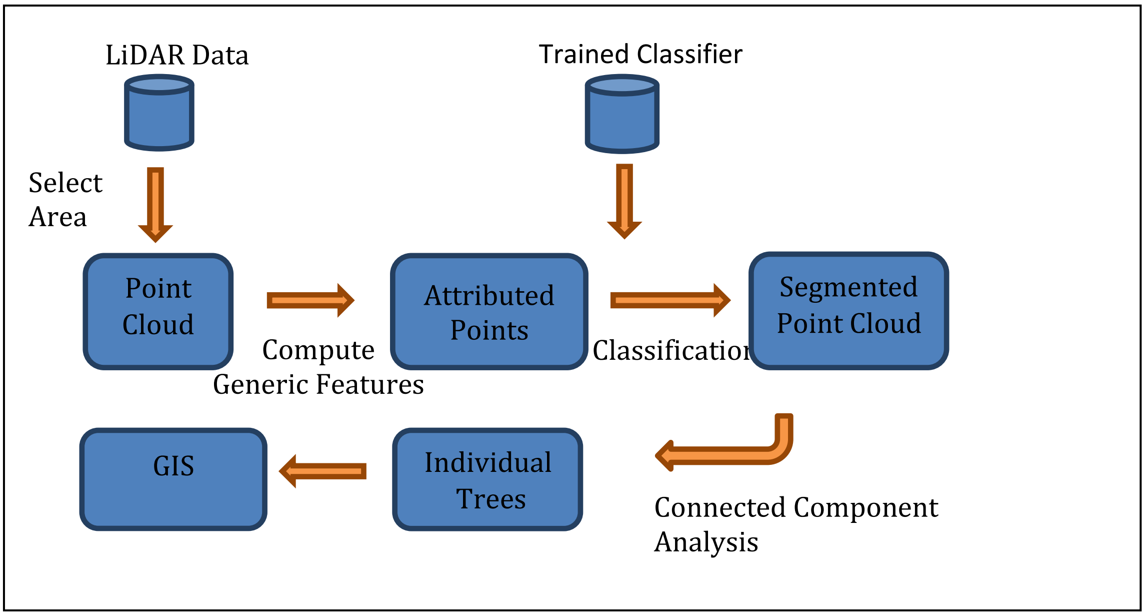


Figure 19. *US2 Workflow*

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.

Preparation component:

Generic system functionalities required:

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April, 2016*
Setting Area of Interest	High	Yes	Implemented	Implemented and integrated

WP4 services required: N/A

WP5 visualization functionalities required: N/A

Feature detection component

Generic system functionalities required: N/A

WP4 services required:

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status I 2015*	*Expected status in April, 2016*
Service # 66 Dimensionality analysis of the input lidar point cloud	High	Yes	Implemented	Implemented and integrated

WP5 visualization functionalities required: see Presentation/Visualization component

Classification & Connected Analysis Component

First alternative

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status I 2015*	*Expected status in April, 2016*
Service # 59 or # 68 Multi-object classification of 3D point clouds (object type must be specified)), in this case: Extract all points sampling trees US2_1	High	Yes	Implemented	Implemented and integrated
Service # 22 Individualization of trees in the classified lidar point cloud US2_2	High	Yes	Implemented	Implemented and integrated

Second alternative

Service name and identifier in eRoom	*Exploitation priority*	*Included in D1.2.3*	*Status I 2015*	*Expected status in April, 2016*
Service # 68 Tree crown recognition from mobile mapping point clouds	High	Yes	Implemented	Implemented and integrated

WP5 visualization functionalities required: see Presentation/Visualization component

Change detection component: N/A

Presentation/visualization component

Generic system functionalities required: N/A
WP4 services required: N/A

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. Data to be visualized:

- List of identified trees
- Associated Point Cloud parts

Output: 3D visualization of Individual tree component extraction / segmented point cloud

Requirement name	*Exploitation priority*	*Included in D1.2.3*	*Status in 2015*	*Expected status in April, 2016*
Visualization of Individual tree component extraction	High	Yes	Not implemented	Implemented and integrated

4 ANNEXES

4.1 SERVICES LIST FOR THE SHOWCASES

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

Components	Exploitation Priority	Planned in D1.2.3	Status in 2015	Expected status in April, 2016
Preparation component				
Generic system functionality				
Import data (upload data and metadata)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI)	High	Yes	Not implemented	Depends on GUI
WP4 requirements				
Service #94: Vector layer partitioning of a point cloud	High	Yes	Implemented	Implemented and integrated
Service #10: Outlier Classification in Point Cloud	High	Yes	Implemented,	Implemented and integrated
Service #35: Resampling of Point Cloud	High	Yes	Implemented	Implemented and integrated
Service #49: Constrained triangulation	High	Yes	Implemented	Implemented and integrated
WP5 visualizaton				
Visualize Vector Layer partitioning of a Point Cloud	High	Yes	Not implemented	Depends of visualization
Feature detection component				
Generic system functionality: -				
WP4 requirements				
Service #48: Multi-resolution triangulation for land monitoring	High	Yes	Implemented	Implemented and integrated
Service #92: Detection of flow lines and drainage basins from triangle meshes	Low	Yes	Not implemented	Implemented and integrated
WP5 visualizaton: see Presentation/ Visualization component				
Change detection component: none				

Presentation/visualization component				
Generic system functionality				
Multi-resolution model visualizer	High	Yes	Not implemented	-
WP4 requirements				
Service #107: Extract the triangle mesh to be visualized at the given LOD	High	Yes	I implemented	Implemented and integrated
Service #49: Constrained triangulation	High	Yes	Implemented	Implemented and integrated
WP5 visualization functionalities required				
Visualize Triangle Mesh	High	Yes	Not implemented	Depends on visualization

Implemented and integrated

Integrated Land Showcase Workflow 2 (LS2): Precipitation analysis

Components	Exploitation Priority	Planned in D1.2.3	Status in 2015	Expected status in April, 2016
Preparation component				
Generic system functionality				
Import data (upload data and metadata)	Essential	Yes	Implemented and integrated	Implemented and integrated
Selection of time interval	High	Yes	Not implemented	Depends on GUI
WP4 requirements				
Service #96: Extraction of rain data	High	Yes	Implemented	Implemented and integrated
Service #40: Approximation by Kriging	High	Yes	Implemented,	Implemented and integrated
Service #67: Approximation by RBF	High	Yes	Implemented	Implemented and integrated
Service #58: Approximation by LR-splines	High	Yes	Implemented	Implemented and integrated
WP5 Visualization				
Visualize precipitation field over the multi-resolution model	High	Yes	Not implemented	Depends on Visualization
Feature detection component				
Generic system functionality: -				
WP4 requirements				
Service #44: Critical Points	High	Yes	Implemented	Implemented and integrated
Service #64: Tracking of critical points	High	Yes	Implemented	Implemented and integrated
WP5 visualization: -				
Change detection component: none				
Presentation/visualization component				
Generic system functionality: -				
WP4 requirements: -				
WP5 visualization::				
Visualize precipitation field over the multi-resolution model	High	Yes	Not implemented	Depends on visualization
Visualize the movement of the maxima of precipitation over the multi-resolution model	High	Yes	Not implemented	Depends on visualization

Implemented and integrated

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

Components	Exploitation Priority	Planned in D1.2.3	Status in 2015	Expected status in April, 2016
Preparation component				
Generic system functionality				
Import data (upload data and metadata)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI)	Low	Yes	Not implemented	Depends on GUI
Area extraction (AOI is extracted from the image)	Low	Yes	Not implemented	Depends on GUI
WP4 requirements				
Service #72: 2D on 2D image registration	High	No	Implemented	Implemented and integrated
Service #70: Radiometric enhancement	Essential	No	Implemented, Hadoop enabled	Implemented and integrated # Service 70or 73
Service #73: Computation of Top Of the Atmosphere reflectance	High	Yes, as part of the Service #73	Implemented and integrated, Hadoop enabled	Implemented and integrated # Service 70or 73
Service #76: Computation of spectral indices	High	Yes	Implemented and integrated, Hadoop enabled	Implemented and integrated
WP5 Visualization: -				
Feature detection component				
Generic system functionality				
Specify reference data (Part of the Service #79: Thematic classification of raster data)	High	Yes	Not implemented	Implemented and integrated
WP4 requirements				
#Service 77: Raster segmentation	High	Yes	Not implemented	Implemented and integrated
# Service 78: Clustering of raster data	High	Yes	Not implemented	Implemented and integrated
# Service 79: Thematic classification of raster data	High	Yes	Implemented and integrated (part of # Service 80)	Implemented and integrated (# Service 79)
# Service 80: Flood and waterlogging detection	High	Yes	Implemented and integrated, Hadoop enabled	Implemented and integrated
# Service 74: Topological Analysis of land 2D data	High	No	Partly implemented	Implemented and integrated
WP5 Visualization: -				

Change detection component: none				
Presentation/visualization component				
Generic system functionality				
Specify RGB colours: The user selects and imports a pre-prepared RGB colour table	Low	Yes	Not implemented	Depends on GUI
WP4 services required: -				
WP5 visualization functionalities required				
Thin client visualization: visualize classification results: - (1) initial raster image (GeoTiff); (2) extracted classification results/thematic rasters; + (3) legend	High	Yes	Not implemented	Depends on visualization

Implemented and integrated

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

Components	Exploitation Priority	Planned in D1.2.3	Status in 2015	Expected status in April, 2016
Preparation component				
Generic system functionality				
Import data (upload data and metadata)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI)	Low	Yes	Not implemented	Depends on GUI
Area extraction (AOI is extracted from the image)	Low	Yes	Not implemented	Depends on GUI
Storage of resulting raster	Low	Yes	Not implemented	Depends on GUI
WP4 requirements				
Service #72: 2D on 2D image registration	High	Yes	Implemented	Implemented and integrated
WP5 Visualization: -				
Feature detection component				
Generic system functionality: -				
WP4 requirements				
Service #76: Computation of spectral indices	High	Yes	Implemented and integrated, Hadoop enabled	Implemented and integrated
Service #69: Raster thresholding	High	Yes	Implemented	Implemented and integrated
Service #71: Convolutional filtering	High	Yes	Implemented	Implemented and integrated
WP5 Visualization: -				
Change detection component				
Generic system functionality: -				
WP4 requirements:				
Service #60: Measuring surface displacement	High	Yes	Implemented and integrated	Implemented and integrated
Service #103: Time-series processing of VHR satellite images for surface deformation detection and measurements	High	No	Partially implemented	Implemented and integrated
WP5 Visualization: -				
Presentation/visualization component				
Generic system functionality:				
Storage of resulting raster and graphics	Low	Yes	Not implemented	Depends on GUI

WP5 visualization:				
Visualization of displacement fields	High	Yes	Implemented in Fat Client	Implemented and intergration in Thin Client

Implemented and integrated

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

Components	Exploitation Priority	Planned in D1.2.3	Status in 2015	Expected status in April, 2016
Preparation component				
Generic system functionality				
Import data (upload data and metadata)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI)	Low	Yes	Not implemented	Depends on GUI
WP4 requirements				
Service #128 Define regular tiling	Low	No New service	Ongoing work	Implemented and integrated
Service #129 Regular tiling Perform regular tiling of data surveys and prepare for approximation	Low	No New service	Ongoing work	Implemented and integrated
Service #122 Tiled point cloud to LR-spline surface Tiled version of service #9	Essential	Yes	Implemented and integrated Are being updated to a multi tile setting.	Implemented and integrated
Service #123 Deconfliction Divides sub parts of the initial data surveys into two groups: keep and remove	Essential	No. New service	Ongoing work	Implemented and integrated
Service #124 Update spline surface Tiled version of service #57	Essential	No	Implemented To be updated to a multi tile setting	Implemented and integrated
Service #125 Trimming of LR B-spline surface Tiled version of service #90	Essential	No	Implemented and integrated. To be updated to a multi tile setting	Implemented and integrated.
Service #85 Stitching of LR B-spline surfaces Modify surfaces generated with respect to a regular tiling to achieve C^1 continuity between the surfaces	Essential	No	Ongoing work. Pre-released exists	Implemented and integrated
WP5 Visualization: -				
Feature detection component				
Change detection component: none				

Presentation/visualization component				
Generic system functionality: -				
WP4 requirements: -				
WP5 Visualization: -				
Fat client visualization: 3D visualization of a combined LR-spline surface consisting of a regular set of surface patches	Essential	Yes, single file version	Ongoing work	Implemented

Implemented and integrated

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

Components	Exploitation Priority	Planned in D1.2.3	Status in 2015	Expected status in April, 2016
Preparation component				
Generic system functionality				
Import data (upload data and metadata)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI)	Low	Yes	Not implemented	Depends on GUI
WP4 requirements: -				
WP5 Visualization: -				
Feature detection component				
Generic system functionality: -				
WP4 requirements				
Service #126 Compare point cloud to LR-spline (Computes the distance between a point cloud and an LR-spline surface) Tiled version of #88	Essential	Yes	Released. Are being updated to tiled data	Implemented and integrated
Service #127 Compare point cloud to LR-spline surface and classify points with respect to distance (Computes the distance between a point cloud and an LR-spline surface) Tiled version of #86	Medium	Yes	Released. To be updated to tiled data.	Implemented and integrated
WP5 Visualization: -				
Change detection component: none				
Presentation/visualization component				
Generic system functionality: -				
WP4 requirements: -				
WP5 visualization				
Distance field visualization	Essential	Yes	Ongoing work to update previous version	Implemented

Implemented and integrated

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

Components	Exploitation Priority	Planned in D1.2.3	Status in 2015	Expected status in April, 2016
Preparation component				
Generic system functionality				
Import data (upload data and metadata)	Essential	Yes	Implemented and integrated	Implemented and integrated
Definition of area of interest (Select AOI)	Low	Yes	Not implemented	Depends on GUI
Area extraction (AOI is extracted from the image)	Low	Yes	Not implemented	Depends on GUI
Storage of resulting raster	Low	Yes	Not implemented	Depends on GUI
WP4 requirements				
Services #9 Generate spline surface from parameterized point cloud.	Essential	Yes	Implemented and integrated	Implemented and integrated
Service #90: Trimming of LR B-spline surface (MS1_1_5). Restrict the valid part of the surface to correspond to the extent of the point cloud by a trimming loop	Essential	No	Implemented and integrated	Implemented and integrated
Service #91: Spline surface to raster	Essential	Yes	Implemented and integrated	Implemented and integrated
Service #72: Resampling and cropping of raster	High	Yes	Implemented	Implemented and integrated
WP5 visualization				
Visualization of the resulting surface raster as greyscale images	Low	Yes	Not implemented	Implemented and integrated
Feature detection component: -				
Change detection component				
Generic system functionality				
Storage of resulting raster	Low	Yes	Not implemented	Depends on GUI
WP4 requirements				
Service #60: Measuring surface displacement	High	Yes	Implemented and integrated	Implemented and integrated
WP5 visualization: -				

Presentation/visualization component				
Generic system functionality				
Storage of resulting raster and graphics	Low	Yes	Not implemented	Depends on GUI
WP4 requirements:-				
WP5 visualization functionalities				
Visualization of displacement fields	High	Yes	Implemented in Fat Client	Implemented and integration in Thin Client

Implemented and integrated

Marine Showcase Workflow 5 (MS5): Lidar WaveForm coastal Feature Extraction

Components	Exploitation Priority	Planned in D1.2.3	Status in 2015	Expected status in April, 2016
Preparation component				
Generic system functionality				
Data load: Import data (upload data and metadata)	Essential	Yes	Implemneted and integrated	Implemented and integrated
WP4 requirements				
Service #111: BWF1_indexing. Performs both sequential and spatial indexing and initial control of the lidar bathymetry survey system data.	Low	No	Implemented	Implemented and integratied in Iqmulus platform
WP5 visualization				
Point cloud visualization	High	Yes	Implemented and integrated	Implemented and integrated
Feature detection component				
Generic system functionality: -				
WP4 requirements				
Service #112: BWF2_WF_attribute_extraction extract parameters from the lidar waveform useful for classification of Seabed bottom types.	Low	No	Pre-release exists Improvements are expected	Implemented and integratied in Iqmulus platform
WP5 visualization: -				
Change detection component: none				
Presentation/visualization component				
Generic system functionality: -				
WP4 requirements: -				
WP5 visualization functionalities				
Visualization of lidar wave form attributes. A color scale visualisation of attribute layer according a point cloud values		Yes, but extended version is needed		

Implemented and integrated

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

Components	Exploitation Priority	Planned in D1.2.3	Status in 2015	Expected status in April, 2016
Preparation component				
Generic system functionality				
Setting Area of Interest	High	Yes	Implemneted	Implemented and integrated
WP4 requirements: -				
WP5 visualization				
Feature detection component				
Generic system functionality: -				
WP4 requirements: -				
Service # 66 Dimensionality analysis of the input lidar point cloud	High	Yes	Implemented	Implemented and integratied
WP5 visualization: -				
Classification and Analysis Component				
First version				
Service # 59 or 68 Multi-object classification of 3D point clouds (object type must be specified)), in this case: Extract all points sampling trees	High	Yes	Implemented	Implemented and integratied
Service # 26 Individualization of trees in the classified lidar point cloud	High	Yes	Implemented	Implemented and integratied
Second version				
Service # 68 Tree crown recognition from mobile mapping point clouds	High	Yes	Implemented	Implemented and integratied
Change detection component: none				
Presentation/visualization component				
Generic system functionality: -				
WP4 requirements: -				
WP5 visualization functionalities				
Visualization of Individual tree component extraction	High	Yes	Not implemented	Depends on visualization

Implemented and integrated