# EOPEN

opEn interOperable Platform for unified access and analysis of Earth observatioN data

H2020-776019

D7.2

Pilots implementation and 2nd prototype evaluation report

Dissemination level:	Public
Contractual date of delivery:	Month 28, 29/02/2020
Actual date of delivery:	Month 29, 24/03/2020
Workpackage:	WP7 Evaluation and pilot implementation
Task:	T7.1 Pilot implementation and evaluation
Type:	Report
Approval Status:	Approved
Version:	6
Number of pages:	117
Filename:	d7.2 Implementation and Evaluation _v6.docx
Abstract	

E O P E N
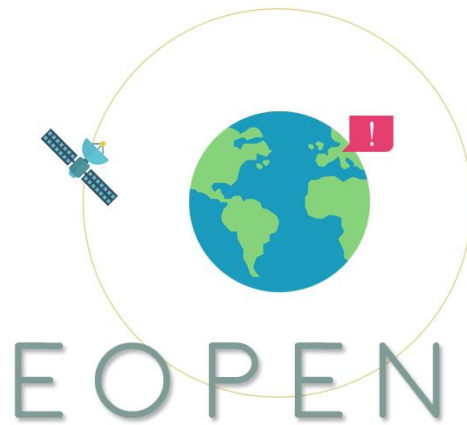
This deliverable reports on the pilots' implementation and the 2nd platform prototype evaluation through the Pilot Use Cases (PUCs). This deliverable is under WP7 and Task 7.1 *Pilot implementation and evaluation* and it is the second of two iterations (D7.1 and D7.2). In this second iteration we report on the established evaluation methodology (D7.1) and on the implementation progress and the evaluation results. The document starts with restating, as in D7.1, the framework of evaluation that defines i) what we measure, ii) who evaluates and iii) which is the implementation-evaluation plan. Then the scientific and technologies activities (Key Results- KRs) of the project are individually reported with respect to their established Key Performance Indicators (KPIs). Subsequently the PUCs report on the current status of implementation and set the timeline for completion. The PUCs are then positioned within the EOPEN framework, elaborating on how its generic modules are utilized in dedicated workflows to address tangible user needs. The test scenarios of evaluation, introduced in D7.1, are finalized. The test scenarios are either fully addressed in D7.2, or a detailed timeline of execution is provided, for scenarios depending on future activities. These will be reported in D7.4.

History

Version	Date	Reason	Revised by	Approved By
1	15/03/2020	Initial Draft		
2	16/03/2020	Further Contributions		
3	17/03/2020	Further Contributions		
4	18/03/2020	Further Contributions	Ioannis Papoutsis	
5	23/03/2020	Review	Gabriella Scarpino, Silvia Scifoni (Serco)	
6	23/03/2020	Final	Vasileios Sitokonstantinou	Ioannis Papoutsis

Author list

Organization	Name	Contact Information
NOA	Vasileios Sitokonstantinou	vsito@noa.gr
NOA	Ioannis Papoutsis	ipapoutsis@noa.gr
NOA	Thanassis Drivas	tdrivas@noa.gr
NOA	Alkis Koukos	akoukos@noa.gr
NOA	Antonis Koutroumpas	akoutroumpas@noa.gr
CERTH	Stelios Andreadis	andreadisst@iti.gr
AAWA	Francesco Zaffanella	francesco.zaffanella@distrettoalpiorientali.it
AAWA	Francesca Lombardo	francesca.lombardo@distrettoalpiorientali.it
AAWA	Gabriele Pastore	gabriele.pastore@distrettoalpiorientali.it
AAWA	Veronica Piffer	veronica.piffer@distrettoalpiorientali.it
FMI	Petteri Karsisto	petteri.karsisto@fmi.fi
SPACEAPPS	Bernard Valentin	bernard.valentin@spaceapplications.com
HLRS	Li Zhong	li.zhong@hlrs.de

Executive Summary

This deliverable, D7.2 *'Pilots implementation and 2nd prototype evaluation report'*, is the second of two parts for the Task 7.1: *'Pilot implementation and evaluation'*. This document builds on top of D7.1 reporting on the implementation progress of EOPEN pilots and the second platform prototype evaluation, both based on the individual Key Results (KRs) and the PUC perspective. This document should be considered as an update of D7.1, following a similar structure.

The document initially reports on the different scientific and technological KRs of the platform. Each KR subchapter includes the scope of the activity, the scientific framework and implementation, the Key Performance Indicators (KPIs) and finally a discussion subsection positioning the KR within the context of the PUCs. The hackathon that took place among the consortia of the Call is then reported as a means of evaluation of the EOPEN platform, in Chapter 5. Then the three different PUCs are individually reported in Chapter 6, including a brief recap of the use case story (D2.1), the timeline of implementation, the monitoring of implementation against the user requirements, the implementation specifics – i.e. the developed products and services and the exploitation of EOPEN KRs within the PUC implementation workflow. Finally, the potential test scenarios, as described in D7.1, are finalized and addressed based on the PUC results and outputs. In certain cases the test scenarios have not been yet exercised, but a detailed plan of execution is reported, clearly elaborated on the means of their fulfilment. All outstanding progress on the implementation and evaluation of the PUCs will be reported in D7.4.

The purpose of the EOPEN Pilot Use Cases (PUCs) is to demonstrate the potential of the 2nd prototype of the EOPEN platform, through case specific workflows, utilizing its components (scientific and technological) to address tangible user needs. The evaluation of the 2nd prototype of the EOPEN platform is performed in the context of the 3 PUCs. The evaluation methodology is designed based on four evaluation axes, as introduced in D7.1, and that is the alignment of implementations with i) the call requirements, ii) the scientific and iii) technological requirements of the project, as specified in the GA, and finally with iv) the collected user requirements, as described in deliverable *D2.2: User Requirements*. In that respect, an additional section, as compared to D7.1, has been added for each PUC, called "Overall Assessment, Lessons learnt and Future plans". In this section, the three PUCs attempt to consolidate their findings, results, evaluations and feedback within the context of the four evaluation axes.

Abbreviations and Acronyms

AOI	Area Of Interest
API	Application Programming Interface
ARI	Adjusted Rand Index
COM	Current Operating Model
DCNN	Deep Convolutional Neural Network
DEM	Digital Elevation Model

DSM	Digital Surface Model
DSM	Digital Surface Model
GA	Grant Agreement
HPC	High Performance Computing
HPDA	High Performance Data Analytics
IoU	Intersection over Union
KPI	Key Performance Indicator
LDA	Latent Dirichlet Allocation
MNDWI	Modified Normalized Difference Water Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NER	Name Entity Recognition
NMI	Normalized Mutual Information
OA	Overall Accuracy
OWL-DL	Ontology Web Language Description Language
PA	Producer's Accuracy
PSRI	Plant Senescence Reflectance Index
RDF	Resource Description Framework
RF	Random Forests
RNN	Recurrent Neural Networks
SAR	Synthetic Aperture Radar
SWIR	Short-Wave InfraRed
TOM	Target Operating Model
UA	User's Accuracy
WFS	Web Feature Service
WMS	Web Map Service
WMTS	Web Map Tile Service

Table of Contents

1 INTRODUCTION

The document is structured as follows. In the second chapter a high level description of the evaluation methodology is provided, setting the questions to be answered in the coming chapters, i.e. of *what we measure, who measures, how we measure.* According to the GA, EOPEN brings in certain scientific and technological activities, established as Key Results (KRs) of the project. The third and fourth chapters elaborate on all scientific and technological activities with respect to their measurable performance indicators but also their exploitation within the context of the PUCs. In the fifth chapter elaborate on the findings and conclusion of the Hackathon and employ those as means of evaluation for the EOPEN solution. In the sixth chapter, we introduce each PUC individually, stating the respective use case stories and the problem to be tackled; and reporting on the implementation progress. Evaluation results are also provided, along with a list of test scenarios. Finally, in chapter six we report on the overall assessment of each PUC, the lessons learnt and any future plans. In this section we will also provide the high level qualitative and quantitative evaluation of the EOPEN platform against the requirements of the call, as seen in the evaluation framework described in Chapter 2.

It should be noted that the three PUCs are rather different and address their individual challenges from alternative perspectives. Those different approaches/perspectives attempt, in a complementary fashion, to showcase the value of the EOPEN platform and its constituent components. For instance, PUC1, being led by the primary stakeholder (AAWA), highlights the **decision making and governance** perspective, supported by the EOPEN solution and its JDIG tool. Complementarily, PUC2 focuses more on the **big data and distributed processing** perspective, effectively utilizing the HPDA infrastructure, to develop large-scale applications as food security monitoring demands. Finally, PUC3 focuses more on the **data visualization and data analytics** dimensions, showcasing the value of the EOPEN User Portal functionalities. For this reason, the evaluation of the 2nd prototype, through the PUCs, should be viewed holistically.

2 EVALUATION FRAMEWORK

The overall evaluation framework for the 1st and 2nd prototype of the platform and the PUCs is elaborated below. Part of this holistic framework and the initial implementation was addressed in D7.1, while this document reports on the implementation progress and concludes the evaluation methodology for each PUC. In order to perform the evaluation of the EOPEN platform we identify the measures, the tools and the people, as described below.

Evaluation measures – Quantitative and qualitative measures based on the described KPIs in the GA and the collected user requirements. Below are described the four main axes of evaluation.

High level Call requirements – The completed PUCs evaluate the EOPEN platform at a higher level, identifying its key competences – showcased through the PUCs - and how these explicitly address the call requirements, such as the ones indicated below.

Enable value adding services on generic data and information storage and processing facilities which can allow public and commercial users effective production environment to interact with and serve their user base without deploying their own fstorage and processing facilities.

Make access to the Copernicus data and information easy and user friendly through scalable dissemination and exploitation software based on international standards.

Foster the establishment of interoperable access facilities to all EU Member States.

Provide user community tools including best-practices.

Optimize the use of Copernicus data by non-traditional user communities to meet societal challenges.

Scientific requirements – Showcase how the following scientific activities of the EOPEN platform were exploited for the purposes of the PUCs. Deliverables 7.1 and 7.2 report on the individual implementations of scientific KRs but also establish and monitor pertinent performance indicators.

- Knowledge discovery and content extraction
- Change detection in EO data
- Concept and Event Detection in Social Media
- Multimodal clustering and retrieval
- Multimodal fusion from diverse data sources for information retrieval
- Clustering of EO and non EO data
- Community detection and tracking
- Animation of the network of user communities
- Knowledge representation and ontology construction for decision making
- Ontology construction and reasoning support
- Linking of open EO data

Technological requirements – Showcase how the following technological activities of the EOPEN platform were exploited for the purposes of the PUCs. Deliverables 7.1 and 7.2 report on the individual implementations of technological KRs but also establish and monitor pertinent performance indicators.

- EO and non-EO data acquisition
- EO data acquisition

- Social media crawling
- Meteo data acquisition
- Standardized interfaces
- High Performance Computing
- Fast and Secure Data Exchange
- EOPEN platform
- DIAS - EOPEN would represent a typical Third Party User interested to exploit the DIAS services to design and operate a Front End EOPEN service.

PUC Requirements

PUC Implementation and Evaluation

Objectives – problem statement and suggested EOPEN solution

Users involved – main PUC stakeholders and their role in implementation and/or evaluation

Timeline of implementation – record achieved and monitor forthcoming milestones

Test scenarios – establish testing and evaluation scenarios to take place during the pilot phase of PUCs

Feedback mechanism – introduce the methods of receiving feedback from the EOPEN platform users and the pertinent scheduled events (i.e. workshops, info days etc.)

Evaluators

Self-assessment – This refers to the performance indicators, qualitative and quantitative, set and monitored by the developers of the product/module/service in question.

Cross-partner assessment (peer) – Partners of the consortium that use but have not developed the Key Result in question evaluate their experience. This can be done either by re-evaluating the self-assessment KPIs, if applicable, or setting other relevant, more user oriented ones.

EOPEN user ecosystem – The end users partners of the consortium (AAWA) and the PUC engaged stakeholders will evaluate the EOPEN platform through their feedback in the form of questionnaires, Q&A sessions, deliverables (in the case of AAWA) etc. that will be based on platform demonstrations and/or hands-on exercises.

3rd party users

External users – This refers to non PUC specific users of the platform. Their feedback will be extracted in the dedicated dissemination events of the project, such as the South Korea ACRS workshop and 'Infoday' of PUC1, Other projects of the H2020-EO-2017 call: A hackathon was held (7-8/11/2019) among five consortia to test the individual platforms (EOPEN, BETTER, CANDELA, openEO, Perceptive Sentinel).

Implementation - Evaluation plan

Two iterations

Iteration 1 – The first deliverable (D7.1) of Task 7.1 focused more on the self-assessment of the individual scientific and technological KRs of the platform and established the overall evaluation methodology (M20)

Iteration 2 – D7.2: The second deliverable (D7.2) of Task 7.1 is more PUC oriented, strengthening the evaluation from the PUC perspective, while fulfilling the overall evaluation framework established in D7.1.

3 SCIENTIFIC ACTIVITIES

Key Result evaluation and establishment of KPI[1]s to monitor

3.1 Change detection in EO data (KR01)

3.1.1 Purpose of Scientific Activity

The Remote Sensing (RS) based Change Detection module's aim is to identify differences in the state of land features for the detection of flooded areas (PUC1), rice fields (PUC2) and of passable roads (PUC1, PUC3) by considering Satellite images either of Optical or Radar origin (mainly Sentinel-1 and Sentinel-2). Some initial information about Change Detection has been included in D4.1, while more details will follow in D4.4.

3.1.2 Description of the scientific framework

As described in D7.1, a baseline approach for the first module has been developed, directly estimating the water-bodies of an input satellite image (PUC1). The second module (PUC1, PUC3) takes as input a Sentinel-2 image inferring whether a road is blocked or not by major flood or snow, i.e. passable or impassable. The third module (PUC2) takes as input Sentinel-1 images to update past land cover maps (rice maps in South Korea) to be used as training datasets for any given inspection year.

In a new approach of the first module (PUC1), elevation information is being used to improve the accuracy. This version of the water detection module has been successfully deployed as a process on the EOPEN platform and is currently running, generating water-masks of the Vicenza area, that are shown on the website using a Web Map Server. Also, a flood detection extension for this module has been developed that is able to detect the existence of a flood incident within a time-series of Sentinel-2 images. The versions of both the road passability and flood detection modules are running on CERTH's side for the moment.

Module 1

For the first module, (PUC1 – water detection), in order to estimate the water-bodies areas, two different approaches were initially developed. The first implementation utilized Sentinel-2 data and the Modification of normalized difference water index (MNDWI), based on the Green and SWIR bands. The second implementation used Sentinel-1 data. In this case, preprocessing via SNAP was required. A thresholding technique on multiple tiles of the image helped at the discrimination of water from non-water areas. Finally, a GeoTiff raster with the water-bodies mask was generated.

The improved version of this first module is using Digital Elevation Model (DEM) files, paired with both the VV and VH bands in a small Deep Neural Network of 3 layers, trained from scratch that is filtering the falsely detected water areas due to high slope. Various neural network architectures have been tested with the VGG-19 demonstrating the best

---

[1] The KPI tables in Section 2 include the field 'relevance to user requirements'. The user requirement IDs included can be found in the ANNEX of D2.2., which has been updated in August 2019; also, updated PUC EOPEN product requirements are described in D1.4 which is a living document

performance. A future version of this methodology will incorporate extra features in the model (e.g. slope and incidence angle) improving the performance when applied to more complex territories.

For the flood detection extension of the first module (PUC1), a time-series of optical imagery is input. The differences of consecutive days are fed to the DNN model and a flood incident is predicted. The VGG-16 architecture demonstrated the best results. An alteration of this methodology will extend to support Sentinel-1 images in order to surpass extreme ongoing weather conditions.

Module 2

In the implementation of the road passability module (PUC1, PUC3), a pre-trained VGG-19 DCNN model was used to further train the prediction model. New coming images were fed to the model to infer whether they are passable or not. In the future, this methodology will be extended to support Sentinel-1 images in order to surpass extreme ongoing weather conditions.

Module 3

In the implementation of the third module (PUC2), the VH backscatter (Sentinel-1) difference between the year of inspection and the year of the land cover map release is used. It is then combined with the water mask produced by the first module to remove obvious outliers, namely parcels of the rice map that are not cultivating rice in the year of inspection. The output product of this change detection methodology does not attempt to accurately classify rice fields for 2018 but merely delete obvious changes in the land cover map of 2015. Therefore, using only rice pixels from the 2015 land cover map we eliminate outliers for the updated 2018 land cover map.

All modules are driven by a process (Umbrella API) that searches for Sentinel-1 and Sentinel-2 products for the dates and areas of interest. For more information on the implementation of all three modules can be found in D4.1.

3.1.3 Results and metrics

Module 1

Three Lakes in Italy were monitored (Garda, Maggiore, Trasimeno). Due to lack of existing ground-truth in the examined PUCs, a first step towards the evaluation was to generate ground-truth data through manual annotation. The shapefiles of the permanent Lakes and Rivers of Italy were used as a base for the annotation. High resolution images of Open Street Maps, Google Maps and Bing Maps were used to correct the initial shapefiles. Evaluation results for the flood assessment of the improved module can be found in Table 1.

For the fine-tuning of the pre-trained VGG-16 DCNN model 267 events in urban African cities were used as a training set. Then 68 real events were used to test the performance of the model. Each event described a different area of urban African cities and consisted of a time-series of Sentinel-2 images of variable size of that area. Results, as provided by MediaEval 2019 Satellite Task "Emergency Response for Flooding Events", can be found in Table 2.

Module 2

For the evaluation of the road passability module (PUC3), multiple pre-trained models were tested, with VGG-19 demonstrating the best overall performance (Table 3), which is 76.6% for the validation set and can be considered a satisfactory result.

Table 1: Water bodies delineation metrics using DNN approach that exploits backscatter and elevation information

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Precision	95,51%	US_V_02 US_V_02b	The model avoids misclassifying water areas as non-water.
Recall	88,69%		The model detects many of the water areas
F-score	91,93%		The model performs well for imbalanced classes (water and non-water)
Time performance	millisecond (ms)		This KPI will be examined in the future.

Table 2: Metrics for F-score in detection of flood events on MediaEval 2019 test set using the pre-trained VGG-16 DCNN model

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
F-score	70.58%	US_V_02 US_V_02b	The model performs adequately for imbalanced classes (water and non-water)
Time performance	millisecond (ms)		This KPI will be examined in the future.

Table 3: Two of the best Neural Networks achieved validation set accuracy for road passability module

DCNN	Learning Rate	Optimizer	Batch Size	Dev. Set Acc.	Valid. Set Acc.
VGG-19	0,01	Adam	128	0,8610	0,7667
ResNet-50	0,001	Adam	256	0,8820	0,7323

Module 3

To evaluate the third module (land cover map update - PUC2), the total number of correctly classified and misclassified pixels of the updated land cover map against a validation dataset (through photo-interpretation) have been calculated. Classification results are presented in Table 4.

Table 4: Updated land cover map to represent the year of inspection

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Precision	98.35%	US_SK_03b US_SK_03c US_SK_04	
Recall	92.28%		
F-score	95.21%		

3.1.4 Discussion

This KR identifies variations of the landscape morphology by exploiting the available Copernicus products for the selected locations. In the future, the selected locations will be defined via search criteria thus allowing the end users to monitor new locations.
This KR is evaluated through self-assessment, as presented above, and at the same time there is a cross-partner assessment evaluation as PUCs make use of it. Specifically, as far as PUC1 is concerned, it allows the monitoring and identification of changes of the landscape due to floods, by developing a water body change detection module.

Moreover, it addresses the needs of PUC2 via preparing refined training datasets for rice classification. Finally, it addresses the needs of PUC3 and PUC1, which involve the monitoring of road conditions during extreme weather events, by developing techniques for road passability due to snow coverage, melted snow and water floods. More information on the utilization of the three modules, in the different PUCs, can be found in Chapter 6.

3.2 Event detection in social media (KR02)

3.2.1 Purpose of Scientific Activity

The Event Detection module aims to identify possible events based on non-EO data. The current implementation focuses on the crawled tweets, derived from the social media monitoring task, and uses the fluctuation of the number of relevant tweets per day to discover events. In the first implementation of the module, a respective notification is produced, but the next version will provide more insights on each event, such as the location of the happening and some representative keywords. The first version of the module runs daily as a process in the EOPEN platform (High level Call Requirement I) and generates JSON files with notifications for any detected events. Some initial information about Event Detection has been included in D4.1, while the complete description will follow in D4.4.

3.2.2 Description of the scientific framework

In order to determine whether an event occurs in a given date, we calculate the *z-score* of that date's number of relevant tweets. The z-score indicates how many standard deviations this day is from the sample's mean, assuming a Gaussian distribution, where sample is the previous thirty days (previous month). If a date's z-score is above a constant threshold, then we can assume there is an event on that day. Bibliography suggests a threshold of 2.5-3.5 and after internal experiments we have settled to a threshold of 3.

3.2.3 Results and metrics

At the time of writing the evaluation of the Event Detection module is in progress. The event detection methodology will be evaluated by running on the collected Italian tweets that speak about floods and by using as ground truth the dates of official flooding incidents in Italy in the years 2017 and 2018, provided by AAWA. In addition, more variations of the algorithm will be evaluated, which concern different thresholds for the z-score and different number of days to be compared. The key performance indicators that will be calculated can be seen in the following table, while the experiments will be reported in D4.4.

Table 5: Performance indicators and scientific metrics for event detection in social media

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Precision	%	US_V_02a4	

Recall	%	US_V_02a5	
F-score	%	US_V_08a	

3.2.4 Discussion

This module discovers possible events exploiting the social media data that has been crawled by the corresponding task. Each dataset is use case-specific by defining different search criteria during the crawling procedure, so events are detected per each use case separately and independently. Every end user that is interested in a certain use case will be only notified about events relevant to that use case. Notifications for PUCs 1 and 3 can serve as early warning for incidents linked to extreme weather conditions, such as floods or heavy snow, while notifications for PUC2 highlight changes in the food industry and encourage interested users to investigate what is commented on social media (Cross-partner assessment).

3.3 Concept extraction in social media (KR03)

3.3.1 Purpose of Scientific Activity

The Concept Detection module aims to extract high-level content (i.e. concepts) from visual low-level information in order to be able to retrieve relevant content and to mark multimodal content as relevant or not to a target event (i.e. flood, heatwave). The current implementation focuses on extracting concepts from the images included in the crawled tweets, derived from the social media monitoring task. This first version of the module runs on CERTH's side, but all the information is stored in a database that the platform has access. Some initial information about Concept Detection has been included in D4.1, while more details will follow in D4.4.

3.3.2 Description of the scientific framework

The initial implementation was based on a framework that used Caffe (Jia et al., 2014) and involved the use of a fine-tuned 22-layer GoogleNet network on 345 SIN TRECVID concepts. However, the framework was implemented in a Windows environment that renders impossible to be transferred to the EOPEN platform. Thus, the progress related to the concept detection module involves the transfer of the module to the EOPEN platform. Subsequently this involves the development of a new framework and the testing of several CNN architectures (i.e. VGG, GoogleNet, ResNet).

3.3.3 Results and metrics

The proposed concept detection methodology will be evaluated, similarly to the initial implementation, in TRECVID datasets but not inside the EOPEN framework, because it requires annotation of a significant large number of images and concepts. Therefore, given that the manual annotation of an EOPEN image collection is labour-intensive, the evaluation will be limited on existing annotated datasets such as TRECVID.

Table 6: Performance indicators and scientific metrics for concept extraction in social media

Indicator/Scientific metric	Value	Relevance to User Requirements
Precision	%	US_V_02a1 US_V_02a2 US_V_02a5 US_F_01b1 US_F_01b3
Accuracy	%	

3.3.4 Discussion

This module assigns high level concepts to images which are collected by the social media crawler (KR12), and it allows filtering of tweets according to specific concepts (i.e. flood, snow) by using the tweet images. Given that the content of social media images differs significantly per PUC, this module allows the retrieval of visually similar images, and thus images of a specific PUC.

3.4 Similarity fusion (KR04)

3.4.1 Purpose of Scientific Activity

The Similarity Fusion module aims to retrieve similar content found inside the EOPEN collection. Given that the module considers different modalities for retrieving similar EO and non-EO content, there is the need to fuse these multimodal objects in a scalable way, taking into account memory and computational complexity, in order to retrieve similar content in response to a query. The current implementation is a single modality retrieval module for non-EO data that considers visual, textual, spatiotemporal, and concepts. The next steps involve fusion of the aforementioned information by considering late fusion techniques, the application of the approach to EO data by considering only the related modalities (i.e. spatiotemporal, concept) as well the transfer of the service to the EOPEN platform. This first version of the module runs on CERTH's side. Some initial information about Similarity Fusion has been included in D4.1, while more details will follow in D4.3.

3.4.2 Description of the scientific framework

The single modalities retrieval modules that were developed and presented in D7.1 (M20). In the meantime, the similarity fusion approach that was considered is a late fusion approach that is able to take into account all the available modalities. Thus, some single modality modules were updated and some new were developed including:

· Spatiotemporal retrieval: This module involves the retrieval the closest tweets in terms of time and location and it considers the metadata either pre-existing in the tweet or produced by other EOPEN modules (i.e. Localization module).

· Visual concept-based retrieval: This module produces apart from the top 10 concepts a 345 vector with the probabilities of the 345 SIN TRECVID concepts. Then, similarly to the visual features, an index is created that considers these vectors and creates and updates on a regular basis an Asymmetric Distance Computation index to achieve fast retrieval.

However, several changes are expected to be realized in the single modality modules in the upcoming period which will subsequently affect the similarity fusion module. Specifically:

· Text-based retrieval: The already existing approach which involves word2vec requires continuous update of the text index which is time-consuming. Therefore, the use of an Apache Lucene for text retrieval and indexing is opted.

· Visual concept-based and visual features-based retrieval: Both modules consider the same DNN network, which is a fine-tuned GoogleNet. Due to the update of the CNN architecture (mentioned in the "Concept extraction in social media" section), changes might be realized in the size of the feature vector of the visual features and concepts vectors.

3.4.3 Results and metrics

At the time of writing, the evaluation of the Similarity Fusion module in the EOPEN dataset has not been carried out yet. In order to evaluate it in EOPEN, ground-truth and domain-specific datasets in the examined PUCs are required, and thus a first step towards the evaluation is to generate ground-truth annotations through manual annotation. Afterwards, the following key performance indicators will be calculated:

Table 7: Performance indicators and scientific metrics for similarity fusion

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Precision	%	US_V_02a2	
Recall	%	US_V_02b3	
F-score	%		
Processing time	ms		
Memory complexity	GB		

3.4.4 Discussion

This module allows the user to retrieve similar EO and non-EO content given a query tweet or EO product. Given that the tweet content, which involves both the text and image included (if it exists) and the EO products are rather different per UC, the similar content that will be retrieved will thus differ per UC.

3.5   Clustering of EO data (KR05)

3.5.1   Purpose of Scientific Activity

The initial purpose of this module is to cluster EO data into groups so that end users can find the interested patterns. However, after further research, this module has started by considering the clustering of Sentinel-2 images time-series to automatically identify rice fields and create labelled data for supervised classification (PUC2) - the algorithm has been developed by NOA, with the support of HLRS. Discussions are under way to exploit such implementation, in PUC1 context which will be documented in future deliverables (D7.4).

3.5.2   Description of the scientific framework

The module for EO data clustering adopts the technology of k-means and w-net. W-net is a deep neural network which is composed of a two-stage U-net and organized in an AutoEncoder-Decoder architecture. Through this model, the EO images can be automatically segmented based on the clustering of pixels and further the rice fields could be identified.

Within the context of PUC2, clustering was used for the generation of a training dataset. The idea is to exploit vegetation indices (NDVI, NDWI, PSRI) calculated from Sentinel-2 times-series imagery as input to a distributed K-Means algorithm (High level Call Requirement I). Temporal interpolation was required to allow generalization and transferability of the model so that it does not depend on the time/date of acquisition; since Sentinel-2 satellites visit each area on different timestamps. Thus, each pixel was assigned to the weighted average value within a fixed time step of a 10-day window. Linear interpolation was also used to derive interpolated values for the pixels classified as cloud, according to the Scene Classification product. The first iteration of K-Means results in two clusters which were used as a water mask. Having extracted pixels considered to be water, a second-level group of K-Means executions takes place for creating a variant number (5-15) of clusters. Each of these clusters is assigned to different observations. Thus, the identification of the rice cluster on each K-Means execution along with the final number of clusters are based on certain metrics, such as Mean Squared Error, Recall and Precision compared to a validation dataset.

3.5.3   Results and metric

The following metrics were used to evaluate the performance (Self-assessment). The results for rice clustering to create training labels is evaluated on its own. Nonetheless, in the PUC2 section we report on the performance of the supervised classification using the k-means based labels.

Table 8: Performance indicators and scientific metrics for Rice Clustering to create labelled data

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments

Precision	97.90%	US_SK_03b	
Recall	89.98%	US_SK_03c	
F-score	93.78%	US_SK_04	
Reduce manual effort for labelling data	3 Person Days	US_SK_10	Manual annotation took 3 PDs, while clustering runs in only several minutes

3.5.4 Discussion

Clustering algorithms are initially developed for defining groups of similar images. However, in image segmentation it is also widely used, especially when clustering the pixels of an image. Towards this direction and using unsupervised machine learning techniques, a set of labelled rice pixels at Seosan and Dangjin districts of South Korea are produced.

3.6 Text clustering of non-EO data (KR06)

3.6.1 Purpose of Scientific Activity

The purpose of this module is to cluster textual information from crawled tweets into groups, which can be perceived as topics that are trending on social media. Currently the module has been implemented, deployed as a service on the EOPEN platform and is available to be called at any time (High level Call Requirement I) . The service receives as an input the use case of interest (e.g., Italian tweets about flood events) and responds with the clusters detected in the 300 most recent relevant tweets, so as to capture the trending topics. Each topic has the tweets it comprises and some keywords that reflect the most common terms mentioned in these tweets. A complete description of the Text Clustering module can be found in D4.2, while any updates will be added to D4.4.

3.6.2 Description of the scientific framework

The module for clustering textual streams of data uses a hybrid density-based framework, which combines the DBSCAN-Martingale process and the well-established Latent Dirichlet Allocation (LDA). After extracting the word *n*-grams of the recent tweets, the robust to noise DBSCAN-Martingale progressively estimates the number of clusters in the given dataset and LDA assigns the tweets to topics.

3.6.3 Results and metrics

The key performance indicators to be used for the evaluation of the proposed text clustering methodology are (Self-assessment):

Table 9: Performance indicators and scientific metrics for text clustering of non-EO data

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments

Normalized Mutual Information (NMI)	%	US_V_02a3 US_V_02a5	This KPI will be examined in the future.
Processing Time	millisecond (ms) or hours (h)		This KPI is examined in relation to the number of cores used in a parallelized implementation.
User-based evaluation	Likert scale (1-5)		This KPI will be examined in the future.

As reported in D4.2, the processing time of a parallelized (multi-core) version of the algorithm has been calculated for different numbers of cores. The results can be summarized in the following table and demonstrate that processing time significantly decreases while the number of cores increases, proving that text clustering is a problem that can benefit from parallelization.

Table 10: Processing Time related to number of cores

Number of cores	Processing Time (h)
1	85.00
2	46.96
4	24.85
8	13.19
16	7.08
32	3.75

Another experiment is currently in progress to calculate the NMI score. Using English tweets that refer to floods, we have created ground truth annotation by making the assumption that tweets that mention the same location on close dates belong to the same cluster. Additionally, we are going to investigate whether clustering can improve the results of the localization methodology (please refer to *Ontology construction and reasoning support framework*).

Finally, through questionnaires addressed to end users of the EOPEN platform (EOPEN ecosystem evaluation and 3rd party users evaluation), we will be able to evaluate how beneficial the developed clustering techniques are to the users.

3.6.4 Discussion

When calling the service for text clustering, the use case has to be defined as input. Then, the module uses tweets that have been crawled specifically for this use case and returns the detected topics. End users have the ability to call the service for the use case they are interested in and get informed about the relevant trending topics on social media. Thus, it can be used for all three PUCs (Cross-partner assessment).

An example for PUC1 follows. The end user is notified by the Event Detection module (KR02) that a potential event is happening with regards to floods, along with some first details, e.g. possible area of the event. In order to gain more insight, the user runs the Topic Detection module and is presented with trending topics, in the visually appealing form of word clouds. Each word cloud contains keywords that explain the topic, e.g. affected places, type and severity of the incident. If the user desires to get more in depth, clicking on the word cloud shows the complete list of the tweets of the cluster, providing access to the original, raw information. In such a way, the end user is able to view what is happening on social media from high-level to low-level information. A very similar scenario can be described for PUC3, where the main subject is the snow coverage instead of flooding incidents.

3.7 Community detection (KR07)

3.7.1 Purpose of Scientific Activity

The objective of this module is to detect end-user communities through their relationship. The current version of the module focuses on social media accounts and their "following" interaction, while it can be easily adapted to focus on users of the EOPEN platform and the similarity of their behaviour. The module has been deployed on the EOPEN platform (High level Call Requirement I) as a service and is currently online. The inputs to the service are the use case of interest and a date range. The response includes the pairs of connected users, the detected communities, and a list of the most influential users. The Community Detection module has been described briefly in D4.1, but further information will follow in D4.4.

3.7.2 Description of the scientific framework

The methodology of this module serves two functionalities: the detection of key-communities and the identification of key-players. After mapping users and their relationships to nodes and edges, the Louvain algorithm is adopted to discover communities (Blondel et al., 2008). On the detected community of nodes, the degree measure is computed per node, defined as the number of its neighbours. The nodes with the largest degree are the key-players.

3.7.3 Results and metrics

The metrics to evaluate (Self-assessment) the Concept Detection method can be seen in the following table and include the modularity of networks, i.e. the strength of division of a network into communities, the time the process needs to run, and the usefulness of the module according to end users.

Table 11: Performance indicators and scientific metrics for community detection

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Modularity	[-1,1]	US_V_02a US_F_01b	This KPI is examined in relation to the use cases and the length of time period.
Processing Time	millisecond (ms)		This KPI is examined in relation to the length of time period.
User-based evaluation	Likert scale (1-5)		This KPI will be examined in the future.

A first experiment concerned the detection of communities in Twitter accounts that mention each other through tweets about floods in Italy. The scope was to observe how modularity and processing time change according to the length of the examined time period, i.e. the number of days in which the tweets were posted. The results are displayed in the next two figures.
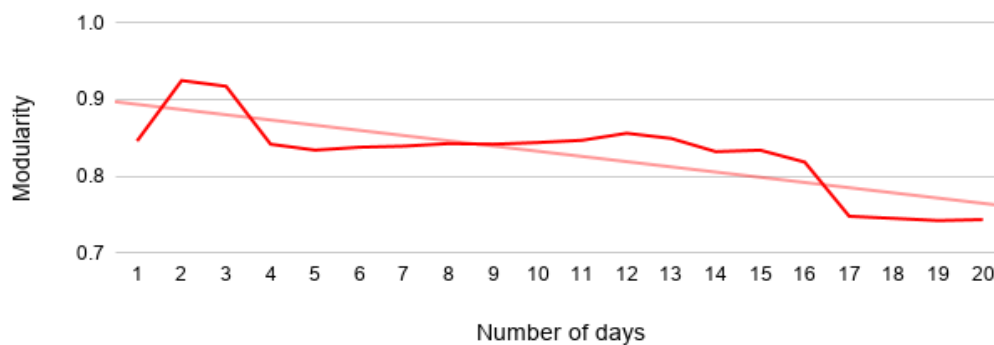
Figure 1: Modularity of community detection for different length of examined time period
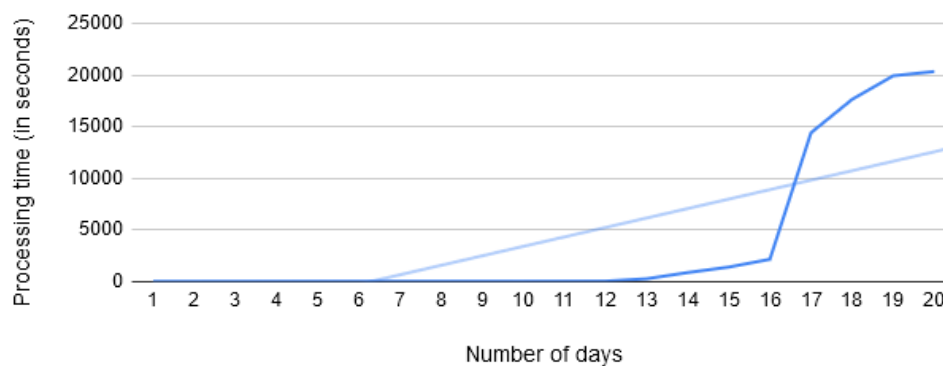
Figure 2: Processing time of community detection for different length of examined time period

Apart from a peak in the early increase of the number of days, modularity varies between 0.819 and 0.856 for most of the different time periods, but starts to decrease when the period exceeds 16 days. This can be interpreted that strong communities in social media tend to be formed between one day and two weeks, while after that period they begin to lose their strength. The mean modularity for all time periods is 0.829, indicating a high performance of the proposed community detection methodology.

Regarding processing time, it takes only a few seconds to detect communities based on tweets posted between one and sixteen days, but then it increases exponentially and requires hours to be completed. For this reason, a repetition of the experiment is planned to be realized on an HPDA infrastructure, in order to examine processing time in a distributed environment and evaluate whether parallelization would benefit the proposed methodology. It should be noted here that the delay is not caused by the detection of communities but by the identification of key-players.

A second experiment that took place involved all the use cases examined in EOPEN to evaluate how modularity changes for different domains and languages. The time period was selected to be ten days for all cases. As seen in the next figure, stronger communities are formed in the use cases that refer to weather incidents (flooding, snow coverage) rather than to food security, which is a more generic topic.

Figure 3: Modularity of community for different use cases

Additionally to the above evaluation, future questionnaires addressed to end users of the EOPEN platform will be able to provide the users' opinion on the visualization of communities (EOPEN user ecosystem evaluation and 3rd party users evaluation).

3.7.4 Discussion

The implemented service searches for formed communities inside a network of user-to-user interactions, where users are social media accounts whose posts have been collected for EOPEN. Since the collections of crawled tweets are use case-based, the results of the service refer to specific use cases (Cross-partner assessment). For PUC1 and PUC3 that concern critical situations due to floods and snow respectively, the identification of communities and key players that are in the centre of these communities leads to the detection of accounts that function as "authorities", i.e they inform/influence other citizens through social media during a crisis of extreme weather conditions. On the other hand, for PUC2 these "authorities" accounts can be further investigated by the end users with regards to their influential role in matters of food security.

3.8 Ontology construction and reasoning support framework (KR08)

3.8.1 Purpose of Scientific Activity

The developed framework provides the appropriate semantic knowledge structures and vocabularies to represent metadata coming from EOPEN modules and other services, like the localisation module, and build intelligent reasoning mechanisms and decision making solutions to extract implicit conceptualisations and interpretations. In the current version of the framework the metadata that are supported come from the localisation, topic identification and the event extraction modules. These metadata are transformed from JSON into RDF format and saved in the Knowledge Base. This version of the framework, as well as the Knowledge Base, has been successfully deployed in the EOPEN platform as dockerised services (High level Call Requirement I and II). More information about the module can be found in D5.1.

3.8.2 Description of the scientific framework

The Ontology construction and reasoning support framework is responsible for:

- generating the appropriate data mapping

- populating the Knowledge Base with the information

- enriching the results using the reasoning module

- retrieving the results using the semantic queries

Furthermore, the localisation module is responsible for:

- retrieving location and organisation entities that appear in social media posts

3.8.3 Results and metrics

Since the semantic framework has not been finalized, some metrics have not yet been utilised. The completeness and conciseness, and the consistency are taken into account at the time that we develop the ontology. The consistency, the pitfalls and the completeness and conciseness are part of the measurement of the quality of the ontology, while the accuracy and precision are associated with the reasoning framework (Self-assessment).(Self-assessment Evaluation). The response time of the framework will also be taken into account (Table 12). Lastly, to measure the performance of the English/ Italian location recognition metadata models and the respective named entity recognition (NER) configurations, values for the precision, recall and F1-score measures were computed (Table 13,Table 14). During the reporting period, important updates have been made to the NER pipeline, and testing of newer transformer-based neural network model implementations for English have improved past results (Table 14). These models are currently being tested with the Italian and Finnish languages for the purposes of PUC1.

Table 12: Ontology-related metrics

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Completeness and Conciseness	Number	PUC1_GA2, PUC1_GA3, PUC1_GA7, PUC2_GB4, PUC3_GC2, PUC3_GC5, PUC3_GC7, PUC3_GC12	The ontology follows the competency questions
Consistency	Number		The ontology is consistent
Pitfalls	String (Critical, important, minor)		
Response time	Number of seconds		1.01100 seconds approximately

Table 13: Performance evaluation of the proposed NER system (EN) vs. the baseline system and a state-of-the-art approach

System (CoNLL2003 dataset)	Precision	Recall	F1-score
Our system in D7.1	90.95	90.94	90.97
Our updated system for D7.2	91.46	92.32	91.88
Best-scoring shared task system: Florian et al., 2003	88.99	88.54	88.76
Baevski, A. et al. 2019	(not reported)	(not reported)	93.5

Table *13* is indicative of the incremental gains in F1-score after employing the new transformer-based (BERT) model. Further improvements will be made during the next reporting period with parameter tweaking/fine-tuning, and ELMO-based model testing.

Table 14: Performance evaluation of the proposed NER system (IT) vs. the baseline system and a state-of-the-art approach

System (EVALITA2009 dataset)	Precision	Recall	F1-score
Our system	75.49	75.60	75.37

Best-scoring shared task system: FBK_ZanoliPianta	84.07	80.02	82.00
Nguyen and Moschitti, 2012	85.99	82.73	84.33

The Italian model has only received slight alterations and as such, results are identical with the ones reported in D7.1. Focus has been placed on updating the relevant available dataset with additional data in order to improve the model's training material. Also, updating the IOB annotation format to a more recent one, like BILOU or BMEWO, is certain to result in F1-score amelioration. Further result reporting will be available in the respective future deliverable (D5.2), due in M33.

3.8.4   Discussion

This module saves metadata from topic extraction in the Knowledge Base. The metadata is converted in semantic format and connected in terms of a semantic graph for all PUCs. Also, semantically similar terms are found in order to detect the similarity among different topics. The localisation module's results benefit all PUCs (Cross-partner assessment), with emphasis placed on Italian location entities in tweets for PUC1, English ones for PUC2 and Finnish ones for PUC3, in order to ensure that the effort done is directly reflected to the users' particular needs.

3.9   Linking Open EO Data (KR09)

3.9.1   Purpose of Scientific Activity

The module is developed as a process that may be integrated in processing workflows to transform the metadata of Earth Observation products (such as Sentinel scenes) from their native format to Linked Data (High level Call Requirement II). The process behaviour is controlled using sets of mapping rules. In the context of EOPEN, mapping rules will be developed allowing to convert XML-encoded Sentinel metadata as well as the JSON records obtained from the Umbrella Sentinel Hub (KR10). Outputs are RDF graphs intended to be stored in the Knowledge Base (graph database, via KR08). The target Linked Data structure is compliant with the GC Discussion Paper OGC 16-074 *EO Metadata Discovery using Linked Data*, published in March 2016.

3.9.2   Description of technological framework

The module makes use of the FOSS tool GeoTriples to perform the conversions. The module is stateless (no need for a database or session management, for example) as it receives the data to convert through input parameters and provides back the generated RDF graphs through output parameters. The actual storage of the generated graphs in the KB must be performed downstream in the workflow.
When deployed, the process exposes an OGC WPS 1.0.0 compatible interface and allows users to issue *execute* requests to it. When integrated in a workflow, all the interactions with the WPS interface are managed by the platform. The Use Case developer only needs to graphically integrate the process in workflows to use it.

During the reporting period, a process for storing Linked Data into a graph database has been implemented. This may be added in workflows and connected right after a conversion process for persisting the generated Linked Data and make it searchable by other applications.

3.9.3 **Functionalities and performance**

The module receives as inputs the EO product metadata (by value or by reference) and a reference to the mapping rules to apply. It returns back the generated RDF graph (by value) or the reference to a file that contains that graph.

At the time of writing, the evaluation of the Linking Open EO Data module has not been carried out yet. In order to evaluate it in EOPEN, specific needs (in terms of data types, link types and search queries) must be identified in the PUCs and implemented as workflows in the platform. Afterwards, the following key performance indicators will be calculated:

Table 15: Performance indicators and scientific metrics for Linking Open EO data

Indicator/Scientific metric	Value	Value achieved	Relevance to User Requirements	Comments
Conversion speed	Number		US_SK_03 US_SK_05 US_SK_06a	The number should be averaged.
Number of EO products (metadata) successfully converted	Percentage of test conversions			Test data to be prepared.
EO products metadata successfully stored in the KB	Percentage of test storage			Test storage scripts to be prepared.
EO products metadata searchable in the KB	Percentage of test searches			Test queries to be prepared.
Successful cross-concept queries	Percentage of test queries			Test queries to be prepared. Example of cross-concepts query involve EO products and tweets.

3.9.4 **How does the technology enable the PUC implementation**

Encoding the EO products metadata using the same specifications as the tweets and the events makes it possible to store, manage and retrieve linked entities in a seamless manner, using either stSPARQL or GeoSPARQL (provided by KR08). For example, it becomes possible to search in a single query for all the messages tweeted and all the satellite imagery acquired in the area and the time frame of a given event.

It also becomes possible to use an RDF browser to navigate in the tweets, events, EO products, and any other concepts that are encoded using the same classes.

4 TECHNOLOGICAL ACTIVITIES

Key Result preliminary evaluation and establishment of KPIs[2] to monitor

4.1 EO data acquisition (KR10)

4.1.1 Purpose of the technological activity

The developed Umbrella Sentinel Hub, under the EO data acquisition Key Result, provides a transparent and seamless access to Sentinel data to the user, enabling the Earth Observation (EO) based implementations of the PUCs. This EOPEN platform service (High level Call Requirement I)offers a single point of access to multiple Sentinel hubs, allowing access to all Sentinel missions, with no geographic restrictions, high performance, and minimal delays. In the current version, five hubs have been connected to the Umbrella hub, while a dedicated API accepts requests for metadata harvested from these sources. Future versions will include a fine tuning of the Umbrella application aiming to optimization and also the implementation of a GUI on EOPEN User Portal.

4.1.2 Description of technological framework

The developed application connects to the APIs of multiple Sentinel hubs, searches and stores new metadata and chooses the most appropriate source from which a requested product will be downloaded from (High level Call Requirement II). The Umbrella application PostgreSQL database frequently collects metadata of newly ingested Sentinel products, keeps up to date by deleting no longer available products and scores the different connected hubs according to their download speed.

4.1.3 Functionalities and performance

The API application creates a REST API via Django views module and allows users to make GET requests to it. In this request, the users are able to define their parameters (mission, product, model, date etc.) based on their needs. The Umbrella hub gives back a result set that contains metadata from the most efficient source to download from. Furthermore, an evaluation methodology has been established and introduced in the following table (Self-assessment):

Table 16: Performance indicators and scientific metrics for EO data acquisition

Indicator/Scientific metric	Value	Value achieved	Relevance to User Requirements	Comments
Number of connected hubs	Number	5	US_V_01c US_V_05a3 US_SK_03	2 more hubs will be added. Connecting with ONDA DIAS will be explored – if applicable

---

[2] The KPI tables in Section 3 include the field 'relevance to user requirements'. The user requirement IDs included can be found in the ANNEX of D2.2., which is a living a document and will be soon resubmitted.

Mean download speed compared to Open Access Hub	Percentage of increase	$\dfrac{8.8(MB/sec)}{2.92(MB/sec)}$	US_SK_06c US_SK_07a US_SK_12a US_F_01a2 US_F_03 US_F_04a US_F_04a3	Ensured timeliness for time sensitive applications
Mean latency compared to Open Access Hub	Percentage of decrease			Ensured timeliness for time sensitive applications
Access to all Sentinel missions	Boolean	Yes		Access to S-1, S-2, S-3 and S-5p allowing the generation of diverse datasets from soil moisture to bathymetry
Search data for a user defined area of interest	Boolean	Yes		
Search data on a specific time range	Boolean	Yes		
Number of broken links	Number	17 broken links in one week from all connected hubs		

4.1.4 How does the technology enable the PUC implementation

This Umbrella Sentinel API functions as the basic building block in all EO based PUC workflows, enabling EO data acquisition in a comprehensive manner. On top of this API, customized PUC-specific processors are built. PUCs used this API successfully (Cross-partner assessment): For the case of PUC2 where time-series of Sentinel imagery is required to capture the phenological stages of the crops, the time-series processor has been implemented by the PUC developers. This is another general process utilizing the Umbrella application to search and download Sentinel-1 and Sentinel-2 time-series, which can be parameterized by the user (start month, end month, max number of images per month etc). PUC1 uses satellite data for mapping flood extent as part of the change detection process. During the same process PUC3 identifies also via satellite images whether a road is passable or not due to flood or snow. Both modules use successfully Umbrella API for searching Sentinel-1 and Sentinel-2 products of the previous day for the desired Areas of Interest (AoI) via processes which are executed on a daily basis

4.2 Social media crawlers (KR12)

4.2.1 Purpose of the technological activity

The Social Media Crawling module is responsible to gather information from Twitter that serves as a non-EO data source for EOPEN. Based on suggestions by end users, proper search filters are used to crawl valuable social media posts. In the current version, after being collected, we detect whether tweets are fake, we try to detect their location if mentioned, and we analyse their images to extract visual concepts and check for nudity. In the final version, we will use textual and visual information of tweets to automatically estimate whether they are relevant or not to the examined use cases. At this point the

crawling procedure and the complementary components only run on CERTH's side, but all the information is stored in a database that the platform has access to. An introduction to this module has been included in D3.1, but future deliverable D3.3 will be dedicated to it.

4.2.2 Description of technological framework

In order to gain access to the global stream of Twitter data, we use Twitter API, a free option to stream real-time tweets. The retrieval options can be keywords that must be included in the posts, open user accounts and locations in form of bounding boxes. Using these search criteria, the API constantly retrieves new tweets in a JSON format and stores them in a MongoDB database. Any information derived from the analysis of the crawled tweets (e.g., extracted locations and concepts, relevancy estimation, nudity detection, etc.) is appended to the JSON before storing.

4.2.3 Functionalities and performance

The evaluation of this module will focus on the visual and textual classification techniques regarding the relevance of collected tweets. The developed models will be language- and use case-specific, so they will be evaluated separately. The metrics to be calculated in the future can be seen in the following table. The automatic estimation of relevance will focus on PUC1 and PUC3, in order to filter out tweets that do not contain information about critical events of extreme weather. The implementation of such models strongly requires training datasets, i.e. ground truth that can be obtained by human annotation. This task has already been completed for both PUCs, so the next step is to train the models and evaluate them.

Table 17: Performance indicators and scientific metrics for social media crawlers

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Precision	%	US_V_02a1	These KPIs will be examined in the future.
Recall	%	US_V_02a2	
F-score	%	US_F_01b	
Accuracy	%		
Response time	millisecond (ms)		

4.2.4 How does the technology enable the PUC implementation

As mentioned above, the social media crawling procedure uses predefined search criteria, which are tailored to the needs of each PUC. For example, for PUC 1 we search for Italian keywords regarding flooding events inside the bounding box of north-east Italy. For PUC 2 we follow accounts relevant to food security and search for Korean words that refer to

agriculture. Finally, for PUC 3 we search for Finnish terms about snow coverage. As a result, every collection of tweets is directly connected to a use case. However, the automatic estimation of relevant tweets will be limited to PUCs 1 and 3, for which we have gathered the human annotation that is needed (Cross-partner assessment).

4.3 Meteo data acquisition (KR13)

4.3.1 Purpose of the technological activity

The developed Weather Data Wrapper Module is a collection of functionalities designed for extracting, transforming, and storing meteorological and climatological data in the EOPEN platform.

4.3.2 Description of technological framework

The module provides a collection of functions and methods, written in Python. EOPEN developers can use the module to automate data ingestion to the EOPEN platform (High level Call Requirement I and II). The data is stored into a common database, from which the user can obtain the requested data via suitable database queries. In addition to PUCs, the module can also serve other EOPEN modules, such as Event Detection module.

4.3.3 Functionalities and performance

The module connects to multiple open APIs which provide meteorological and climatological data, extracts the data from the services, and transforms the data into JSON data format. During the project, the development is focused on the needs of the use cases, however, the module is designed to be reusable for the future needs of EOPEN users. Self-assessment is presented in the following table:

Table 18: Performance indicators and scientific metrics for meteo data acquisition

Indicator / Scientific metric	Value	Value achieved	Relevance to User Requirements	Comments
Number of connected services providing meteo data	Number	4	US_V_08 US_SK_02 US_SK_03	The value achieved covers all use case data requirements.
Module is extendable to include additional APIs or services	Boolean	Yes	US_SK_05 US_F_01 US_F_07 US_F_08 US_F_09 US_F_10 US_F_11	Extendability is not directly related to the needs of the use cases. However, the module must allow for inclusion of new services as required by future EOPEN users.

4.3.4 How does the technology enable the PUC implementation

The module connects to open APIs, which provide the meteorological and climatological data as requested by the Use Cases. In order to limit the requests made to the services, this module is designed to be a lower-level functionality. Namely, the data ingestion and

transformation are developed as separate processes from the Use Cases; the Use Case developers should query the EOPEN database in their processors. For the Use Cases, the module will be developed to connect to FMI Open Data service (PUC1, PUC3), Korea Meteorological Administration Open API (PUC2), Copernicus services (PUC1, PUC3), and NASA Langley Research Center (PUC2).

4.4   Standardized Interfaces (KR14)

4.4.1   Purpose of the technological activity

The use of standard interfaces, including formats and protocols, is the key to bring interoperability among independently designed and developed systems and environments (High level Call requirement III). To be fully interoperable, a system must adhere to standard interfaces on all its ends, including back-ends interfaces used to communicate with remote services and data providers, and its front-ends allowing client users and applications to interact with it in a seamless manner.
Standardized interfaces thus allow systems to not operate in an isolated manner, to benefit from data and services offered by other compatible systems, and on its turn to offer its data and services to client applications.

4.4.2   Description of technological framework

The EOPEN platform is an environment that allows integrating, configuring and executing any kind of services (High level Call requirement I). An example of custom service is the harvesting and processing (geo-location, concept extraction, community detection) of tweets. Other services automatically harvest meteorological products encoded in netCDF, convert them to JSON, and store the results in MongoDB. The platform also integrates processes that allow publishing products (such as GeoTIFF files) as GIS layers through GeoServer. The adoption of standards must thus not only be considered as the platform level but also at the processes and applications level.
From the user point of view, it is important to be able to discover and access the available data, as well as discover and use the available services in a standard manner.

4.4.3   Functionalities and performance

Table 19: Performance indicators and scientific metrics for Standardized Interfaces

Indicator/Scientific metric	Value	Value achieved[3]	Relevance to User Requirements	Comments
Ability for a user to authenticate using a standard protocol	Boolean	Dev. Platform: True	US_SK_14	Relevant protocols: OAuth2, OpenID Connect, CAS

---

[3] Most indicators are "Mostly True" or "Mostly False" at this stage. Indicating "True" here does not mean we do not need to still improve the solution and indicating "False" does not mean it will not be developed later. It merely records the current status.

		Dashboard: False		
Ability to discover the available collections and products through a standard interface	Boolean	False	US_F_05a	Relevant interfaces: OGC CSW, OpenSearch, OpenAPI, SPARQL
Ability to access the product data files through standard interfaces	Boolean	True	US_F_05a	Relevant interfaces: http, ftp
Ability to access GIS layers through standard interfaces	Boolean	True	US_SK_06b	Relevant interfaces: OGC WMS/WMTS, WMS-T, WFS, WCS
Ability to remotely execute an application through a standard interface	Boolean	False	US_F_05b	Relevant interface: OGC WPS

4.4.4 How does the technology enable the PUC implementation

PUC stakeholders already use tools to perform their tasks. EOPEN brings enhancements in providing new products and services. The use of standard formats and protocols allows PUC users to interconnect the tools they are already using with services run in the EOPEN platform and they will be able to fetch the data available in EOPEN to use them in their tools.

Concrete examples include the possibility to visualise GIS layers (OGC WMS, WFS) served by EOPEN in any standard-compliant GIS client such as QGIS and ArcGIS. Opening water body masks encoded as raster data (GeoTIFF) or vector data (Shapefile, GeoJSON) is also possible in user-chosen software. Moreover, these standards allow visualising and fusing data served by EOPEN with user owned data and any other data served by third-parties.

4.5 High Performance Computing (KR15)

4.5.1 Purpose of the technological activity

The goal of this module is to integrate HPC and HPDA infrastructures into the EOPEN platform for deployment of embarrassingly-parallel tasks and data analytics applications. Specifically, Cloudify orchestrator will be implemented to enable access to both High Performance Computing (HPC) and High Performance Data Analytics (HPDA) infrastructures in a seamless and user-friendly manner.

4.5.2 Description of technological framework

To enable the full utilization of the computing resource of HPC and HPDA through Cloudify orchestrator, further Cloudify plugin needs to be developed. The Cloudify plugin is composed of two parts: Cloudify HPC plugin and Cloudify HPDA plugin.

Cloudify HPC plugin, which is designed to organize the interaction between portal and HPC systems, consists of three main modules: workflows, tasks, cli_client and workload_manager. It was already developed in the former GoeGSS project and has been implemented in EOPEN. A sample BASEMENT simulation for hydro- and morphodynamic modelling has successfully run on the HPC at HLRS. The Cloudify HPDA plugin has been implemented to support the integration of HPDA and EOPEN platforms. The submodules of workload_manager and cli_client in HPDA plugin are significantly different from the ones in HPC plugin, which enable the connection to and utilization of Apache Spark.

To effectively process and analyze Big Data on the HPDA system, a stack of software frameworks is established. Apache Hadoop is enabled for processing Big Data at large-scale. Apache Spark, which allows in-memory processing and a large set of libraries for Big Data processing, including common clustering and classification tasks is also enabled. Cray Graph Engine is supported to provide the ability to perform graphic analysis. Besides, Tensorflow is supported to enable deep learning algorithms performed on the image analysis. Further information is delivered in D6.1.

4.5.3 Functionalities and performance

The performance will be measured via a monitoring framework such as ATOM, which is going to be deployed, and that performance measurements will be provided in deliverable D6.3.

Table 20: Performance indicators and scientific metrics for concept extraction in social media

Indicator/Scientific metric	Value	Value Achieved	Relevance to User Requirements	Comments
Access to HPC	Boolean	True	US_V_01 US_V_02 US_V_08d US_F_08c1 US_F_10a1 US_F_10b1 US_F_11a1	Access to HPC enabled via ssh and Cloudify
Access to HPDA	Boolean	True		Access via ssh is enabled and the development and implementation of Cloudify HPDA plugin is done
At least three application kernels parallelized	Number	5		A distributed random forest and a distributed k-means has been implemented by NOA on HPDA. Besides, a distributed text analysis (LDA), clustering of twitter images and a deep neural network to segment the EO image are also implemented on HPDA.
Improved application performance by at least 15%	Percentage	N/A		Prerequisite for benchmarking is the deployment of ATOM which is to be done.

At least 5 workflows ported to HPC/HPDA	Number	5		A clustering workflow to create training data
				A Random Forest workflow for rice classification using the training data provided by the first workflow
				A Twitter Image clustering workflow
				A Twitter topic extraction workflow
				An EO image segmentation workflow
Workload manager as a component	Boolean	True		Employed Cloudify as orchestrator for HPC/HPDA;
				Integrated within the EOPEN platform
Adopted state of the art data analytics frameworks	Boolean	True		Apache Spark, Cray Graph Engine and TensorFlow, etc. are available for processing.

How do we measure

For monitoring HPC jobs at HLRS, CrayPAT as a preinstalled tool is adopted in that previously monitoring on "Hazel Hen" as described in D7.1. However, from the beginning of 2020, HLRS will have a new HPC infrastructure 'Hawk', HPE Apollo. *Extra-P* and CUBE are installed on 'Hawk as an automatic performance-modeling tool and a performance report tool respectively.' These tools can collect different performance information which can be used for the analysis includes: call graph profile, communication statistics, time-line view for Communication and IO, activity view, pair-wise communication statistics and text reports. Based on the analysis, the user can easily monitor the load imbalance, excessive communication, network contention, excessive serialization and I/O Problems. Except from the HPC monitoring stated above, ATOM will be implemented next to monitoring the HPDA framework at HLRS. The objective of the ATOM Monitoring Framework is to enable application optimization based on the understanding of both software non-functional properties and hardware quality attributes with regards to performance and energy consumption.

The monitoring workflow is divided into two parts as shown in the next Figure. The devices keep being monitored by the **MF-Client** (Infrastructure-level), and the instrumented applications keep monitored using the **MF Library** (Application-level). The metrics collected are sent to the **MF-Server**, which tasks are to store the metrics and to provide an interface for query and analyse them.
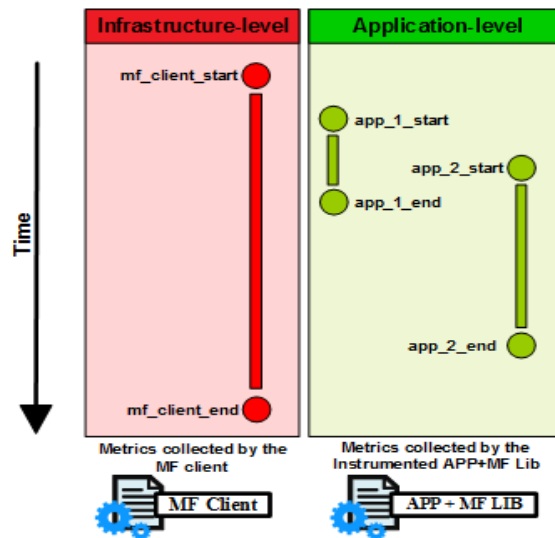
Figure 4: Infrastructure-level and Application-level monitoring.

The metrics that can be acquired cover a wide range of functions that target different aspects of the hardware, such as memory, IO, processor, GPU, and network utilization as well as energy consumption. Next figure shows also a classification of metrics:
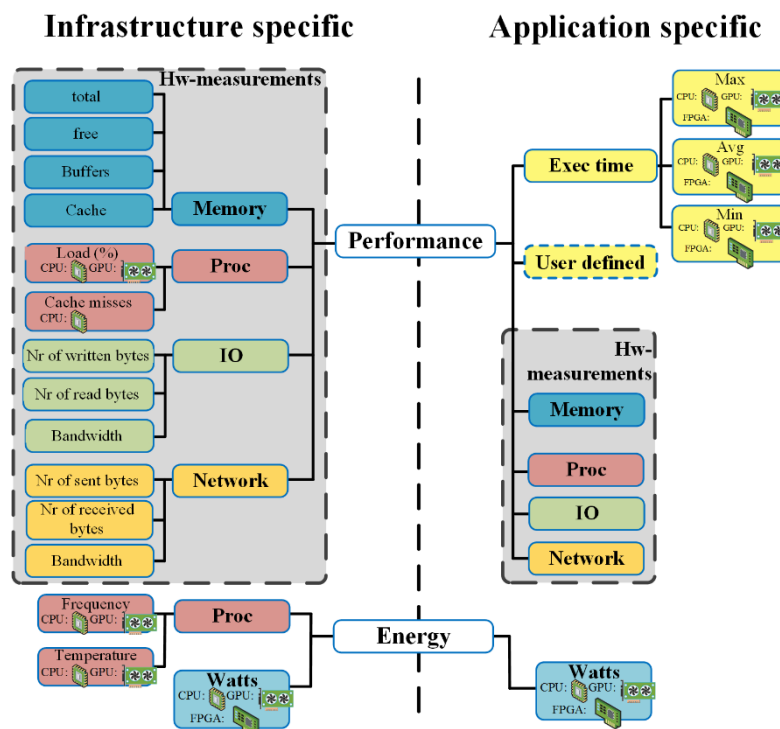
Figure 5: PHANTOM Monitoring Client metrics' taxonomy

For more detailed information, please refer to: **https://github.com/PHANTOM-Platform**.

4.5.4 How does the technology enable the PUC implementation

For PUC1, it is required to provide faster and more effective emergency responses to extreme weather by increasing the speed of risk analysis and by improving in-field coordination of emergency response. This requirement can only be satisfied by exploiting

large-scale distributed computing power provided by HPC and HPDA. EOPEN builds on state-of-the-art methods in HPC and HPDA, and as the basic building block in all EO based PUC workflows, the extreme performance of the computing systems will be exploited by the EOPEN analytics platform. The High Performance Computing and High Performance Data Analytics (HPDA) can provide two different strands for PUC: batch job processing for static workloads as well as stream processing for live data analytics. One great example is PUC2 that deals with a long time-series of Sentinel imagery that extends over multiple years. Such applications, at the national scale, constitute big data problems and can be solely tackled by high performance computing systems - see more information in the PUC2 section.

4.6 Fast and Secure Data Exchange (KR16)

4.6.1 Purpose of the module

The purpose of this module is to enable a fast and secure data exchange so that the full bandwidth of the underneath high-speed data connection could be utilized. Moreover, a secure transfer and storage will be ensured through the maintenance of standards in encryption. Besides, data locality is of utmost importance in order to achieve maximum performance, thus input data are required to be transferred in a secure and fast manner to these systems before processing can start.

4.6.2 Description of technological framework

A secure and fast data transfer from and to HLRS infrastructure is realized through SCP. As the most commonly used manner for data transfer, SCP uses Secure Shell (SSH) for data transfer and uses the same mechanisms for authentication, thereby ensuring the authenticity and confidentiality of the data in transit.

Besides, as the Sentinel images for the PUC datasets are refreshed at quite low frequency, therefore as an addition to transferring data through scp, the Umbrella API which was implemented by NOA will provide the metadata to HLRS and then the products will be downloaded with a simple script directly at HLRS. Thus the workload for transferring data from NOA to HLRS is eliminated.

4.6.3 Functionalities and performance

Table 21: Performance indicators and scientific metrics for fast and secure data exchange

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Service Coverage	%	US_V_02a US_V_02b	Number of sites which are offering data services (goal is to have a 100% coverage)

Service Availability	%	US_V_02c	Uptime is taken into consideration for monitoring the service availability. If the service is required to be available all the time, then the goal is to have 100%.
		US_V_02d1	
		US_V_05a	
		US_V_05b	
		US_V_09a	
		US_V_11	
		US_SK_07	

4.6.4 How does the technology enable the PUC implementation

Data exchange among sites is a fundamental requirement for all PUCs. SCP is commonly used as a data transfer tool that enables secure and fast file transfer between HLRS and the users of its infrastructure. Also, the download script which can fetch the metadata and download links provided by the Umbrella API will be implemented at HLRS locally. Therefore, the workload of data transfer would be eliminated, and the efficiency of processing is improved.

4.7 The EOPEN Platform (KR17)

4.7.1 Purpose of the technological activity

The EOPEN Platform provides core components that allow orchestrating the execution of processes organised in workflows. The platform includes built-in processes that allow performing generic tasks such as sending email notifications and registering a product in a catalogue. Common libraries and toolboxes are also available in the core platform that allow developing custom algorithms and import them as processes.

The EOPEN Platform is accompanied by a generic Web-based user interface allowing to interact with the platform, request for workflow executions, and for discovering and visualising the available data.

The EOPEN Platform is meant to be used by developers for implementing new modules and applications at setup-time and used as a back-end for collecting, processing and serving data to end-users during normal operations.

4.7.2 Description of technological framework

The EOPEN Platform is based on the ASB Framework initially developed for ESA. In particular, the ASB core components included in the EOPEN Platform are used by the module and Use Case developers to import their algorithms (as processes and services), create, test and execute processing workflows, and publish these as applications (processors). In the course of the EOPEN project, a number of modules (some are implementing the KRs) must be integrated allowing the preparation of the Pilot Use Cases.

4.7.3 **Functionalities and performance**

The core functionalities of the EOPEN Platform reside in the ability to import algorithms, configure workflows and execute processors. These features are available through a Web interface. The performance may be related to the amount of modules integrated and available for implementing each of the Use Cases.

Table 22: Performance indicators and scientific metrics for the EOPEN platform

Indicator/Scientific metric	Value	Value achieved	Relevance to User Requirements	Comments
Amount of integrated services	Number	9*		
Amount of operational processes	Number	12*		
Amount of operational processors	Number	>5		
Amount of scheduled processors	Number	6		
Amount of available database engines	Number	3	US_F_01	PostgreSQL (RDBMS), MongoDB (Document), GraphDB (Triplestore)
Amount of product types managed in the platform	Number	>8	US_F_01	Tweets, Sentinel-1/2/3 images, water bodies, SMOS Freeze-Thaw, GlobSnow SWE L3A, rice paddy fields, ...
Amount of processes executable in the HPDA	Number	0	US_F_05b	Proofs-of-concept have been implemented but no process currently integrated in EOPEN workflows. The PUC2 rice classification pipeline is in progress to become executable through the platform.
Amount of DIAS platforms available for deploying EOPEN processes	Number	up to 4 out of 5	US_F_01d	Support for all but one (WEkEO) has been demonstrated. Due to the limited IT resources we have in the DIAS environments (mostly trial nodes), we cannot keep all of them alive all along the EOPEN project.

Amount of product types accessible in the Dashboard	Number	>5	US_SK_14	Tweets, water bodies (from S1 and S2), rice paddy fields (from RNN), rice paddy fields time series.
Amount of product types accessible through APIs	Number		US_SK_06c	This includes custom and standard APIs, including database API, OGC Web Services, etc.

* As documented in D6.4 EOPEN 2nd Prototype Developer Portal.

4.7.4 **How does the technology enable the PUC implementation**

The EOPEN Platform, with its core components and its integrated modules allows PUC-specific processing chains to be implemented and executed (immediately or differed) and the collected and generated data to be displayed to the end-users. In this context, the usefulness of the Platform not only resides in the built-in features but also on the various modules that will be integrated (some being directly related to the KRs).

5   HACKATHON AS MEANS OF EVALUATION

This section reports about the preparation and attendance of the H2020 Earth Observation Big Data Hackathon event, held in Frascati on 7-8 November 2019, as well as its follow up. The event has been organised by five EU-funded Horizon 2020 projects that are developing cutting-edge software to process large amounts of Earth Observation (EO) data; it was open and free, aimed to give the opportunity to learn about and gain hands-on experience with various innovative EO solutions being explored by the following five projects: BETTER Data Pipelines, CANDELA platform, EOPEN platform, OpenEO API, and PerceptiveSentinel's EO-LEARN Python library. Additionally, the hackathon served as an evaluation activity for the EOPEN platform, showcasing the interoperability of the platform as a whole and the reusability of its constituent components (3rd party users evaluation).

5.1   Success stories

The EOPEN session of the hackathon should be considered overall very fruitful, since there were many cases of successfully completed exercises, involving participants from various projects.

In detail, two partners of the CANDELA project managed to execute through the EOPEN platform a water body detection algorithm using a client provided by the BETTER project for getting the satellite data (first exercise). The generated water mask was previewed (as a PNG file) in the EOPEN interface.
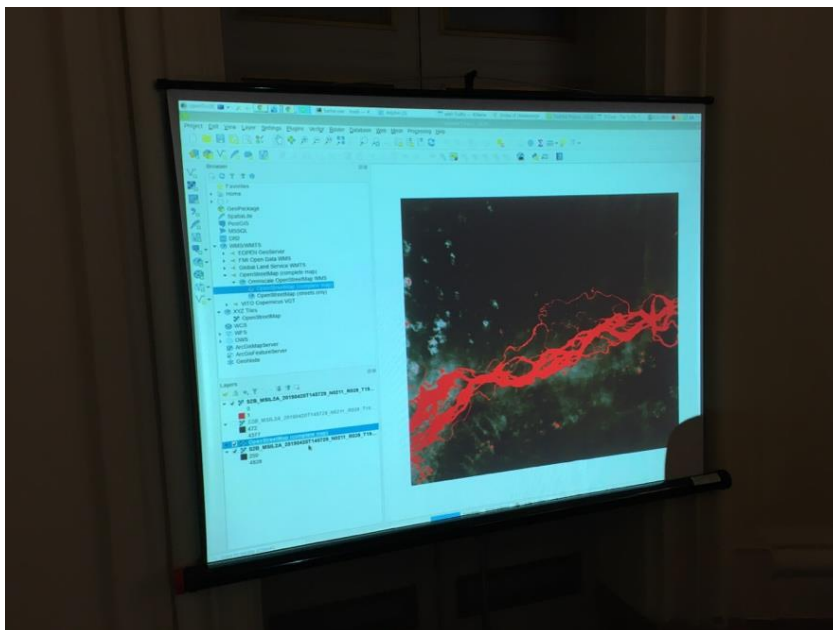
A person from the openEO project succeeded the second exercise, by logging in to the EOPEN platform and generating a water mask with the use of an openEO client for products and EOPEN extensions for water body detection. Again, the mask was previewed in the Web browser.
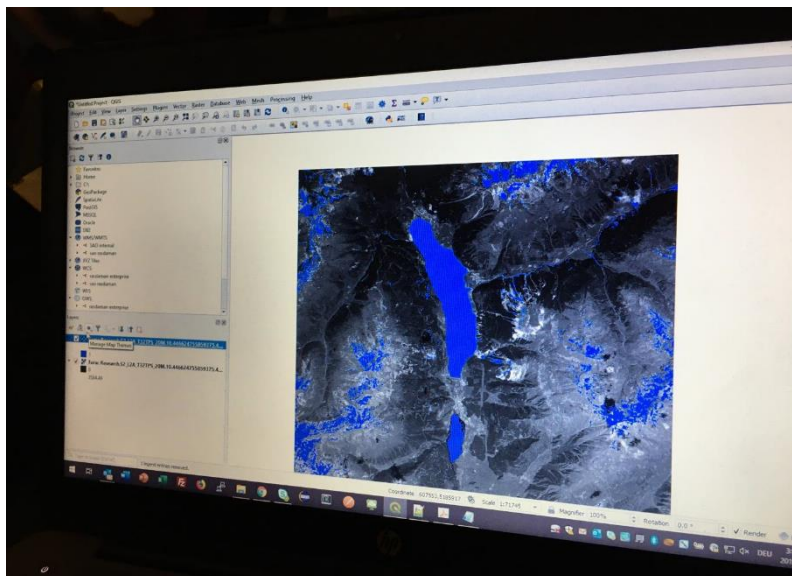
Regarding the third exercise, people from BETTER and CANDELA projects achieved to download tweets through the EOPEN service and expressed their interest to further utilise them in the future. Additionally, some technologies relevant to non-EO data, such as topic and community detection, were demonstrated by EOPEN. During the fourth exercise there was a presentation of executing an algorithm on the ONDA-DIAS platform in order to generate water masks. Due to time limitations, there was no hands-on activity by the participants in this exercise.

Finally, the fifth exercise has been executed in parallel to the others. It is not always straightforward to adapt an algorithm in a few hours. This challenge has been taken by a member of the openEO team to enhance the openEO client process (from exercise 2) to make it more robust and feature-full.

The result has been successfully executed in the EOPEN Platform for searching for and downloading Sentinel-2 products from the EURAC openEO provider. The resulting algorithm is now available as an example script in the official openEO Python Client repository on GitHub[4]. It should also be noted here that just the participation of EOPEN in the hackathon event attracted the attention of external parties, raising unexpectedly the project's publicity.

5.2 Questionnaire and feedback analysis

After the activity, a questionnaire has been distributed electronically. The collected feedback is provided and analysed below. The questionnaire included five questions which have been answered (either partially or completely) by 9 people. The questions and answers are collected hereafter:

---

[4] https://github.com/Open-EO/openeo-api

1. **How easy was it to import a python script to the EOPEN developer platform?**
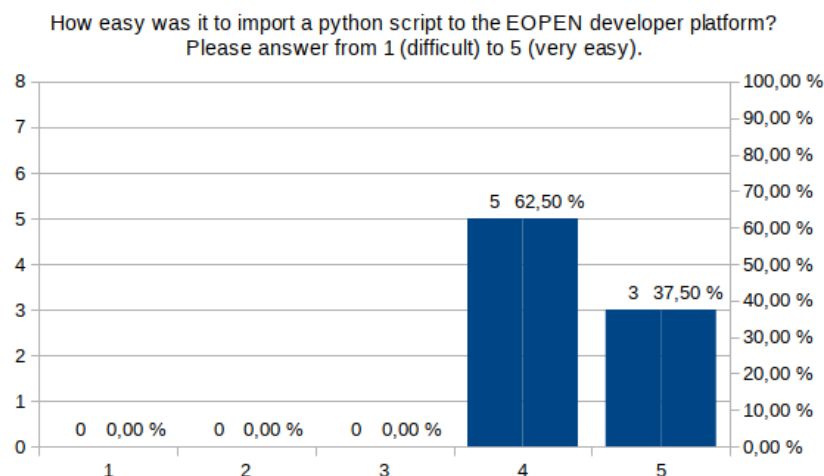   Please answer from 1 (difficult) to 5 (very easy).

Figure 6: Answers regarding easiness of importing Python script to the EOPEN developer platform

This result is very encouraging as it means we are on the right track in what concerns the usability of the platform, at least for the import of new processes. Nevertheless, we must keep in mind that the exercise has been performed with the active support of experienced users and developers. The challenges reported hereafter give a more detailed indication on the aspects that must be improved to enhance the platform and make it more attractive.

2. **Which was the most challenging part in executing an algorithm in EOPEN?**
   Answers:
   a. Checking the intermediate results and fast debugging
   b. Wifi connection
   c. Bug fixing! It is not difficult, but time consuming due to the permanent recreation and redeployment of containers.
   d. The comprehension of the input (maybe provide not only an input title, but also a description) + need to change of page for every step (I take time to search how execute the algorithm + I forget to save changes)
   e. The most challenging part was having a successful script using openEO API

Two (a, c) out of the five answers concern the difficulty to figure out what is going wrong when a process execution fails and the fact that importing new process versions is time consuming.

Answer d. shows that the graphical user interface is not enough self-descriptive and intuitive. More information must be provided to the users to guide them in the various elements of the interface.

Answer b. reflects the need to have a stable Internet connection to be productive in using the EOPEN Platform.

Answer e. relates to the difficulty to write an openEO client script. As the EOPEN Platform executes the script as-is, this issue is not under the control of the platform. The platform must however make it easy to access the information about the failed executions.

3. **Are you interested in exploiting Twitter geo-localised data in your own platform?**
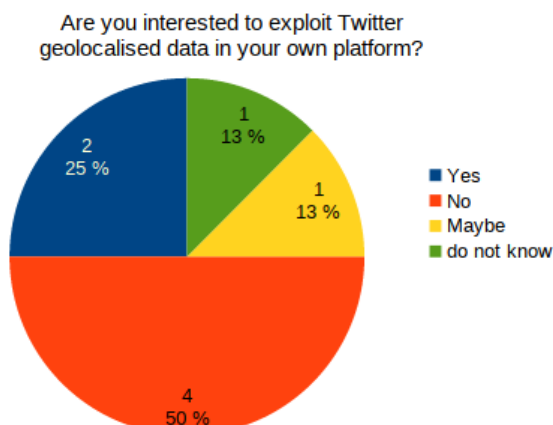

Figure 7: Answers regarding interest on exploitation of Twitter geo-localized data

Approximately one third of the responders are or could be interested in including Twitter posts in their platform, which is quite considerable, given the fact that the projects focus on Earth Observation data and tweets are a different type of data. Regarding the 13% that are undecided, maybe the usefulness of Twitter should have been better communicated to them during the respective exercise.

4. **Do you prefer JSON files or RDF data in getting EO and non-EO data?** The 8 responders to this question all prefer JSON. This is reasonable, since the usage of the JSON syntax is widespread and it offers compatibility between different systems/applications. In addition, based on the responses, it is estimated that not all projects involve semantic technologies in their framework.

5. **How likely is it for you to use the EOPEN platform to deploy and run a custom algorithm in the future?** Rate from 1 (unlikely) to 5 (very likely)
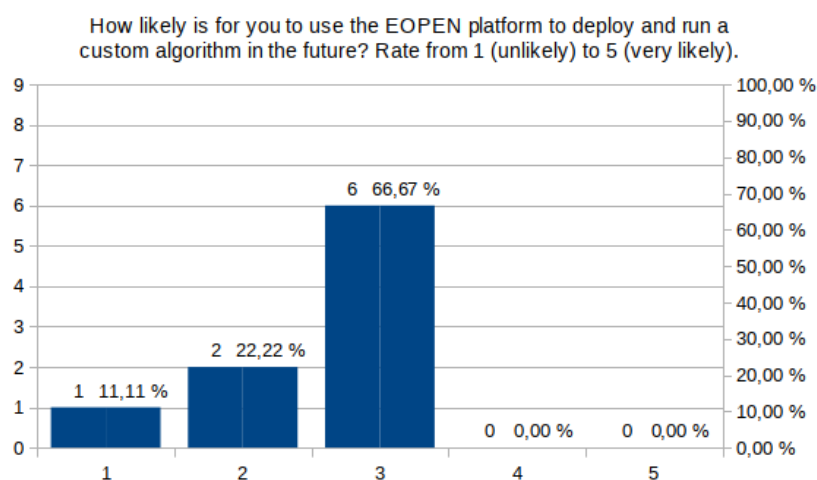

Figure 8: Answers regarding the possibility of deploying an algorithm to the EOPEN platform

This result shows that there is a non-neglectable interest for the features provided by the EOPEN Platform but substantial effort is still required to enhance its attractiveness as none of the responders has decided yet to use it to run a custom algorithm. In particular,

we have observed that the Platform is not yet intuitive and user-friendly enough to be used by first time users at least without being supported by experienced users.

As far as the EOPEN project is concerned, the dashboards which will be implemented for each Pilot Use Case are supposed to be user-friendly, whereas the Hackathon exercises were thought for developers - perhaps developers could benefit from looking at already implemented Use Cases for gaining a view about what can be achieved through the platform.

Nevertheless, the usability of the user interface will remain a major focus in the remaining of the project. Another one will be to make it easier to integrate third-party libraries and services such as the ones developed by the EO Big Data projects.

5.3 **Future collaborations**

After the completion of the hackathon, there has been a mutual interest between projects to initiate some collaborative actions. Now that each project has a clear understanding of the tools and capabilities offered by the others, chances that future collaborations will take place have become much greater.

In the exercises detailed above, the EOPEN Platform has been used to interact with openEO backend providers (namely VITO and EURAC) to search for and download Sentinel‑2 data. Another way to integrate EOPEN and openEO is to use EOPEN as a backend provider accessible through an openEO compliant interface. This would require developing a driver on top of the EOPEN Platform that implements the openEO API. The effort required is not neglectable but it is worth investigating this possibility in the future.

As far as it concerns the social media data, EOPEN and BETTER are already in discussion on how to combine work from the two projects. EOPEN provided 1,000 tweets in English about floods, which share at least one detected location and visual concepts extracted from their images. BETTER is planning to exploit these two attributes (location and concepts) for a meaningful analysis and visualisation of the results. Furthermore, there has been some initial interest from the CANDELA project.

The capabilities offered by the EOPEN Platform in terms of custom process integration and distributed workflow execution through a fully Web-based interface has also attracted the interest of a senior researcher at ASI (Agenzia Spaziale Italiana) who proposed to integrate a well-defined and validated workflow in the Platform.

6 PUC IMPLEMENTATION AND EVALUATION

6.1 Joint Decision & Information Governance Architecture Framework

6.1.1 Overview

The Joint Decision & Information Governance Architecture Framework (JDIG) is an architecture framework designed for EOPEN to support service improvement in the PUC's. Based upon TOGAF 9.2, it builds Current (COM) & Target (TOM) Operating Decision-making models, to identify how services are currently implemented in the PUC, and how they could be implemented using EOPEN; through a combination of People, Process, Information, Governance & Technology components.

6.1.2 Baseline Evaluation

As part of the JDIG each PUC underwent a baseline Holistic Capability Assessment to identify the maturity of its capabilities to deliver their desired services, as seen in Figure 9, Figure 10 and Figure 11 below. Rationale behind the evaluation and the process to define these are described in detail in *D2.3: Current Operating Decision-Making Models.* The colour coding is as follows:
- Green – Optimal Functionality
- Amber – Satisfactory but can be improved
- Red – Improvement needed
- Black – Not applicable/ Not enough information provided
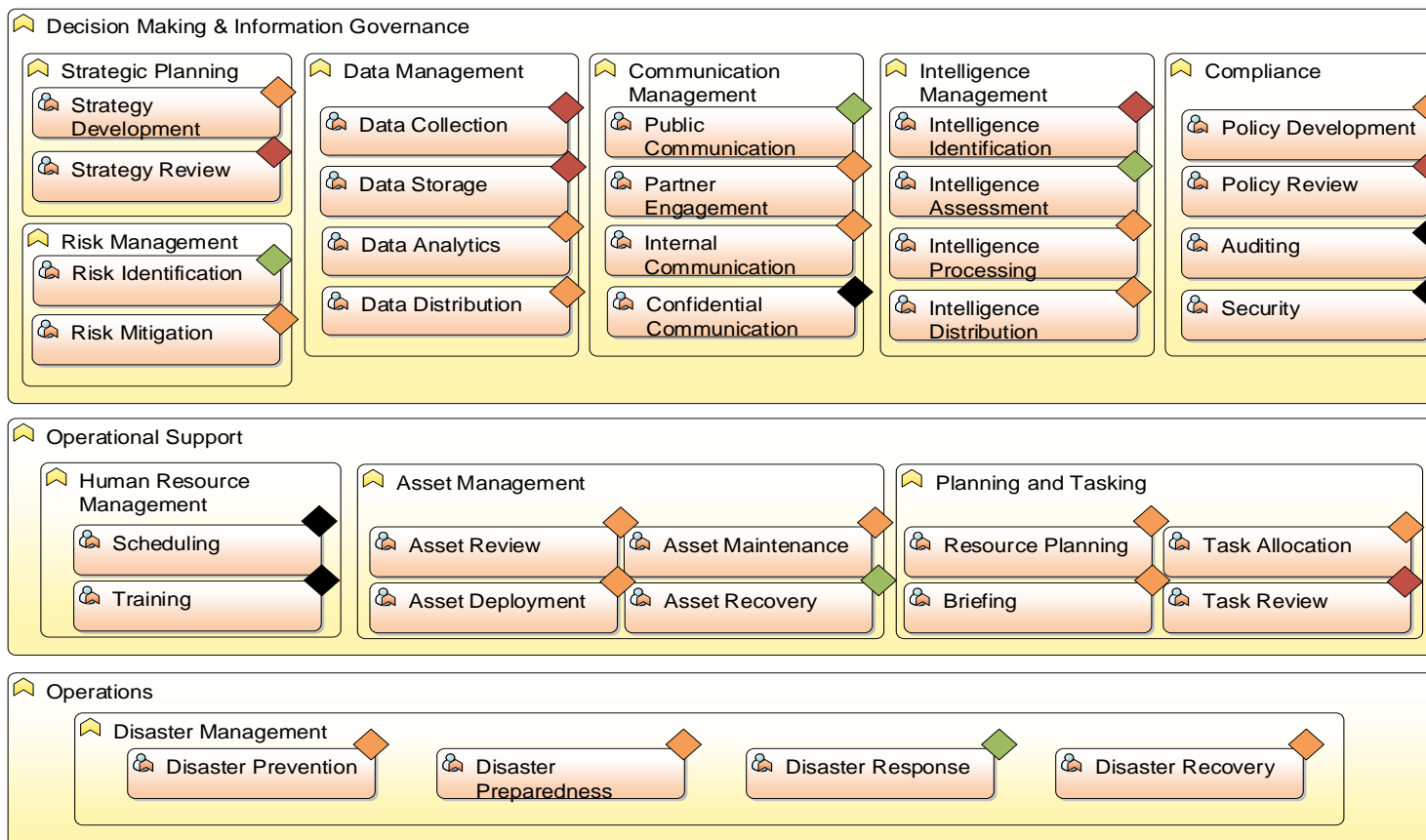
PUC1: COM Holistic Capability Assessment

Decision Making & Information Governance

Strategic Planning
- Strategy Development
- Strategy Review

Risk Management
- Risk Identification
- Risk Mitigation

Data Management
- Data Collection
- Data Storage
- Data Analytics
- Data Distribution

Communication Management
- Public Communication
- Partner Engagement
- Internal Communication
- Confidential Communication

Intelligence Management
- Intelligence Identification
- Intelligence Assessment
- Intelligence Processing
- Intelligence Distribution

Compliance
- Policy Development
- Policy Review
- Auditing
- Security

Operational Support

Human Resource Management
- Scheduling
- Training

Asset Management
- Asset Review
- Asset Deployment
- Asset Maintenance
- Asset Recovery

Planning and Tasking
- Resource Planning
- Briefing
- Task Allocation
- Task Review

Operations

Disaster Management
- Disaster Prevention
- Disaster Preparedness
- Disaster Response
- Disaster Recovery
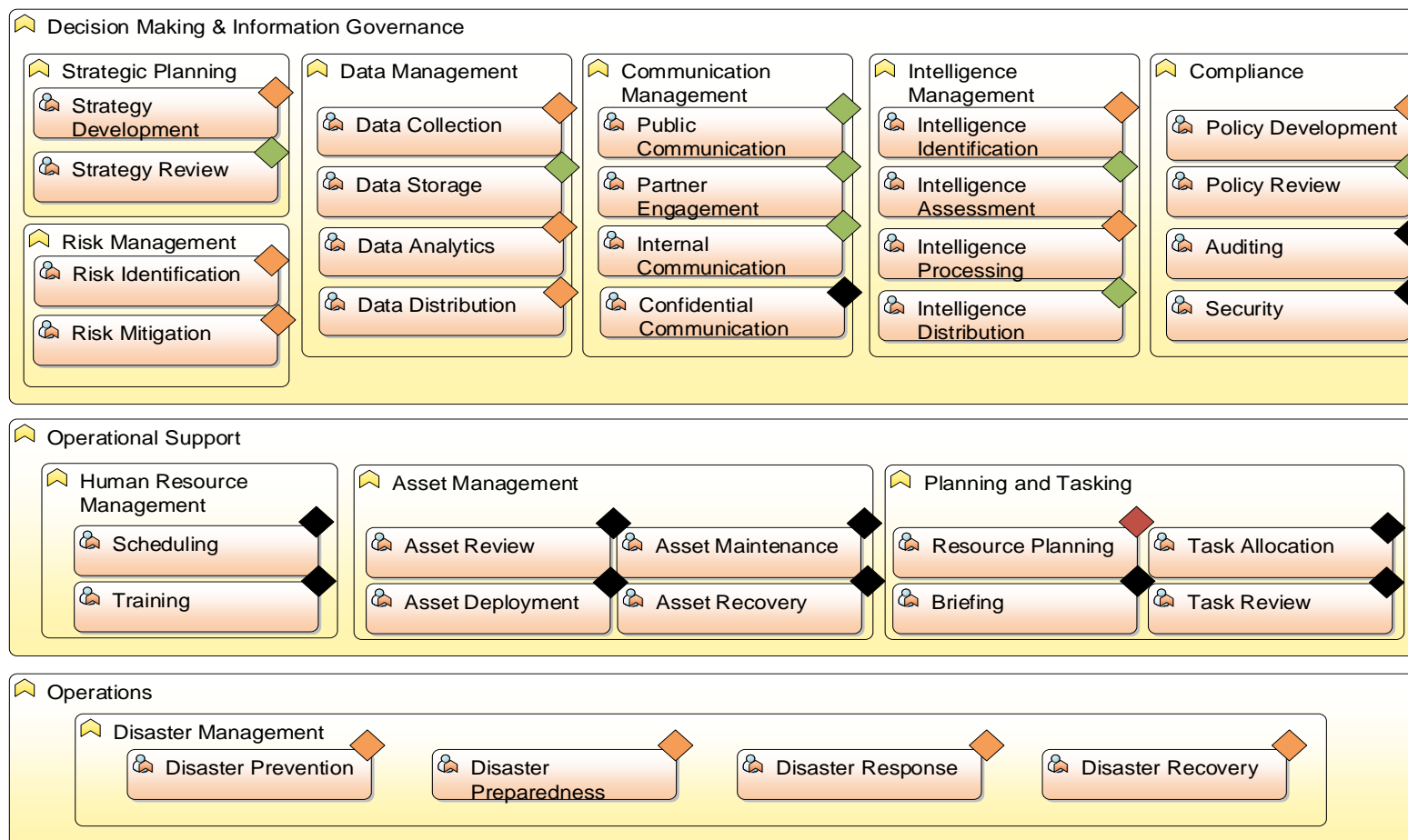
Figure 9: PUC1 - COM Holistic Capability Assessment

Figure 10: PUC2 - COM Holistic Capability Assessment

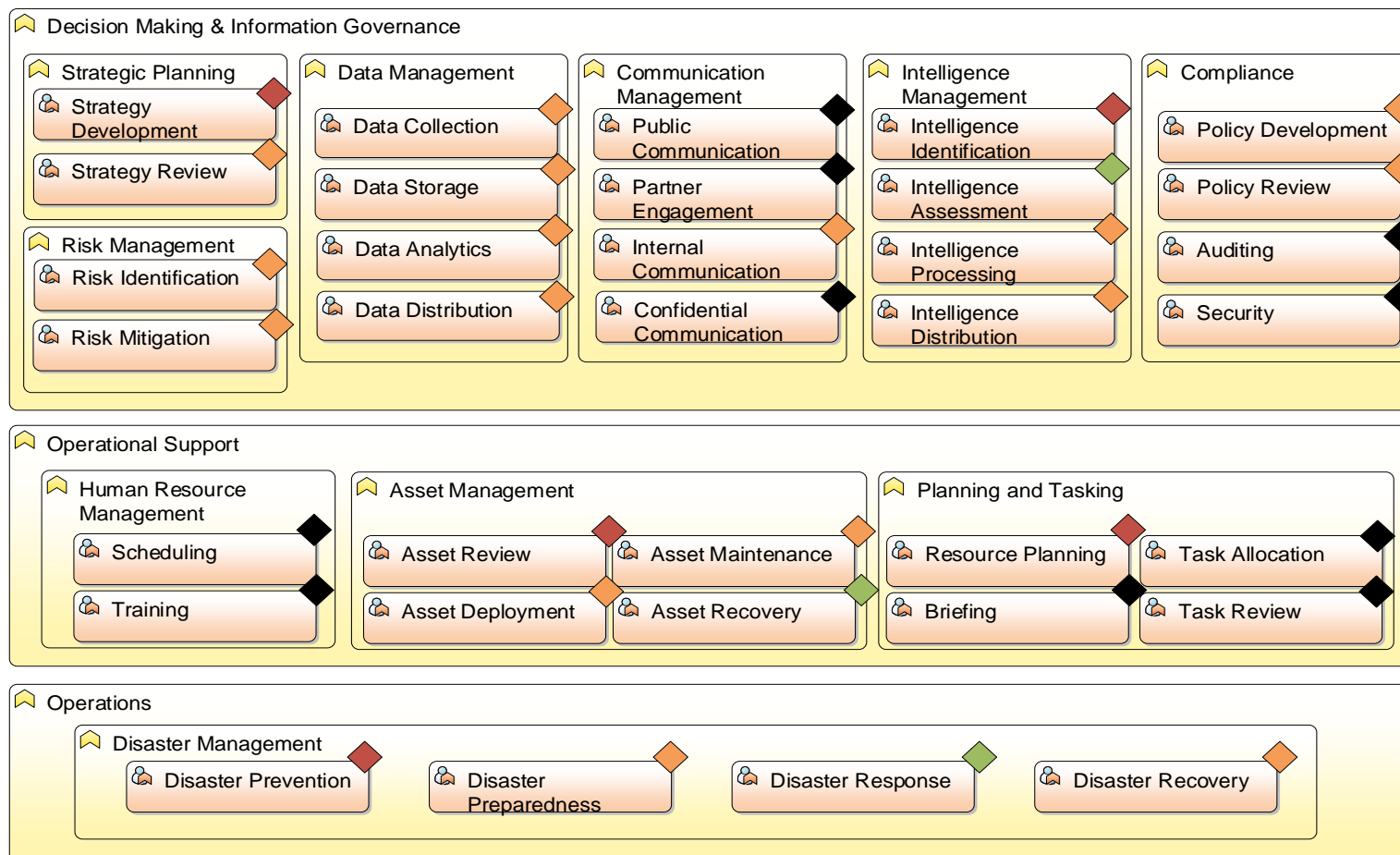## PUC3: COM Holistic Capability Assessment

Figure 11: PUC3 - COM Holistic Capability Assessment

6.1.3 Future Evaluation

As part of the JDIG TOM process, PUC's will undergo a final Holistic Capability Assessment, to verify improvements provided by the EOPEN platform have enhanced their capabilities. This capability assessment will be based upon the outputs of the individual PUC evaluations described in the following chapters.

The testing scenarios utilised by the PUC's to support their individual evaluation will be a realisation of the TOM processes created by the JDIG. These scenarios will provide end users with valuable insights into how the EOPEN platform will perform and function in a near to real-life situation, ensuring an accurate functional evaluation.

6.1.4 Evaluation Methods

For the JDIG capability assessment EOPEN will utilise a likert scale, against each of the capabilities listed below (5.1.5). The Likert scale will have axes of -2 to +2, with 0 being neutral, -2 being greatly reduced capability and +2 being greatly improved capability. This will be presented in the form of a questionnaire, to end-users of each of the PUCs as part of the pilot testing, in conjunction with any PUC specific questions. To be deemed successful EOPEN aims to achieve an **average score of +1** for all capabilities which were **red or amber** for the PUC baseline evaluation. For EOPEN to have achieved beyond expectation, the project aims to achieve an **average score of +1** for all capabilities with an existing **green** evaluation.

6.1.5 Desired Improvements

EOPEN aims to provide direct and indirect improvements to the below capabilities. Direct meaning EOPEN will provide this capability itself; Indirect meaning EOPEN does not provide this capability, but it is improved as a result of a direct improvement.

EOPEN aims to provide direct improvements to the following capabilities across all PUC's:

- Data Management (Data Collection, Data Storage, Data Analytics & Data Distribution)
- Risk Management (Risk Identification)
- Communication Management (Public Communication, Partner Engagement, Internal Communication)
- Intelligence Management (Intelligence Identification, Intelligence Assessment, Intelligence Processing, Intelligence Distribution)
- Planning & Tasking (Task Allocation, Task Review)

EOPEN aims to have indirect improvements to the following capabilities across all PUC's:

- Strategic Planning (Strategy Development, Strategy Review)
- Risk Management (Risk Mitigation)
- Compliance (Policy Development, Policy Review)
- Asset Management (Asset Review, Asset Maintenance, Asset Deployment, Asset Recovery)
- Planning & Tasking (Resource Planning, Briefing)

● Disaster Management (Disaster Prevention, Disaster Preparedness, Disaster Response, Disaster Recovery)

6.2 **PUC 1 – Flood risk assessment and prevention**

6.2.1 **Use Case story**

The test scenario of PUC1 was built based on the flood event of 2010 in Vicenza; the storyboard provided in Deliverable 2.1 underlines the most important figures involved in the emergency management and their role in the "Command Chain" (High level Call requirement V). EOPEN provides a unique platform for all actors involved in the emergency management; the use of satellite images and ground truth information in a scalable environment is the added value (High level Call requirement I) that the PUC1 wants to test.

During emergency all forces work together with a common target that is the emergency support and safety for citizens; the main problem is the fragmentation of information due to the number of actors involved. During an emergency it is very difficult to share information in traditional ways (papers, emails, fax and so on) and it requires a lot of time, so the decision maker can't have a precise and timely view of the emergency. Another issue is that at the end of emergency the lack of a unique database results in losing acquired information from the event.

The lack of powerful computer infrastructure is mainly a limit for the use of satellite images that require Earth Observation knowledge and an appropriate ICT infrastructure. What works is mostly the command chain that is known by all, during emergency every actor works separately with their team, information and decisions are shared at the manager/director level that are physically inside each organization (from COC to DICOMAC).

Table 23: PUC 1 main stakeholders

Stakeholder	Description
ARPAV	This administration is the environmental agency of Veneto Region; inside this office there is also the Copernicus contact person for Veneto Region
Regione Veneto Difesa Suolo	This Administration is the office of Veneto Region Administration responsible for land use, water, environmental planning, and waste management. This office is also responsible for cartography.
Regione Veneto Protezione Civile	This Administration is the office of Veneto Region Administration responsible of Civil Protection Volunteer (at regional scale, so about formation, guidelines etc.), for Bulletins (Alert bulletins)
Genio Civile di Vicenza	This Administration is the operative office of Veneto Region Administration in water management (river maintenance, river project, dikes etc.)
Corpo Nazionale dei Vigili del Fuoco di Vicenza	The firefighters of Vicenza.

Provincia di Vicenza protezione Civile	This is the provincial office of Civil Protection, similar to Veneto Region Civil Protection office but at province level.
Comune di Vicenza	Municipality of Vicenza.
Consorzio APV	This Administration is a Land reclamation authority responsible for ''Alta Pianura Veneta'' basin.
Consorzio Brenta	This Administration is a Land reclamation authority responsible for ''Brenta'' basin.
AAWA (Autorità di Bacino dei fiumi Isonzo Livenza Piave Brenta-Bacchiglione)	This Administration is the higher water authority in Veneto, Trentino-Alto-Adige and Friuli region; it is also responsible for some international basins like Timavo Basin (Slovenia). It is an office directly dependent from the Italian Environmental Ministry.

One of the most important stakeholders, involved in flood management, is AAWA that is also part of the EOPEN Consortium. All important stakeholders were involved in EOPEN during several meetings to "co-design" the platform based on the requirements reported in D2.2. During M23, AAWA organized an event (Infoday) to demonstrate the first prototype of EOPEN. During this event AAWA provided stakeholders with an overview of the platform. At the end of the "Infoday" AAWA provided a second iteration of questionnaires to get a feedback from participants to ensure that most of the user requirements of D2.2 were included and the prototype is aligned with their specifications. Questionnaires sent to stakeholders after the Infoday provided many important feedbacks for EOPEN; in particular the global evaluation of the first EOPEN prototype obtained over 4/5 marks (EOPEN user ecosystem and 3rd party users evaluation).

Implementation

Flood detection product (Water Presence Map, ref. D1.4)

The flood detection product utilized in PUC1 is part of the change detection module and more information can be found in D4.1. Additionally the relevant performance indicators for the flood detection product are mentioned under the Section 3.1 , earlier in this document. The product will also integrate during the validation phase the tweet analysis products developed by CERTH. Tweets will be used as proxies of ground truth information to ensure that flooded areas are correctly identified. Tweets detected as relevant by KR02 are geolocalised and a map in the EOPEN User Portal will show different layers (tweets, flood maps etc.) to give to decision makers all information needed in a comprehensive manner.

Hydraulic model (prototype)

The hydraulic model prototype was prepared in M21 for tests; and it was tested on the platform as a call service. Over the next months, a workflow will be developed to allow users to load model parameters (geometry, boundary conditions etc.) and run the model through the platform. All outputs and that is the depth and velocity maps will be stored and visualized on the platform and can be useful to predict potential flooded areas. The model will be updated during the integration phase (Hydrologic-Hydraulic, M25-30) with

the hydrological module to produce an Early Warning system that can work as a predictor of flooded areas, so as to request in advance pertinent satellite products.

Flood detection product (final) (Water Presence Map, ref. D1.4)

The final version of the flood detection product follows the test and the development of the baseline product. This product will be able to detect flooded areas from optical and SAR imagery but will also include advanced filtering and noise cancelling post-processing techniques. The algorithm will combine both tweets and the results of the hydrological-hydraulic model. The model will work as a predictor for flooded areas.

Questionnaires

A second iteration of questionnaires was produced after the "Infoday" to evaluate the status of the platform and to acquire feedback from stakeholders, in respect to their original requests. Another iteration of questionnaires will be sent after the development of the visual interface of the EOPEN platform, before the final exercise.

Based on the "Infoday" questionnaires emerged a general evaluation of the EOPEN platform, with a score of 4.3/5 and several important conclusions and realizations in regard to the usefulness of particular components of the EOPEN platform. In particular the highest score (up to 4.8 / 5) were on the:

- ability to visualize images from social media=> 4.5/5
- availability of webgis function => 4.8/5
- ability to define of AOI => 4.8/5
- ability to merge different layers (road, floods etc.) => 4.5/5
- ability to use pre-defined processors for non EO users => 4.5/5

The functionalities and capabilities offered by the EOPEN platform to run processes and execute pipelines in a scalable environment was greatly appreciated. Most of the Italian stakeholders claimed that they could use the EOPEN platform and its core functionalities.

6.2.2 Timeline for PUC 1 implementation

Table 24: PUC 1 implementation timeline

Milestone	Completion Month	Status	Comments
EOPEN User Requirement Meeting	M6,	Done	Meeting with PUC 1 stakeholders in Vicenza about EOPEN User Requirement
Questionnaires	M8	Done	
Requirements gathering	M8	Done	Both user and data requirements are defined. These may be refined during the project based on the feedback of the end users.
Development of baseline flood detection method	M16	Done	Developed by CERTH

Integration of the baseline flood detection (ref. Change Detection module) with the platform	M19	Done	Developed by CERTH
Hydraulic model (1st prototype) (Basement)	M21	Done	Transferred to the platform (execution tested manually as proof of concept)
Questionnaires	M24	Done	2nd iteration during infoday in September
Flood detection product (baseline) (Water Presence Map Ref. 1.4)	M23	Done	Collaboration with CERTH
Hydrologic-Hydraulic model	M26	In progress	Under calibration (at AAWA premises)
Development of refined/sophisticated flood detection method (Water Presence Map ref. D1.4)	M26	Done	Defined the workflow to run it
Query on AAWA products – workflow integration	M28	in progress	The tool to query AAWA products (generated at their premises) in EOPEN platform
Integration of refined/sophisticated flood detection module with the platform (Water Presence Map (Ref. D1.4) (tbc)	M28	In progress	
Flood detection product (final)	M28	Done	Collaboration with CERTH, the 2019 quota will be useful to test the algorithm

6.2.3 Monitoring of PUC 1 implementation against the User Requirements

NOTE: The list of products appearing in the following table has been updated and specified in the D1.4 Table 4 – second names are the updated ones.

Table 25: PUC 1 Implementation against the User Requirements

Use case G-id	Users of the EOPEN platform:	Status
PUC1_GA1	Must be provided with capabilities for data dissemination and integration of EO data, weather information and relevant social media text and images.	Development (1)

PUC1_GA2	Must be allowed geo-localisation of social data and real time control.	Done (*)
PUC1_GA3	Should be provided with an intuitive online platform with the possibility to visualise EO data (e.g. from webcam or mobile phone) and possibility to send text message (e.g. SMS or tweet) and analyse its semantics for meaningful automatic decision-making.	Development (2)
PUC1_GA4	Must be enabled to merge different administrative database and formats in a unique platform with all data shared.	Done
PUC1_GA5	Should be enabled to implement EWS and add maps in the platform.	Development (3)
PUC1_GA6	Should be provided with an intuitive and robust interface.	Done
PUC1_GA7	Must be provided with an interactive archive of each event; all data from social network communities and from satellites should be stored in a specific database to provide a history of each event.	Not dev.
PUC1_DA1	DEM/DSM 1m (e.g. Airbus Pleiades) from Copernicus	Delivered (1)
PUC1_DA2	Snow maps with a resolution < 20m	Delivered (2) (**)
PUC1_DA3	Soil moisture maps with resolution < 10m	Delivered (3) (**)
PUC1_DA4	Flood maps (EOPEN AAWA AMICO Early Warning System Flood Forecast/AA EWS FF)	Development (4)
PUC1_DA5	Damage maps (EC EMS Damage Map (or others)/EC EMS DM (tbc))	Development (5)
PUC1_DA6	A- Water presence maps WPM (B- Water Presence Change Monitoring WPCM)	Development (6)
PUC1_DA7	Bathymetry of coast, lakes, rivers	not prov.
PUC1_DA8	Orthophoto with resolution of 50 cm (e.g. WorldView4)	Delivered (7) (**)
PUC1_DA9	Vegetation presence	Delivered (8) (**)
PUC1_DA10	Land cover	Delivered (9) (**)
PUC1_DA11	LAI and other vegetation indexes	Delivered (8) (**)
PUC1_DA12	Other maps (thermal or multispectral data ready to be processed) with high resolution	Not. Prov
PUC1_DA13	Weather forecast	Delivered (10) (**)

User requirements (GA_1 to GA_7)

* Done with the baseline product, which will be updated during the project. All work is performed by CERTH with the help of AAWA as a co-designer.

** Workflows to access them through the EOPEN Platform has to be implemented in the current year.

1- PUC1_GA1: This capability will be better developed in the final version of the visual interface, as of now there are different interfaces for the visualization of maps and layers and for the processing; specifically exists a dashboard with tweets and a GIS Viewer with other data.

2- PUC1_GA3: The final visual interface will be better developed, as of now there are different interfaces for the visualization of maps and layers. The User requirement GA_3 requires that users will visualize the information provided through the platform in a unique interface that will allow users to overlay different layers, visualize tweets (also the text or media) and other information. It will be developed inside the platform as a workflow to send automatic messages (alarms) to users based on weather forecast variables exceeding a certain threshold.

3- PUC1_GA5: The hydrological-hydraulic model is under calibration by AAWA that has provided a first version before the Infoday limited to a test area; the hydraulic solver is also going to be updated (Basement version 3.0.1) and it will run on GPUs instead of classic CPUs. The hydraulic model is an important part of AAWA EWS; now this module runs on a traditional Workstation with high computation cost and time. The new version, developed and used inside (use "on demand") the platform using GPU's can increase the performance of the whole EWS with a reduction of the computational time (as shown in the tables below for the Circular Dam Breach benchmark).
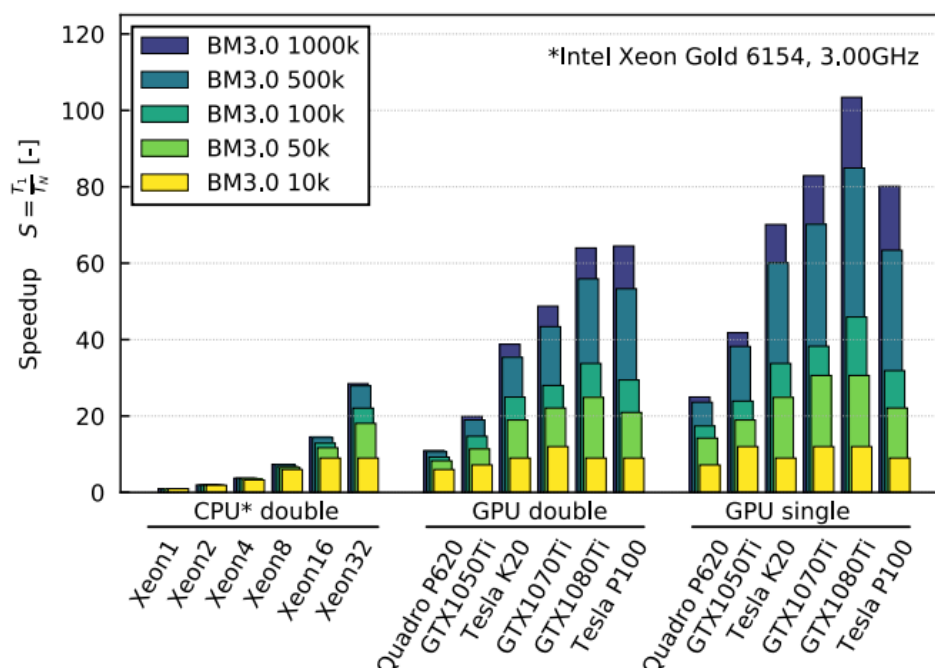

Figure 12: Speed Up for different mesh size, 10K to 1000K elements, CPU vs GPU

The figure above describes the computational time required to perform a hydraulic simulation with Basement 3 (BM3) with different resolution of the computational mesh and different processor types. The increasing of the mesh resolution from 10k to 1000k increases the computational times, also the types of processor influenced this variable; however regardless of the mesh size and the processor type, running a parallel CPU simulation, option allowed only in the newest version of the solver, is much faster of the traditional single solver, used in the previous versions.
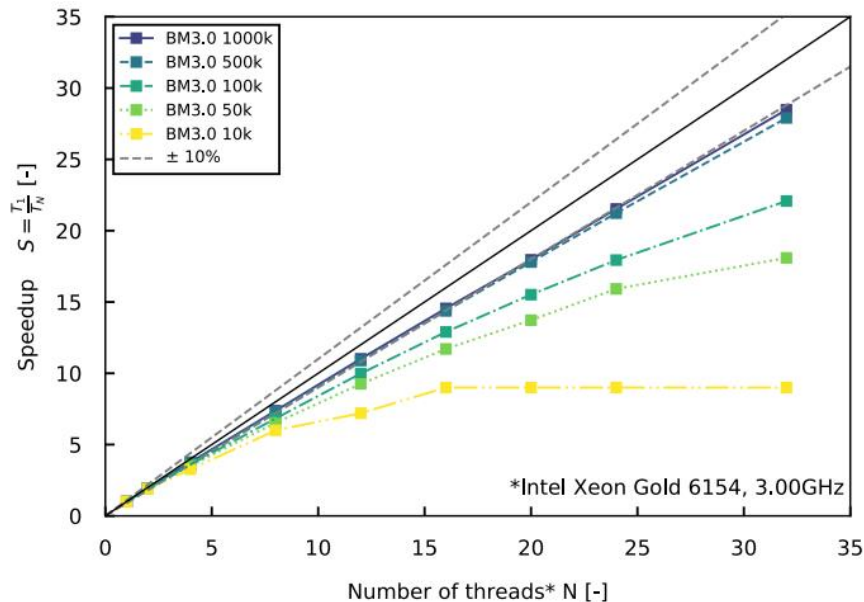
Figure 13: Speed Up for different mesh size, 10K to 1000K elements for a parallel computation with increasing numbers of cores of a fixed type (Intel Xeon 6154 3.00GHz).

The figure above describes the computational time required to perform a hydraulic simulation with Basement 3 (BM3) with different resolution of the computational mesh in case of a parallel CPU simulation. Raising the number of parallel cores used for the simulation increases the speed up factor; in particular for high resolution mesh (1000k) the speedup factor is almost linear increasing the number of cores; while for low resolution mesh there is a maximum number of cores for gaining significant time reduction. It is clear from this figure that the updating of the hydraulic solver from a serial to a parallel computing reduces drastically the calculation and elaboration time.

Datasets (relate to status in the PUC1 requirements table)

1- The DEM/DSM will be derived from AIRBUS TERRASAR-X/TAN-DEM HR Spotlight products
2- Snow maps at different resolution (snow cover < 20m, Snow Water Equivalent > 20m)
3- Resolution > 10 m
4- Maps under development (hydraulic model under update)
5- Access to EMS acquired in May 2019
6- Access to EMS acquired in May 2019
7- HR imagery from WorldView satellite (quota 2018, 2019)
8- Provided many different index (LAI, NDVI etc.)

9- Corine Land Cover Map 2018
10- Weather forecast from FMI (HIRLAM)

6.2.4 PUC1 Implementation

The most important part of EOPEN platform for PUC 1 is the possibility to run an algorithm to detect floods. The WPM (Water presence maps) derived from this algorithm are useful for decision makers to see the situation on the territory. Since AAWA is also the most important stakeholder, those maps allowed a more precise calibration of hydraulic models.

In 2018 the VAIA storm caused many damage inside AAWA District (North East of Italy); through the EOPEN platform flooded areas were derived and compared to other services (mainly Copernicus EMS); in the following part are reported the results of the analysis.

VAIA storm lasted from the 27th of October to the 30th of October 2018.

1) Flood detection products generated by AAWA pre-platform (nov. 2018) with flood detection coupled with ground truth information from local media:

Figure 14: Flood detection products generated by AAWA pre-platform (November 2018)

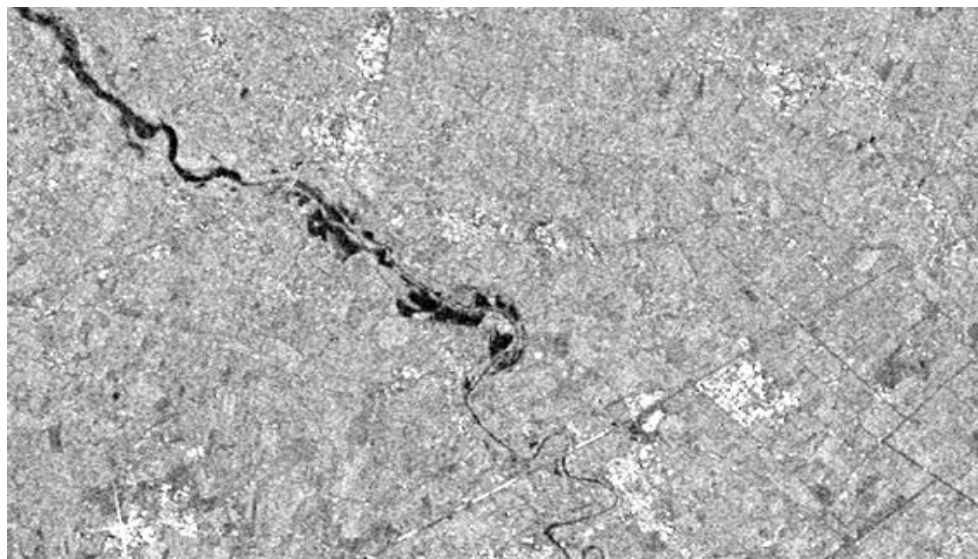2) Sentinel 1 product generated offline with the algorithm by AAWA

Figure 15: Sentinel 1 product generated offline from AAWA

3) WPM generated inside the platform

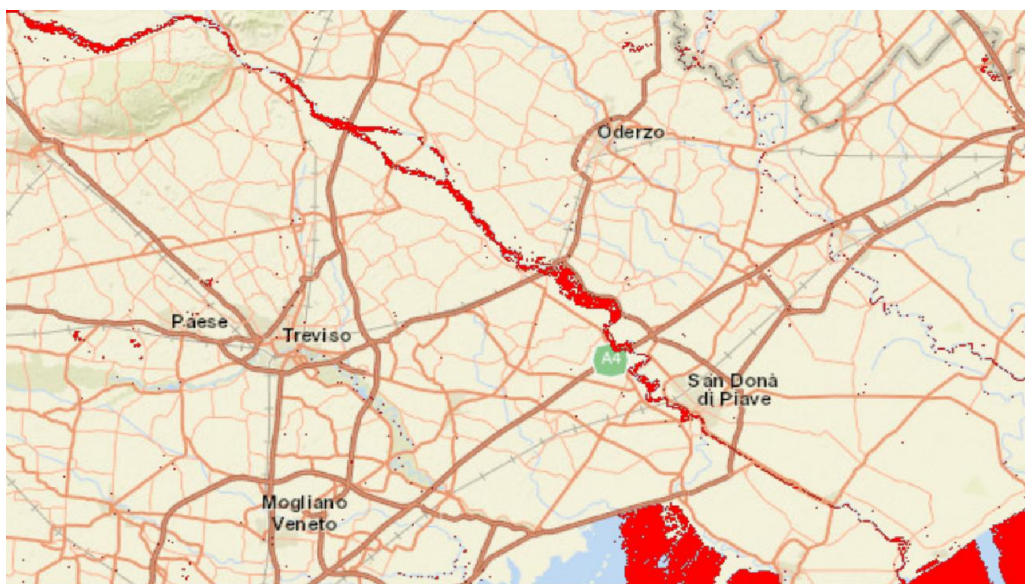

Figure 16: WPM generated inside the platform

Water bodies detected by EOPEN algorithm on the 31st of October 2018.

4) EMS product: EMSR332_09VENEZIA_01DELINEATION_MONIT01_v1, flooded area detected are inside Piave river floodplains (second picture, part inside black lines is the Piave river flow area)
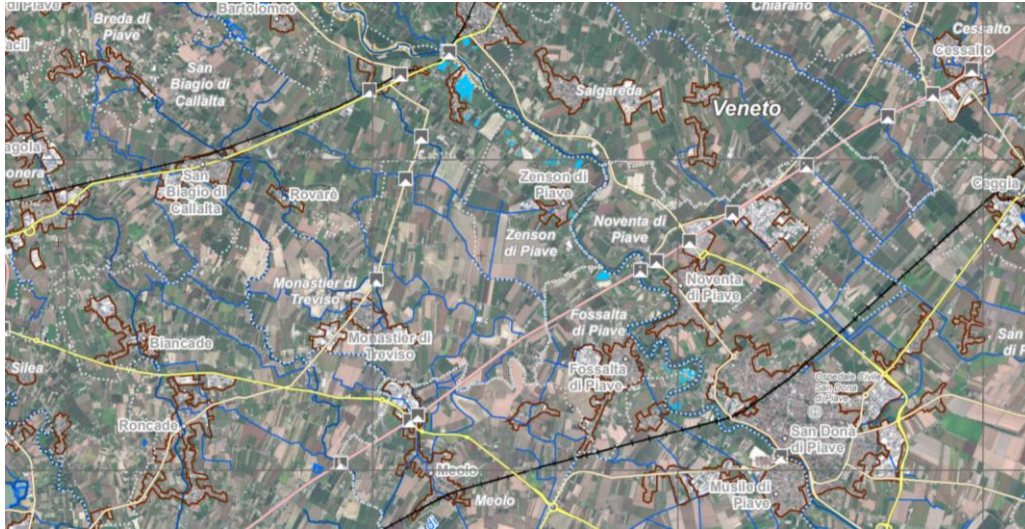
Figure 17: EMS Product of flooded area inside Piave river floodplains

Figure 18: Piave river flow area

5) WPM inside EOPEN for other purposes (Dam monitoring), Corlo dam 10-11-2018, EOPEN WMP and original image, 95% of water detected.
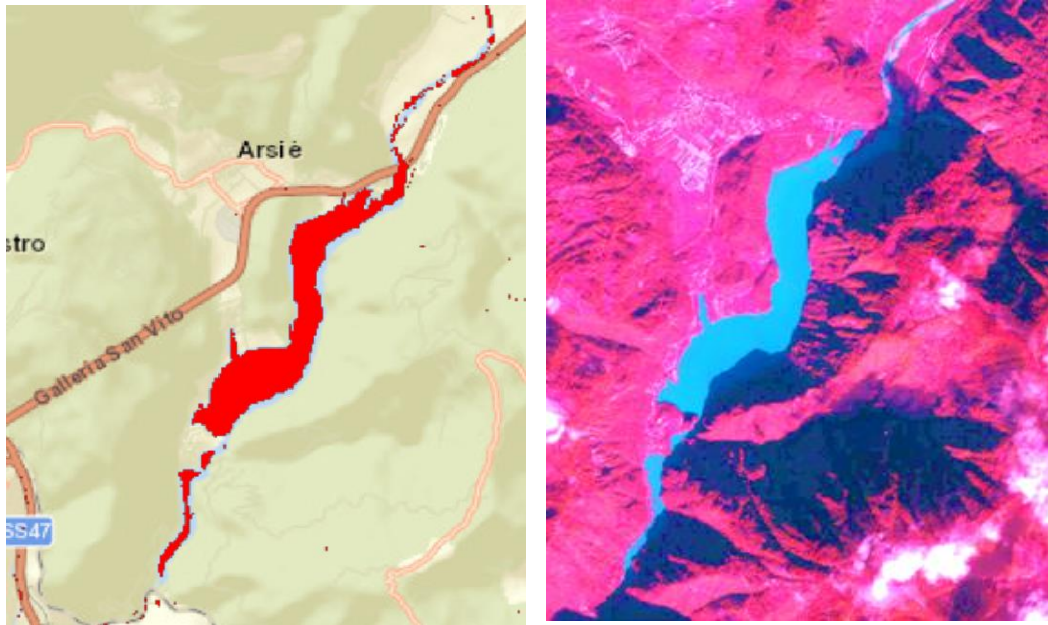
Figure 19: WPM inside EOPEN for other purposes (Dam monitoring), Corlo dam 10-11-2018, EOPEN WMP and original image, 95% of water detected

6.2.5 PUC 1 Implementation workflow

Images below show the workflow of PUC 1 and the workflow that will be implemented for the integration of AAWA EWS (Early warning system) inside EOPEN with some specific information also about the data format that will be exchanged.

AAWA EWS ingests Hirlam weather forecasts (KR13), soil moisture maps, land use maps and maps of vegetation variables (LAI) from Copernicus land Service and provide discharge values of basins inside the Eastern Alps river basin district. AAWA EWS encompasses a hydrological module running continuously and a hydraulic module running "on demand". The hydraulic model developed with the Basement ETH software will run inside the EOPEN platform and will provide important information about the forecast of areas affected by floods in terms of spatial extension, flow velocity and water depth useful to be used along with tweets (ground truth) and future satellite observations.

The hydraulic model run will be triggered by an alert generated in the EOPEN platform. The alert depends on the exceeding of specific thresholds in the last Hirlam forecast.
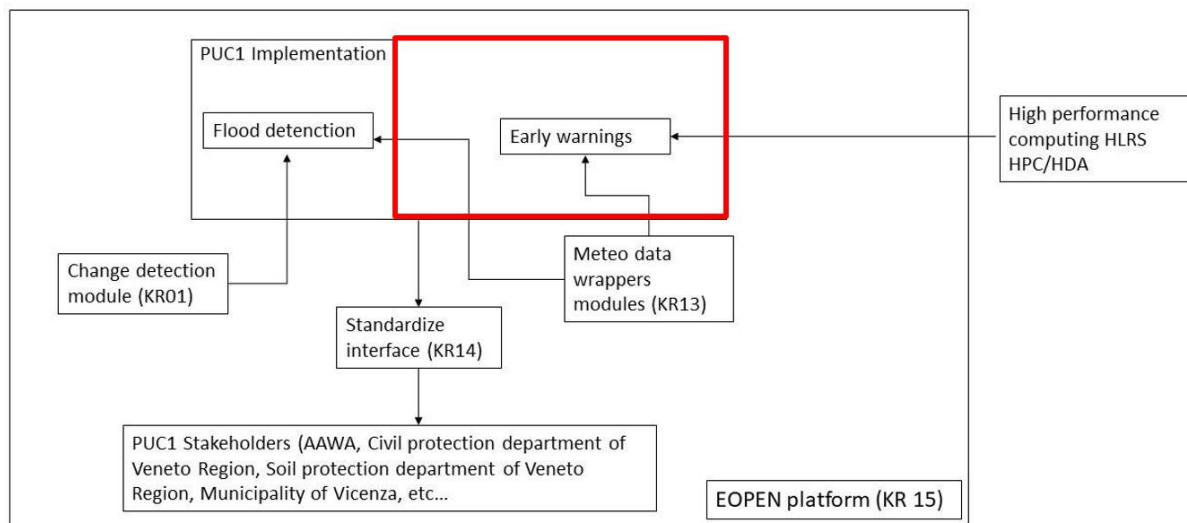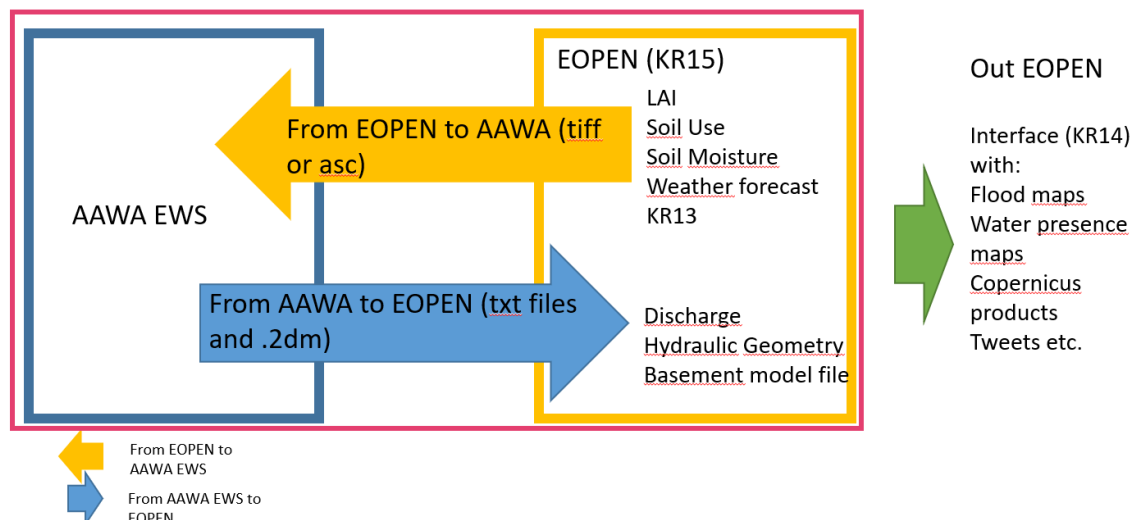
Figure 20: PUC1 implementation workflow

Figure 21: AAWA EMS integration

6.2.6 **Potential Test Scenarios**

Scenario 1 A,B

The first scenario deals with the automatic and near real-time identification of flooded areas through the flood detection product (Water Presence Map) (Water Presence Map) [Note that in the 2nd prototype the Water Presence Map product (PUC1_DA6_a) includes both, flooded areas and permanent water bodies - refinement is under way to generate the actual flooded area, excluding permanent water bodies, which is the area of interest for AAWA, as final version of this product.].

We will evaluate the accuracy of the detection algorithm; if these results can be merged or layered to the administrative database and formats integrated in the EOPEN platform; if the information is presented in a clear and understandable way that could potentially lead to a time reduction to make a decision and subsequent actions during a real flood.

The scenario, part A, consists in the delineation of an AOI with the tool available inside the platform, find the product, select the product and visualize it. After this passage, the user should be able to upload on the viewer its own layers to visualize for example the municipalities affected by flood or the fluvial areas provided by AAWA.

The accuracy of the product will be evaluated with the comparison between the polygon extension of water areas from Italian official cartography (provided by ISPRA) and rasters generated in the platform

Table 25: Evaluation parameters and metrics regarding Scenario 1

Parameter	Evaluation	Metrics
accuracy of the detection algorithm	Using official perimetration of waterbodies by ISPRA we will compare EOPEN products with that	% of area detected

For expert users will be developed a scenario B, where each user will be able to build the workflow inside the platform and generate a product. After this point the user will develop the scenario as described in part A.

Scenario 2

The second scenario is about the flood prediction and the early warning products, involving the flood detection product and the hydraulic model; in particular, EOPEN platform identifies areas at risk of imminent flooding based upon satellite and in-situ data and provides warnings of the likelihood of a flood event based on Hirlam weather forecasts.

In this scenario, it will be evaluated i) the accuracy of the detection of the imminent flooded areas from the satellite images, ii) if these results can be merged or layered to the administrative database and formats integrated in the EOPEN platform, iii) if the data can be integrated with information from social media.

In regard to the early warning product, we will evaluate the time required for the platform to provide new or updated forecasts, as compared to with the current system used, focusing also on the speed up factor of the hydraulic model due to the parallel GPU's and CPU's computation, in comparison to the current serial computation.

Moreover, it will be evaluated if the prediction and early warning products are easily understandable and if they provide a clear outline of the situation, as these aspects could be translated, during a real flood, in time reduction to make a decision and subsequent actions.

This scenario is mainly oriented to AAWA users; since AAWA is also an important stakeholder, this scenario will demonstrate the importance of EOPEN inside AAWA office. The operator, once an event occurs, should access the EOPEN User Portal and inquiry the last Water Presence Map. Water presence maps can show the real situation of the territory and if any flood occurred. To support this task and predict possible floods AAWA will use 2 functionalities of the platform: the meteorological criteria that define an "Event" and send a notification and the hydraulic EWS. Once defined an area that could be affected by floods, AAWA will run on the platform the hydraulic model with as input,

the variables (discharge) derived from AAWA EWS. The hydraulic model will provide an estimation of the flooded area useful for the comparison with satellite data.

Flooded areas derived both from satellite images and hydraulic models will be compared with some ground truth acquired by the social media analysis to confirm the situation on the territory and to provide the decision makers with important datasets.

This scenario will be evaluated as described in the Table *26* below:

Table 26: Evaluation parameters and metrics regarding Scenario 2

Parameter	Evaluation	Metrics
the accuracy of the detection of the imminent flooded areas from the satellite images	Using past events where AAWA owns the delineation of flooded areas made by technicians will be comparing maps inside the platform with those data. Using recent events where there are any surveys, AAWA will compare WPM products with maps already generated by other services (EMS, Civil Protect, etc.)	% of area detected
the merged or layered of the administrative database and formats integrated in the EOPEN platform	This part will be evaluated by using own data (shapefiles, rasters, etc.) inside the EOPEN User Portal	Yes/no criteria
the integration of information from social media		Yes/no criteria and number of tweets detected
the speed up factor of the hydraulic model		speed up factor (AAWA computational time/ HPC computational time)

Scenario 3

This scenario will exploit the EOPEN features of providing flood damage based upon the scale of flooded land and user input data. We will evaluate the precision of the estimated damage, and if the data can be merged or layered to the administrative database and formats integrated in the EOPEN platform. Finally, we will also assess the time and workload reduction to prepare damage reports.

Scenario 4

This scenario deals with the real time monitoring of environmental variables (LAI, Snow cover, Land Use) and meteorological forecasts (i.e. rainfall), in addition to having access to a database of historical measurements.

We will evaluate the User Experience, in particular regarding the ease to access to the historical data; the time required for the System to provide the real-time datasets (e.g. weather forecasts) and to update it once new data arrives; if the interface is clear and robust; if the data can be merged or layered to satellite data and the administrative database and formats integrated in the EOPEN platform; if the data can be integrated with social media derived information.

The users should be able to access, visualize and download several products also from the past, archived inside EOPEN.

Table 27: Evaluation parameters and metrics regarding Scenario 4

Parameter	Evaluation	Metrics
User Experience, in particular regarding the easiness to access to the historical data	the user experience will be evaluated comparing the total time spend in the research of products from different services and the number of accounts needed for access them	time spent on EOPEN/ time spent on the web research; number of EOPEN account/number of account needed
the time required for the System to provide the 'real time' measurement and to update it once new data arrives	This can be evaluate considering the time needed for EOPEN to acquire and visualize, for example, the last FMI weather forecast	seconds
the interface is clear and robust	User should be able to find and visualize several product and download them	Yes/no criteria
the integration of information from social media		Yes/no criteria
if the data can be integrated social media derived information	This parameter will be evaluate visualizing tweets detected for a specific area	number of tweets detected

Scenario 5

This scenario deals with the platform capability that provides the users warnings and notifications of potential extreme weather events.

We will evaluate the User Experience, in particular: the easiness to access data; if the information is clear, easily understandable and if it provides a clear outline of the situation of the emergency.

We will also evaluate the availability of the notification and the time reduction to make a decision and take subsequent actions consequently.

This scenario will start with the development inside the platform of a specific workflow to allow users to set rain rate thresholds on subbasins. The user should, after that, receive an alert message when there is an exceeding of threshold. After that, the user should run the hydraulic model inside the platform and monitor the situation with tweets.

Table 28: Evaluation parameters and metrics regarding Scenario 5

Parameter	Evaluation	Metrics
User Experience, in particular regarding the easiness to access to the historical data	the user experience will be evaluated comparing the total time spend in the research of products from different services and the number of accounts needed for access them	time spent on EOPEN/ time spent on the web research; number of EOPEN account/number of account needed
Provide warning and notification	This part will be evaluated counting the number of correct/incorrect warning sent and unsent during 2 month of observations	number of warning sent/ number of correct warnings; number of warning sent/ number of warning unsent
the integration of information from social media		Yes/no criteria and
if the data can be integrated social media derived information	This parameter will be evaluate visualizing tweets detected for a specific area	number of tweets detected
in time reduction to make a decision and take subsequent actions	This parameter will be evaluate by the comparison between COM and TOM	time before (COM)/time with EOPEN (TOM) this parameter will take into account also learning factors

The evaluation of those scenarios will be implemented in next months and will be reported in D7.4 at M36, starting from the first scenario that is already testable. AAWA

will test all those scenarios as it is the most important stakeholder of PUC1 and will organize an exercise (foreseen also in the GA) where it will test the platform with 3rd party stakeholders.

6.2.7 Overall Assessment, Lessons Learnt and Future Plans

PUC 1 is focused on flood management, during a flood emergency; the most important problem is the lack of information because of the different formats, data sources and timing. EOPEN platform was made under specific requirements acquired directly from stakeholders that are directly involved in the flood management during an emergency, and reported in D2.2. PUC 1 stakeholders include experts that are both familiar and unfamiliar with EO, so they constitute a diverse audience to test the EOPEN developments.

The EOPEN platform can make accessible a plethora of data from different sources (PUC1_GA1, PUC1_GA4), and several Copernicus services, through a scalable system. For EOPEN expert users and non-expert users there are many different tools to make services accessible to everyone; in particular non expert users can use pre-developed algorithms by scientific partners (CERTH, NOA etc.), and can also develop specific workflows within the platform.

EOPEN can make available meteo forecasts with specific time series tools, useful for users and also useful for AAWA EWS. The current status of the platform can visualize information from FMI (forecast KR13), from Copernicus land service (LAI, Land Use) (PUC1_DA2, PUC1_DA3,PUC1_DA9,PUC1_DA10,PUC1_DA11) and data from social media in a web GIS service (PUC1_GA2, PUC1_GA7) that is one of the visualization tools used in EOPEN User Portal.

EOPEN can also provide information and data from the change detection module; water presence maps are already available on the platform. The algorithm runs on the platform analyzing data from Sentinel 1 and Sentinel 2 missions. The Water Presence Maps are produced and stored automatically within the EOPEN platform, so non expert users can visualize them on the web GIS interface (Opensphere). More expert users can use the geoserver utility. EOPEN can also be used within common GIS software, such as a WMS service. This functionality allows users to merge different datasets (PUC1_GA4) and partially satisfy the needs of an intuitive interface (PUC1_GA6).

During the Infoday (M23) all those functions were shown to stakeholders also with some specific parts (technical) that described how to implement a processor or a workflow using the EOPEN platform. Additionally, a workflow to run the hydraulic module of AAWA EWS will also be developed and integrated in the EOPEN platform. This workflow was tested by AAWA and will be developed and integrated during M30. Also the ability to use data from the platform directly inside AAWA EWS will be explored; all maps generated with AAWA EWS will be stored and visualized inside the platform (PUC1_GA5) on the web GIS interface. Finally, the social media crawling and geo-localization is already developed inside the platform.

During the 3rd year of project AAWA will produce some training material supported with KU partners, will organize another demonstration of the platform with the new functionality of the visual interface and of the second prototype and will organise the field exercise.

The "field exercise" will be developed with stakeholders involved in the user requirements meeting held in Vicenza in 2018 and during the Infoday; they are the decision makers that act during a flood emergency. The structure of the exercise will be defined during the 3rd year; AAWA will also organize a demonstration open to external users to demonstrate the platform during the COWM conference in 2020 (activities will be later defined).

Table 29: PUC 1 field demonstration timeline (*)

Milestone	Completion Month	Status	Comments
Training session setup	M31	In progress	Setup of training sessions, we will contact stakeholder to organize training sessions.
Training sessions	M33	In progress	Trainig sessions with stakeholders
EOPEN Field demonstration	M34	In progress	Field demonstration with stakeholders

(*) The timetable could change, due to the Covid-19.

6.3 PUC 2 – EO datasets to monitor Food Security in South Korea

6.3.1 Use Case story

PUC 2 deals with the food security monitoring in South Korea, focusing on the production of rice. Rice is systematically overproduced in South Korea; resulting in large storage costs, while at the same underproduction of other major grains, making the country largely dependent on imports. Based on the collected user requirements (D2.2) and their subsequent translation into EO based products and services, it was decided to implement a food security monitoring system founded on the accurate mapping of rice and the succeeding monitoring of its growth, for the year of inspection, meeting the High level Call requirement V. The table below lists the main stakeholders of PUC 2 products and services.

Table 30: PUC 2 main stakeholders

Stakeholder	Description
Korea Rural Economic Institute	A national agricultural policy research institute focused on the development of agriculture, rural areas and the food industry. The institute is responsible for agricultural monitoring, Free Trade Agreements, World agriculture information, agricultural policies, Overseas crop market information, returning to farm support, international cooperation such as the Korean Agricultural Policy Experiences for Food Security(KAPEX).

National Institute of Agricultural Sciences	A research institute under the Rural Development Administration, which is in charge of testing and research to develop agriculture into a sustainable and competitive biological industry. The role of NAAS is generation of new income by exploration of the values of agricultural resources and maintaining a healthy agricultural ecosystem, crop protection and technical development on practical use of biological resources, development of production techniques for safe agricultural products, development of commercialization techniques for globalization of agricultural and Korean foods, and so on.
KU Lab of Environmental Ecological Planning and Policy Studies(EEPPS Lab, Korea University)	A laboratory of environmental ecological planning and policy studies is managed by Prof. Seong Woo Jeon. The laboratory is focused on environmental ecological planning, adaptation for climate change and environmental policies.
CJ CheilJedang Corporation	The largest general food manufacturer in South Korea, carrying out bio business and manufacture of food, medicine, and feedstuffs and other ingredients such as sugar, flour and oil. Keeping up with market trends CJ cheiljedang analyses social and environmental impacts. They are also dedicated to reinforcing innovative ideas and technological capacities in the fields of food, food ingredient, bio, and feed and livestock. Moreover, they identify the global trend of technologies in the food and bio sectors, and explore new promising businesses.
KU Lab of Agrifood Distribution and Trade(ADT Lab, Korea University)	The laboratory is managed by Prof. Seng-Ryong Yang and focused on agricultural marketing and price analysis. The laboratory presents monthly world food security reports written by the national food security index.

6.3.2 Timeline of PUC 2 implementation

Table 31: PUC2 implementation timeline

Milestone	Completion month	Status	Comments
Questionnaires	M8	Done	

Requirements gathering	M8	Done	Both user and data requirements are defined. These may be refined during the project based on the feedback of the end users.
Recurrent Neural Networks based rice classification product	M16	Done	National scale application, over multiple years of inspection (2016, 2017, 2018).
Random Forest based rice classification product	M17	Done	For the South Korean provinces of Dangjin and Seosan
Change detection – land cover map update methodology (baseline)	M17	Done	Update of the 2015 land cover map
Integration of RNN based rice classification in the EOPEN platform	M18	Done	
End to end integration of the RNN based rice classification product – from data acquisition using the EOPEN Umbrella API to producing rice maps in the front end	M21	Done	
Integration of change detection algorithm	M23	Done	
Integration of the RF based rice classification algorithm in USTUTT's HPDA	M23	Done	Large-scale application using distributed processing
Development of Rice Status Indicator and Rice Yield Estimation (RYE)	M23	In progress	Timeseries of vegetation indices and crop growth indicators. [5]
Prototype demonstration	M24	Done	During ACRS 2019 conference and EOPEN's special session, held in South Korea, PUC2 presented the current status of PUC2 to the audience. The audience included an official of the PUC2 stakodler KREI.

---

[5] Below follows a disclaimer, as reported in D1.4

(*) The ultimate product is considered "high risk high gain". The product refers to the large-scale rice yield estimation for South Korea. Since, ground truth data, such as fertilization usage, cultivating practices, high resolution meteorological or soil data, are not freely available we will only make use of Sentinel data.

6.3.3 Monitoring of PUC 2 implementation against the User Requirements

Table 32: Monitoring of PUC 2 implementation against the User and Data Requirements

Use case G-id	User of the EOPEN platform:	Current Status
PUC2_GB1	Must be provided with an agriculture monitoring system (based on earth observation data) that will provide more accurate cultivation area mapping and production estimates coupled with statistical data.	Done
PUC2_GB2	Must be provided with a crop monitoring solution/platform that will integrate spatial and statistical data.	Done
PUC2_GB3	Should be enabled to overlay the GIS data of "soil atlas" with remote sensing data. **Hint**: *National Institute of Agricultural Sciences provide a system called "Soil Atlas" but this system shows the agricultural land area according to land usage which does not show the actual crop production n area and location.*	In progress NOTE: All PUC 2 products will be provided standard formats and will be completely reusable. Both downloadable but also available on the platform that allows users to import their own data as well.
PUC2_GB4	Should be provided with services that account for big data handling, meteorological data coupling with statistical yield estimations and satellite data.	Done Products: Paddy Rice Mapping, Rice Status Indicator
PUC2_GB5	Must be enabled to integrate different platforms to the governmental statistics system.	N/A
PUC2_GB6	Must be provided with data at farm level (e.g. crop type classification, etc.)	Done
PUC2_GB7	Must be enabled to download agriculture related information through web.	Done
PUC2_GB8	Should be enabled to receive agriculture related information through reports.	N/A NOTE: Information will be provided in the form of raster and vector layers, which is deemed to be far more efficient, intuitive and useful than a report.

PUC2_GB9	Should be enabled to integrate the produced agriculture related information into the national statistics system.	In progress The output products will be provided in appropriate format for integration	
Use case D-id	**List of Datasets requested by stakeholders**	**Status**	**Description**
PUC2_DB1	High resolution remote sensing imagery (a. Timseries of vegetation indices and crop growth indicators, b. Timeseries of rice maps at 10 m spatial resolution – ref. D1.4	Done	Timeseries of vegetation indices, crop growth indicators and rice maps at 10 m spatial resolution
PUC2_DB2	Detailed meteorological observation data	In progress	Air temperature, Wind speed and direction, Precipitation, and Solar Radiation
PUC2_DB3	In field inspection data	Done	Land cover maps available with 10 yr interval. These maps are updated through change detection for the year of inspection
PUC2_DB4	Farmers' claims data	Not available	
PUC2_DB5	Accurate yield statistics[6]	Done	Rice Status Indicator
PUC2_DB6	EO based production status		
PUC2_DB7	Statistical data on national fertilizer usage	Not available	
PUC2_DB7	Statistical data on crop yields	Not available	Rice Status Indicator (Biomass and Yield Indicators)
PUC2_DB8	Social media	Done	Event detection module

6.3.4 **Implementation**

PUC2 deals with food security monitoring, which demands the large-scale application of its products and services. PUC2 uses exclusively long time-series of high resolution imagery from the Sentinel missions, which results in large volumes of data to be processed. Indicatively, the combined time-series of Sentinel-1 and Sentinel-2 imagery for the districts of Seosan and Dangjin, use in the RF based rice classification, amount to more than 80 GB. Therefore, it was apparent from the start that PUC2 requires the HLRS HPDA infrastructure to effectively manage and process such big data.

NOA invested significant effort to familiarize with the HPDA environment, set up the required libraries and implement the distributed classification pipelines of k-means and Random Forest. The process proved tedious and frustrating but ultimately produced

---

[6] See D1.4 - Note in Table 7

results of high value, finding solutions for all big data challenges – i.e. data management, distributed processing, generalization, transferability and computational complexity.

Paddy Rice Mapping

The first principal component of the implementation of PUC2 is the accurate, efficient and large scale mapping of rice in South Korea. Korean Ministry of Environment provides land cover maps in a fishnet grid, as shown in the figure below. The grid cell size of the level-3 land cover map is about 6.4 square km.
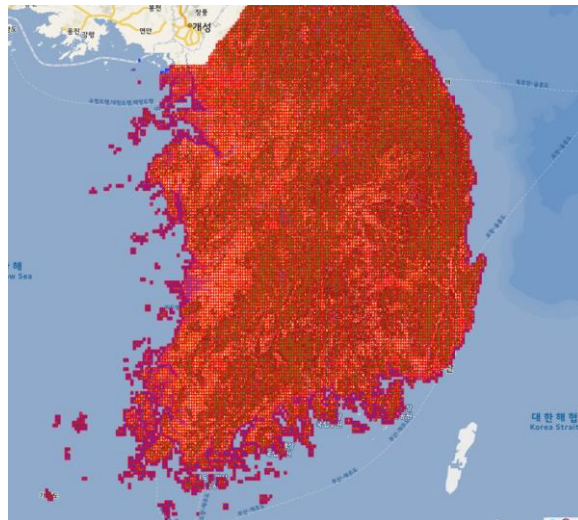

Figure 22: Fishnet grid of land cover maps in South Korea

Below we show the provinces of Dangjin and Seosan, whose land cover maps of 2015 were used to train the classification models.
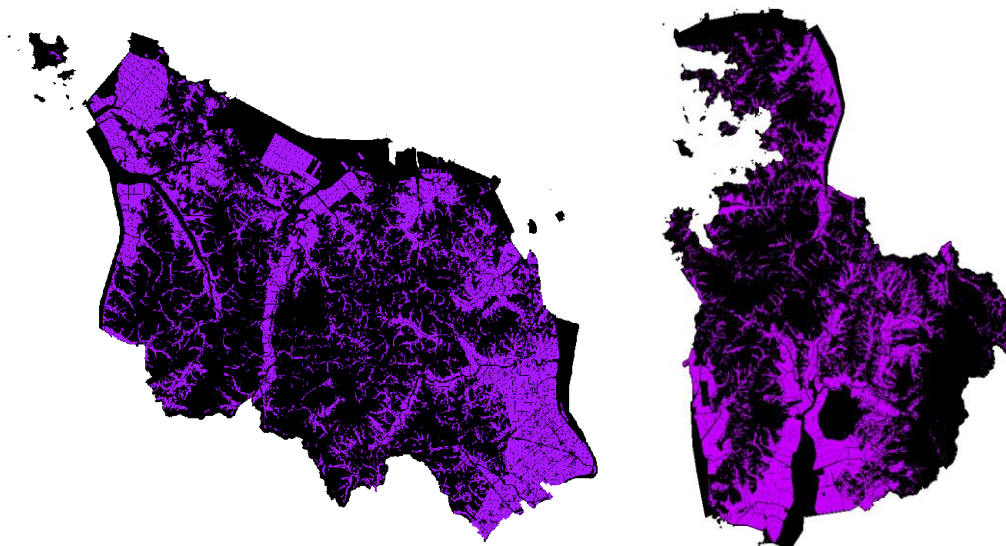
Figure 23: Provinces of Dangjin (left) and Seosan (right) showing rice fields with purple based on the level-3 land cover map distributed in 2015

The rice mapping classification framework for PUC2 involves two different approaches to be compared in terms of effectiveness, efficiency and generalization. The first involves the utilization of a time-series of Sentinel-1 images from 2015 to train a recurrent neural network model using the land cover maps, as presented above, and then fit the model on a Sentinel-1 time-series from 2016, 2017 and 2018. The second method uses a combined input dataset of both Sentinel-1 and Sentinel-2 imagery to train a Random Forest classifier based on a generated updated land cover map for the year of inspection (currently 2018) or based on a land cover map created from scratch using a clustering method. The methods for creating training datasets used for the classification process are presented next.

Generation of Training Data

A.   Land cover map update

This development has been already fully reported in deliverable D4.1 EO assisted change detection, and is fully described in Section 3.1 of this document. Detailed land cover maps at the parcel level, distinguishing between rice and non-rice cultivations, are produced every few years in South Korea. However, in order to systematically produce, for any given year, the products of rice mapping, rice status indicator, biomas indicator and yield indicator - updated land cover maps are required for the appropriate training and validation of the machine learning algorithms utilised. Thus, as explained earlier, an updated land cover map for 2018 is generated as a training set for the supervised rice classification algorithm.

B.   K-Means algorithm

Training data is essential for supervised machine learning algorithms. However, especially in remote sensing, ground truth data is not always available. In addition, ground truth data may be available for a certain district or region preventing the generalization and transferability of a model to another area, of different characteristics. Hence, the generation of a training dataset, given the scarcity of freely available ground-truth data, is of utmost importance. In this scenario a dataset has been generated assigning each pixel to a value that shows if it is rice or not; to achieve this, a k-means algorithm was used. As described in Section 3.5 , several executions of distributed k-means run on temporally interpolated vegetation indices in order to identify rice fields. This final rice cluster yields a dataset for the supervised rice classification, resolving the absence of ground truth data.

C. Data augmentation

In order to overcome the shortage of labelled data, a data augmentation technique is proposed, which attempts to simulate the phenological differences of rice in other regions. Figure 24 illustrates this method, with which each acquisition is shifted one time period backward as well as one time period forward, thus increasing the volume of the training data by three times.
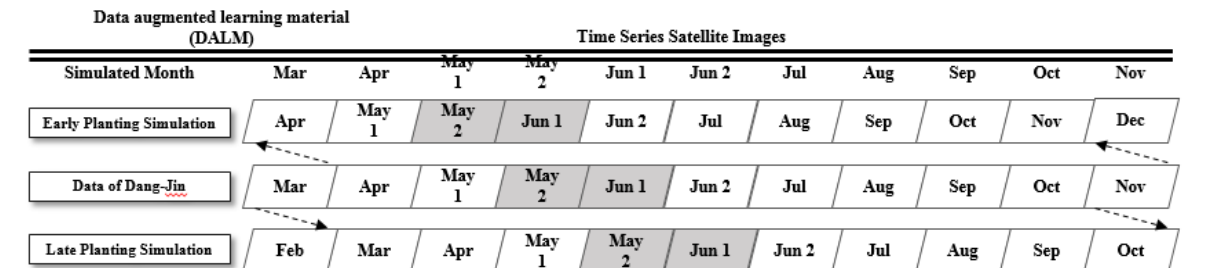
Figure 24: Data augmentation technique

Paddy Rice Mapping

Moving to the classification process two different approaches were implemented, the first one based on recurrent neural networks and the second on distributed random forests. Two different methods are produced to examine the relevant differences in processing complexity (RNN is faster), accuracy (RF is more accurate), and generalization (better generalization with the extended feature space of RF solution). These methods are described below:

Method 1 – Recurrent Neural Networks

The first approach for rice paddy classification implements a Recurrent Neural Network (RNN), which is a time-series analysis method. The model was developed using a time series of Sentinel-1 (IW: VV) images, and trained with the dataset of 2015's Dangjin city land cover map, using the data augmentation method C. Dangjin city is the greatest rice producer in South Korea. Validation was performed in two ways. First of all, detailed pixel-by-pixel comparison was performed in Dangjin and Seosan cities by using Level-3 land cover product, distributed by Korea Ministry of Environment. The accuracy was recorded at 91.46%.

To verify its general applicability in South Korea, the result was also compared with rice paddy area statistics of 160 cities in South Korea, which is distributed by Korea Statistics. A linear regression between the model result and statistics was "y = 0.8791x –470.5" (y: model result, x: statistics), and the R2 was 0.9691. The algorithm has been fully integrated (end-to-end) on the EOPEN platform.

Method 2 – Random Forests

Using a long time-series of Sentinel-1 VV backscatter products, along with cloud free Sentinel-2 imagery and pertinent vegetation indices, we create a very deep feature space for classification. The figure below illustrates the high level workflow of the methodology. The workflow starts by searching for the appropriate input data using EOPEN's Umbrella API, which was developed as part of task T3.1. EO data acquisition and more information the service can be found in D3.1. The ingested time-series of imagery is then appropriately preprocessed and useful vegetation indices are computed for each Sentinel-2 acquisition (NDVI, PSRI, NDWI and SAVI)[7]. A subset, 10-20% of the updated land cover map is then used to train the Random Forest model, which it thereafter applied on the remaining dataset.

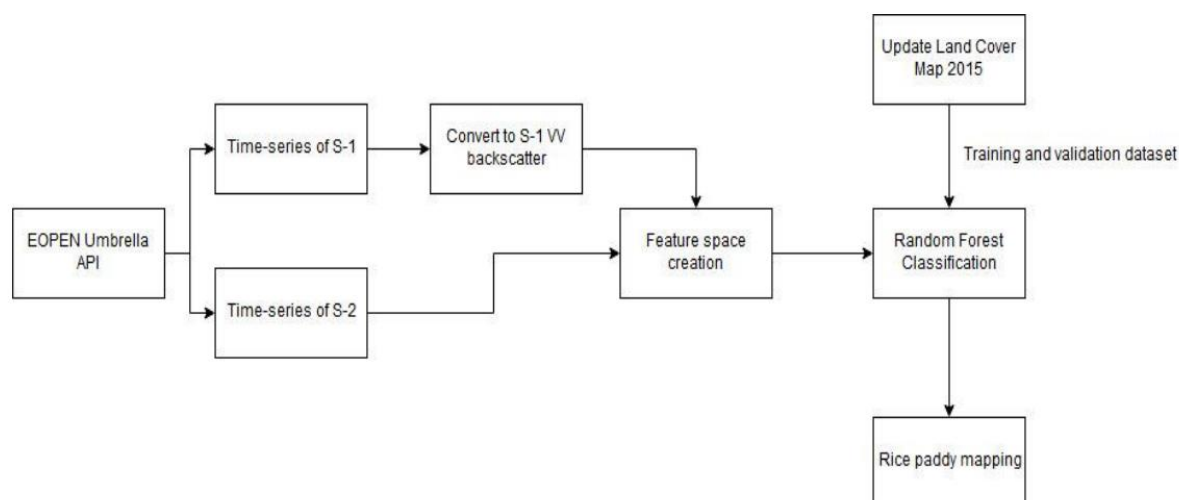[7] https://www.sentinel-hub.com/develop/documentation/eo_products/Sentinel2EOproducts

Figure 25: Random Forest rice paddy classification workflow

The rice classification method has been deployed in HLRS's High Performance Data Analytics (HPDA) module, exploiting the distributed processing solutions (Apache Spark) in order to achieve a large-scale scale application (High level Call requirement I). This approach constitutes a Big Data problem in terms of volume of data (multiple TB of information), processing complexity and scale of application and thus cannot be realized in conventional machines. Table *33* illustrates the time performance of the Random Forest Classifier using 1 node compared to the time performance using 8 nodes. The parameters of the algorithm were set to high values in order to increase the complexity, and thus the time comparison will be more precise.

Table 33: Number of nodes vs processing time

Number of nodes	Processing time (s)
1 node	17,826
8 nodes	7,022

Rice Status Indicator and Rice Yield Estimation

The RSI methodology was implemented using the TIMESAT software (Eklundh et al., 2015). TIMESAT is a software-package, which analyzes satellite derived time-series, using regression fitting approaches. In addition, TIMESAT software provides three different fitting approaches, which are based on least square fits to the upper envelope of the time-series data. For the case of the rice status indicator, Normalized Difference Vegetation Index (NDVI) time-series were analyzed in order to obtain growth information for rice. Among the three available processing methods, the adaptive Savitzky Golay filter was used, as it captures subtle and rapid changes in the time-series ,providing a better understanding for the beginning and for the ending of the growing season. Moreover, the adaptive Savitzky Golay filter, as a form of moving averaging, succeeds in smoothing the

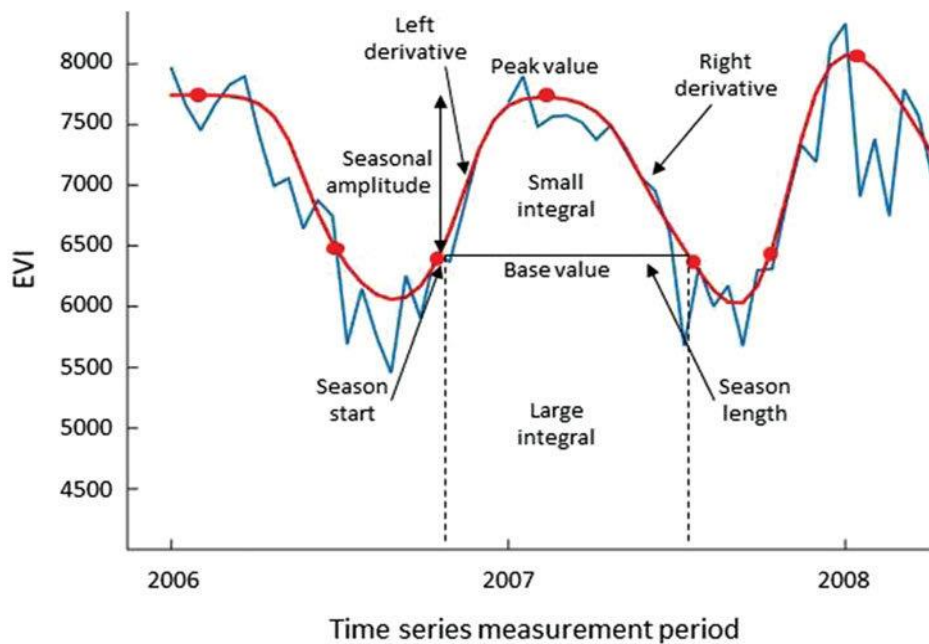NDVI growth trend curves, and at the same time fitting the time-series without reshaping the raw data.

Figure 26: Example of TIMESAT smoothing process using EVI time-series and the calculated phenological metrics

In order to analyze the NDVI curves, TIMESAT needs even spacing between the provided time-series. For the purpose of creating a robust cloud free time-series, a number of interpolation methods were investigated. Ultimately, the weighted average and linear interpolation methods were utilized in order to create decadal time-series from the raw data acquisitions. The weighted average interpolation method fills the fixed timestamps (dekadal time-series) that fall within +- five day range, from the collected satellite derived imagery, and then linear interpolation is followed to fill the remaining decadal dates.

A Median Filtering method was employed with multiple iterations for spike removal, as outliers may seriously degrade the final function fit and interfere with the final width and weight of the annual NDVI curve.

TIMESAT required a user-defined percentage level parameter, which describes the season start (emergence) and end (harvest). In particular, TIMESAT requires a user specified threshold level of the distance between left minimum level and the maximum of the smoothed curve for the time of the start of the season (SOS) and the level of the distance between the right minimum level and the maximum for the time of the end of season (EOS). In order to overcome this, and create a generalized crop growth retrieval methodology, two more indices were investigated. Analyzing the time-series of Normalized Difference Water Index (NDWI) and the Plant Senescence Reflectance Index (PSRI), and taking into account the temporal cultivation practices relevant for the rice crop in South Korea, the SOS and EOS parameters were determined. Specifically, the transplanting (emergence) practice was detected through the NDWI time-series analysis for specifying the SOS and the harvesting through the PSRI for the EOS.
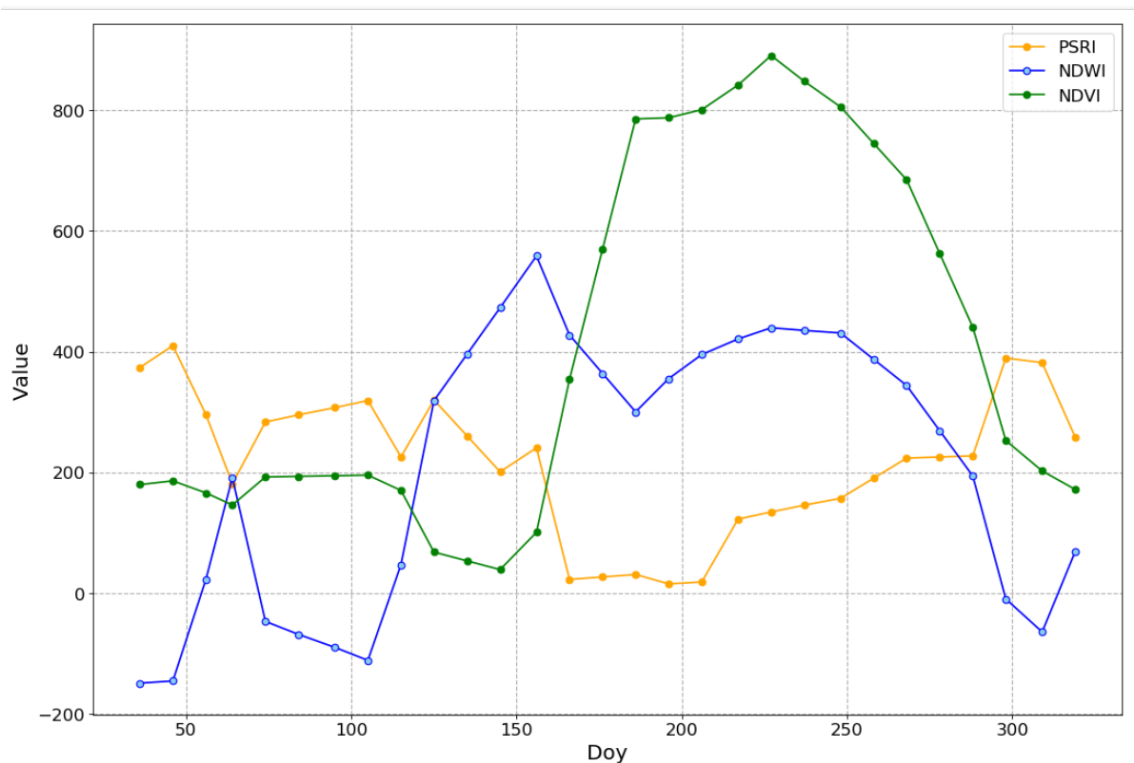
Figure 27: Interpolated time-series of NDVI, NDWI and PSRI of one parcel

The products of the TIMESAT software are eleven phenological metrics that were computed as seasonality parameters and are presented at the following table:

Table 34: Explanation of Phenological Metrics

Phenological Metric	Explanation
Time of the start of season (SOS)	Time for which the value of the index has increased to a defined level (percentage of the distance between the left minimum level and the maximum)
Time of the end of season (EOS)	Time for which the value of the index has decreased to a defined level (percentage of the distance between the right minimum level and the maximum)
Length of season (LOS)	Time from the start to end of the season
Base level (BL)	Given as the average of the left and right minimum values

Time for the mid of the season (MOS)	Computed as the mean value of the times for which, respectively, the left edge has increased to the 80% level and the right edge has decreased to the 80% level
Largest data value for the fitted function during the season (MaxV)	It may occur at different time compared to MOS
Seasonal amplitude (Amp)	Difference between the maximum value and the base level
Rate of increase at the begging of the season (RIBOS)	Calculated as the ratio of the difference between the left 20% and 80% levels and the corresponding time difference
Rate of decrease at the end of the season (RDEOS)	Calculated as the absolute value of the ratio of the difference between the right 20% and 80% levels and the corresponding time difference. The rate is thus given as a positive quantity
Large seasonal integral (LSI)	Integral of the function describing the season from the season start to the season end
Small seasonal integral (SSI)	Integral of the difference between the function describing the season and the base level from season start to season end

The rate of increase in NDVI during the beginning of the season can be related to the physiognomy of the vegetation, and the green up rate, as it describes velocity at which the crop moves from the germination stage to the growth stage of the plant. On the other hand, the rate of decrease describes the rate of senescence at which the crop shifts from the flowering phase to the ripening stage, where the harvesting begins. Additionally, the LSI can be used as a proxy of the relative amount of vegetation biomass without regarding the minimum values, and the SSI as a proxy of the relative amount of vegetation of biomass while regarding the minimum values (Jönsson et al., 2004).

Moreover, apart from the metrics described above, two more integrals, which are used as indicators, are introduced. The definition of the Start of Season and the Peak of Season can be characterized as crucial, as the left integral defined in combination with the base level, establish a total accumulated biomass indicator of the plant. At the same time, after the finish of the phenological cycle of the crop, the Peak of Season and the End of Season along with the base level, define a crop production indicator (Araya et al., 2018). In addition, high values indicate smaller reduction in the accumulated biomass after the flowering stage of the plant, which are associated with a slower and longer process of

producing shoots and fruits, whereas small values represent a shorter process that has not been completed filling shoots and associated with low yield.
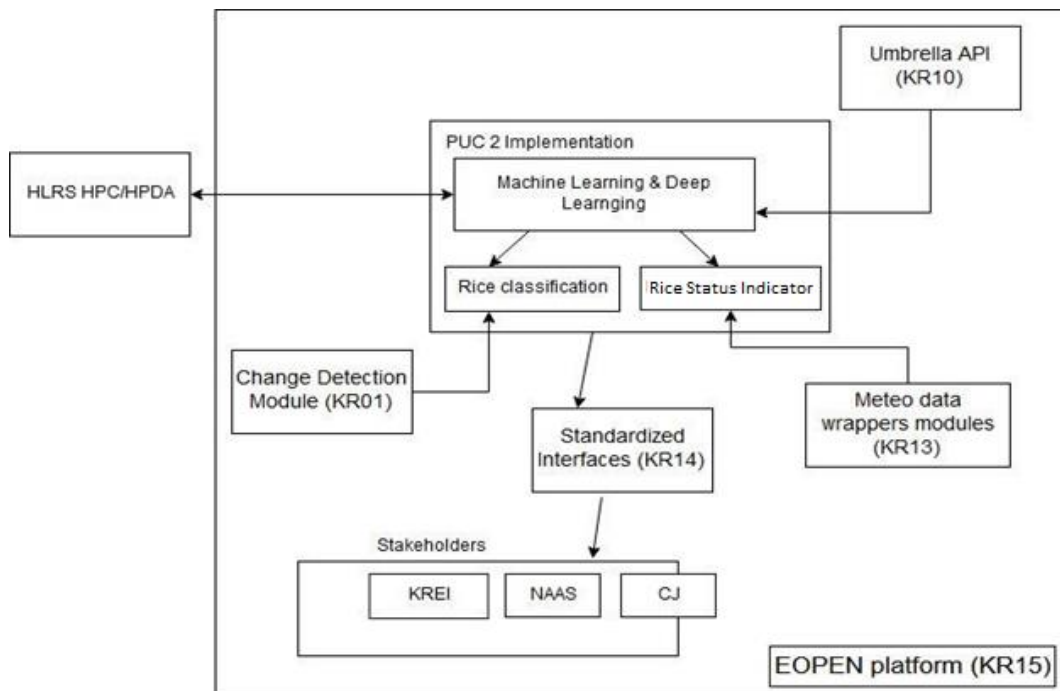
6.3.5 PUC 2 Implementation workflow

Figure 28: PUC2 Implementation workflow

6.3.6 Test Scenarios

Four scenarios were implemented during the evaluation stage. The first one is related to the recurrent neural networks approach while the next two to the random forest method. The results of both rice classification methods and of the rice status indicator will be presented in the form of four test scenarios. All of the rice classification scenarios are validated against a dataset of approximately two million pixels that has been annotated through elaborate photointerpretation, using three different evaluators.

Scenario 1

The first scenario involves the RNN rice paddy detection algorithm that trains based on the land cover map of 2015. The trained model is applied to a time-series of Sentinel 1 images, at national scale, for the years 2017-2019. This way the generalization of the model is evaluated, identifying the significance of the accuracy decrease in the succeeding years of 2015.

RNN rice paddy detection algorithm that trains based on the land cover map of 2015 was applied to the entire nation for the years of 2017-2019 in order to validate its general applicability over diverse spatio-temporal dataset. The years of 2015 and 2016 were not applied because of not enough time series satellite images or qualitative problems appeared on the images. The three consecutive rice paddy detection results were

validated with national rice paddy area statistics (Figure 29), which is produced by a "crop production survey" on the basis of a field survey of more than 50 thousand sample areas.
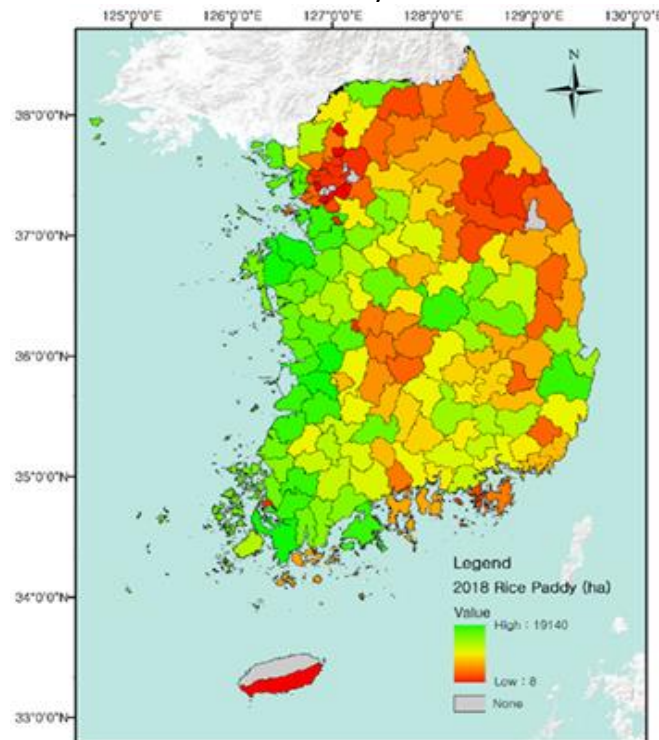

Figure 29: National area paddy rice statistics

The result of this scenario indicated that the algorithm had performed successfully for all of the consecutive years, with explanatory power more than 0.96; while there exists minor over-estimation trends, especially in 2018 and 2019. The rice paddy areas in 160 cities were constantly decreasing from 2017 to 2019 with recording about 748, 728, and 724 thousand ha for each consecutive year, but the detected areas were 748, 829, and 819 thousand ha. According to the visual interpretation, most of the misclassifications had occurred near the boundaries of rice paddies. In addition, trend lines with high explanatory power indicate a high possibility of classifying boundaries to rice paddy. Therefore, if such overestimation trend continues, multiplying a coefficient to the result could be considered to match it with the national statistics.
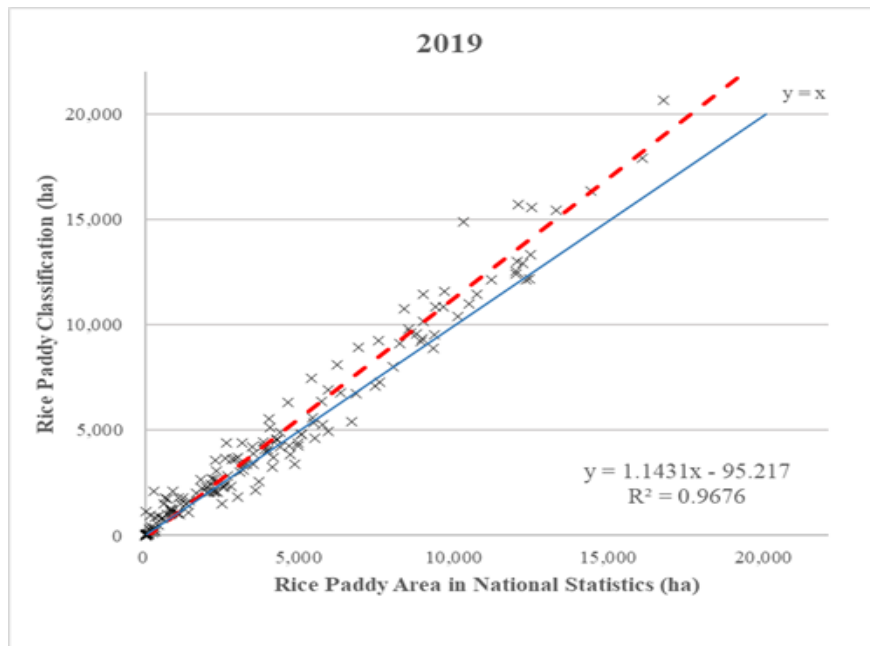
2017

$$y = 1.0561x - 313.65$$
$$R^2 = 0.9752$$

2018

$$y = 1.1186x + 88.672$$
$$R^2 = 0.9631$$
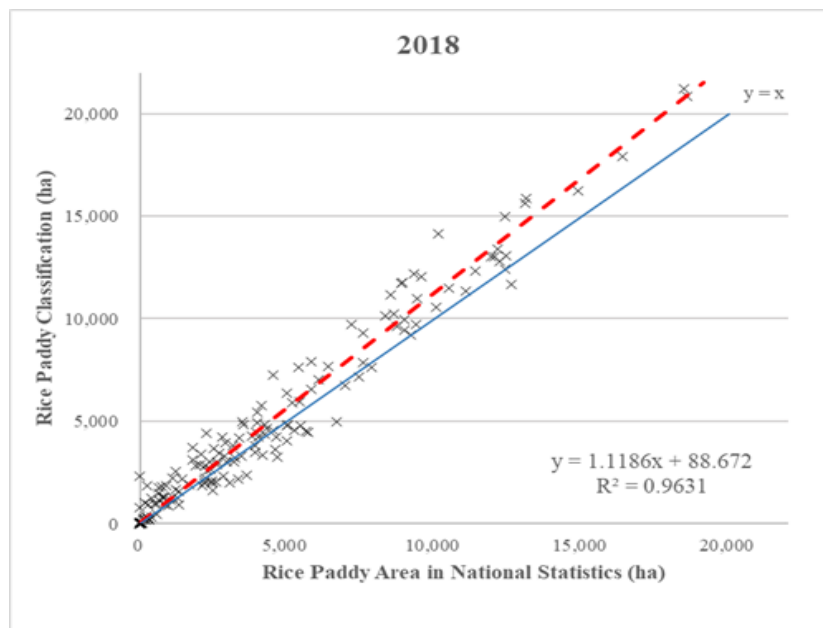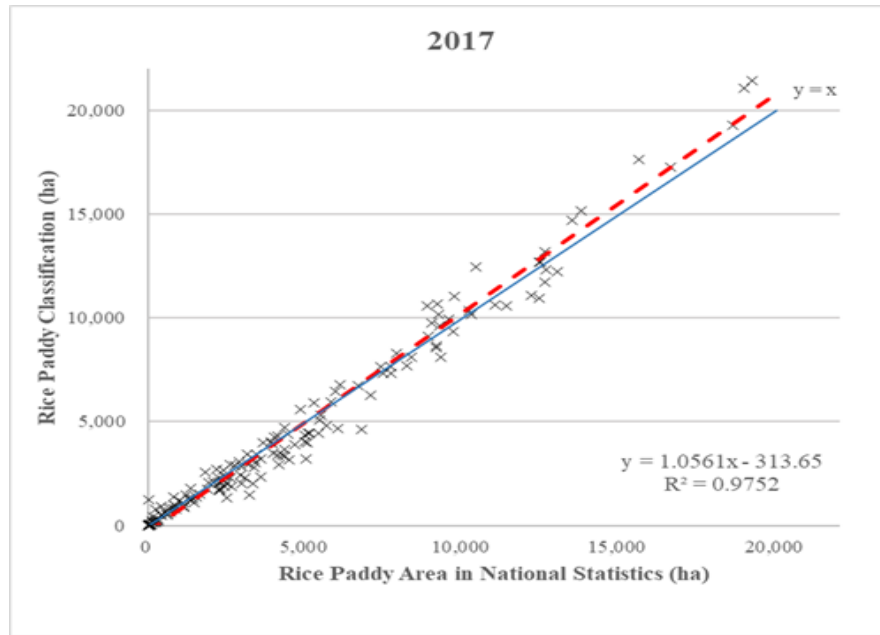
Figure 30: Results for rice paddy trends in 2017, 2018 and 2019

Scenario 2

The second test scenario involves the Random Forest rice classification approach that utilizes the dynamic land cover map update for training. We trained the model using the generated training data of Seosan and Dangjin and applied it on other regions. Additionally, the Seosan-Dangjin trained model is tested in other provinces against locally trained models. This way we can identify an operational framework for rice mapping, evaluating the generalization and transferability of the scheme.

A random forest rice paddy classification algorithm was trained based on the updated land cover map of Seosan (see data augmentation method A). The most important parameters of this algorithm are the maximum depth and the number of trees. Multiple combinations of these two parameters were tested in order to use values that avoid both underfitting and overfitting. Figure 31 and Figure 32 display the precision and recall scores obtained when we the Seosan trained model to Seosan and Haenam provinces[8]. Number of trees parameter set to 9 and the depth parameter set to 8 lead to the best generalized model. Moreover, between the two parameters, depth presents a more significant effect to the performance and the generalization of Random Forest for this problem. Higher values create overfitting while lower create some underfitting.

---

[8] When the Seosan trained model refers to 30% of the Season land cover map, the remaining 70% is used for testing the model.
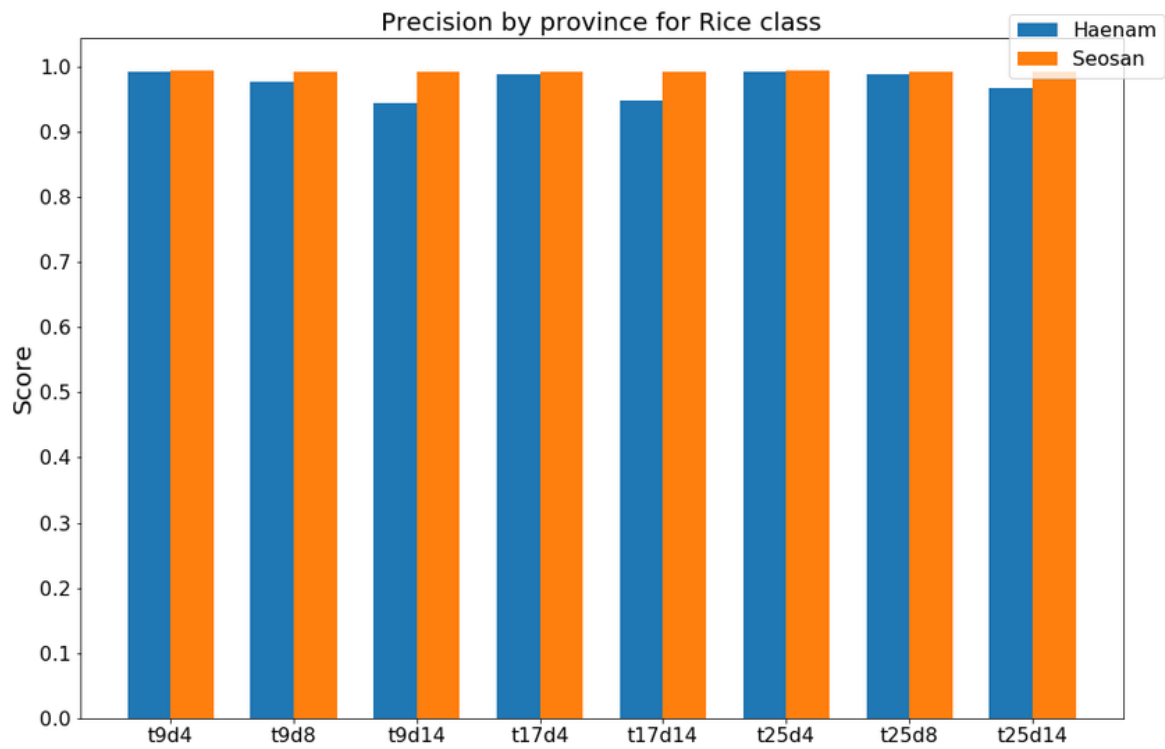
Figure 31: Precision scores obtained by predictions of random forest models trained on Seosan province using the updated land cover map (training data A), both in Seosan (orange) and Haenam (blue), for different values of number of trees (t) and depth (d) parameters.
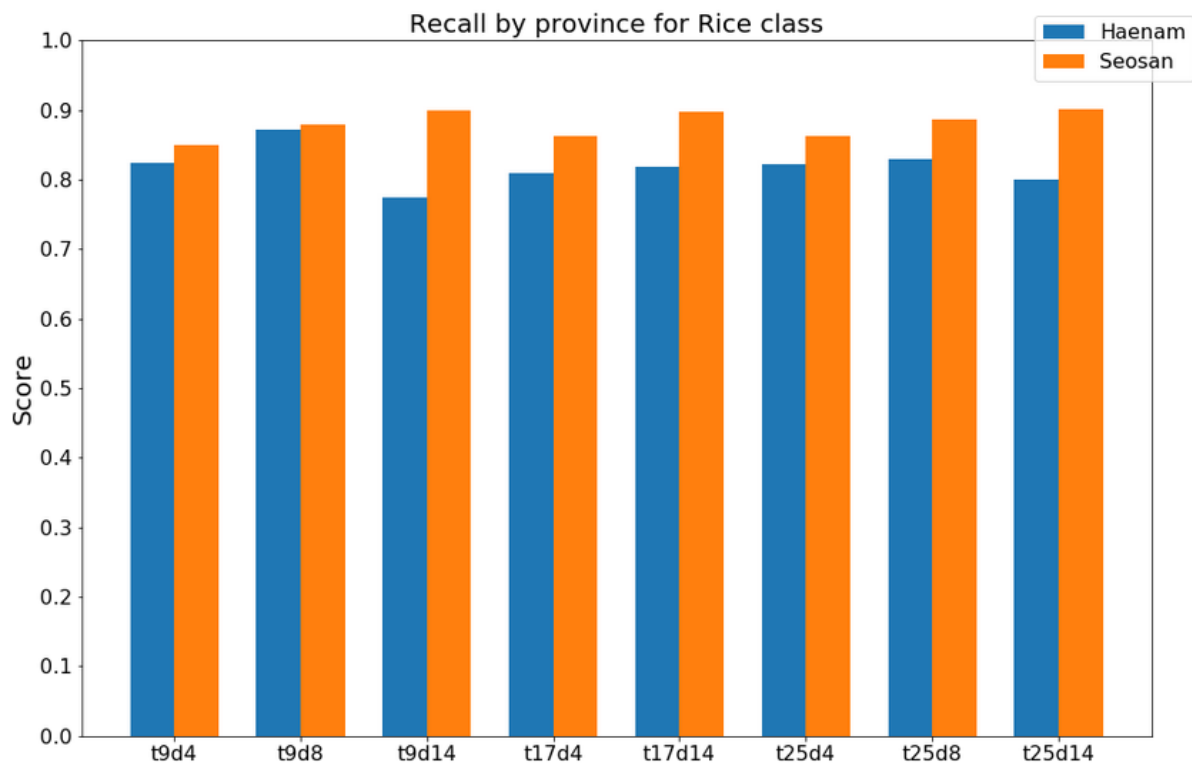
Figure 32: Recall scores obtained by predictions of random forest models trained on Seosan province using the updated land cover map (training dataset A), both in Seosan (orange) and Haenam (blue), for different values of number of trees (t) and depth (d) parameters.

Additionally, we have plotted the time evolution of precision and recall for models that are trained with an increasing number of features (new acquisitions). This was done to discover how accurate the model is on each period of the year. At first the input of the model consists of acquisitions from February until May and for each next run, acquisitions of the next month are concatenated to the input. As Figure 33Figure 34 indicate, the generalization of the random forest model performs poorly early in the year, but after August it produces very accurate results.
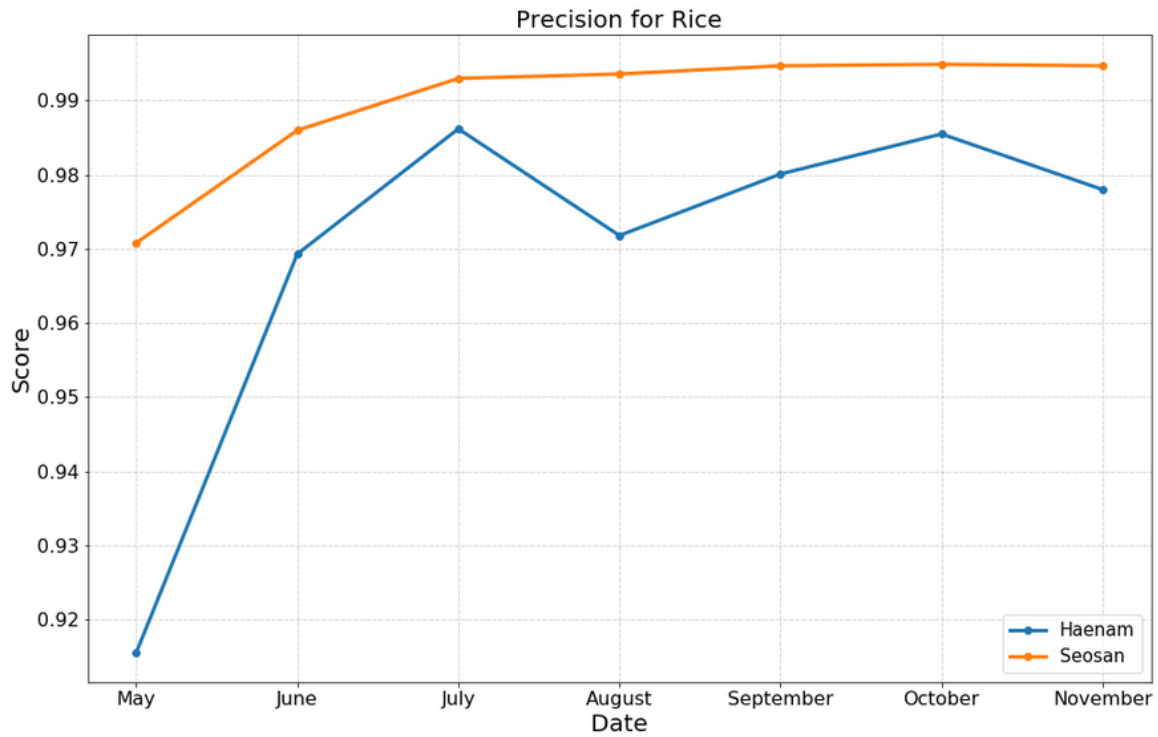
Figure 33: Precision of Rice class of predictions obtained by evolutionary Random Forest models trained on Seosan with the updated land cover map. The blue curve refers to prediction of Haenam province and the orange curve to those of Seosan province.
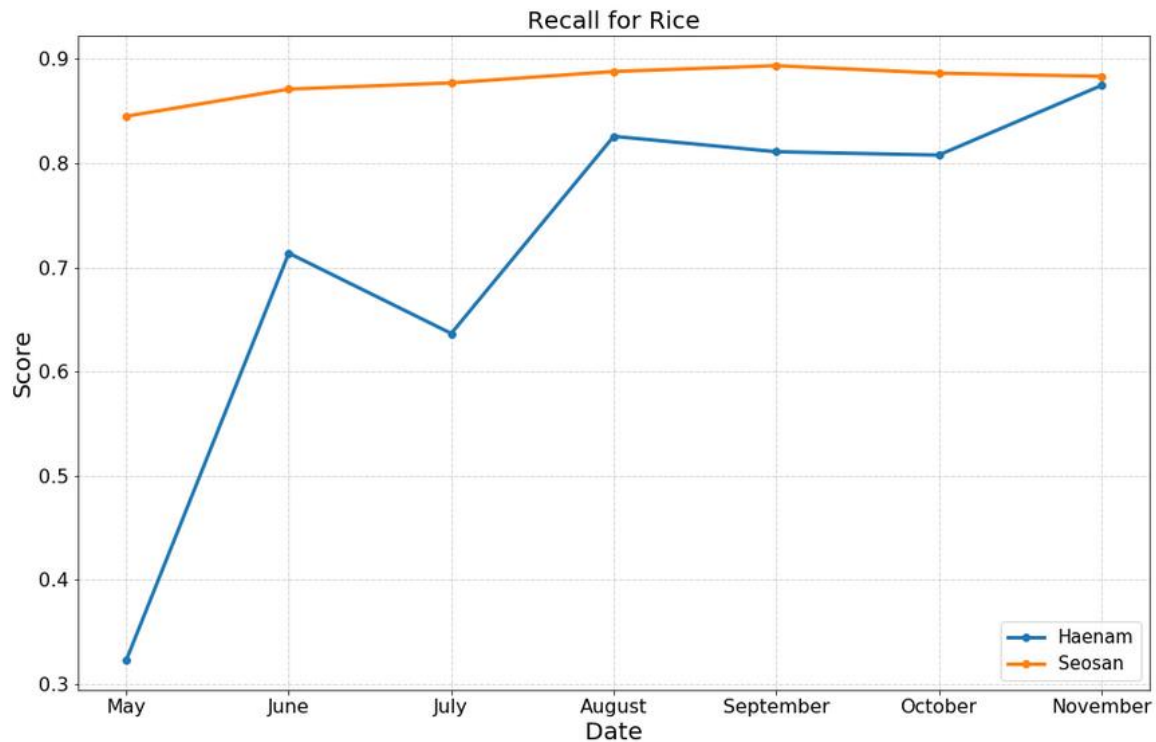
Figure 34: Recall of Rice class of predictions obtained by evolutionary Random Forest models trained on Seosan with the updated land cover map. The blue curve refers to prediction of Haenam province and the orange curve to those of Seosan province.

Scenario 3

Following the example of Scenario 2, this scenario will fill the gap of unavailable updated land cover maps of other South Korean provinces by using generated data from K-Means algorithm. This test scenario also involves the Random Forest rice classification approach that utilizes the K-Means Clustering methodologies to create a training dataset for the year of inspection. Train the model using the generated training data of Seosan and Dangjin and apply it on the rest of the country. At the same time, the Seosan-Dangjin trained model will be tested in other provinces against locally trained models. The generalization of the model will be thereby evaluated.

Random Forest rice paddy classification algorithm was trained based on the land cover map for the year 2018, which was created using k-means clustering method, as mentioned at KR07. Different combinations of number of trees and depth parameters were also tested in this scenario, in order to find values that lead to the most generalized model. Figure 35 and Figure 36 present precision and recall scores obtained when applied to each model both on Seosan and Haenam provinces, as well as these metrics on the clustering results. Number of trees parameter was set to 7 and the depth parameter was set to 4 to give the best results.
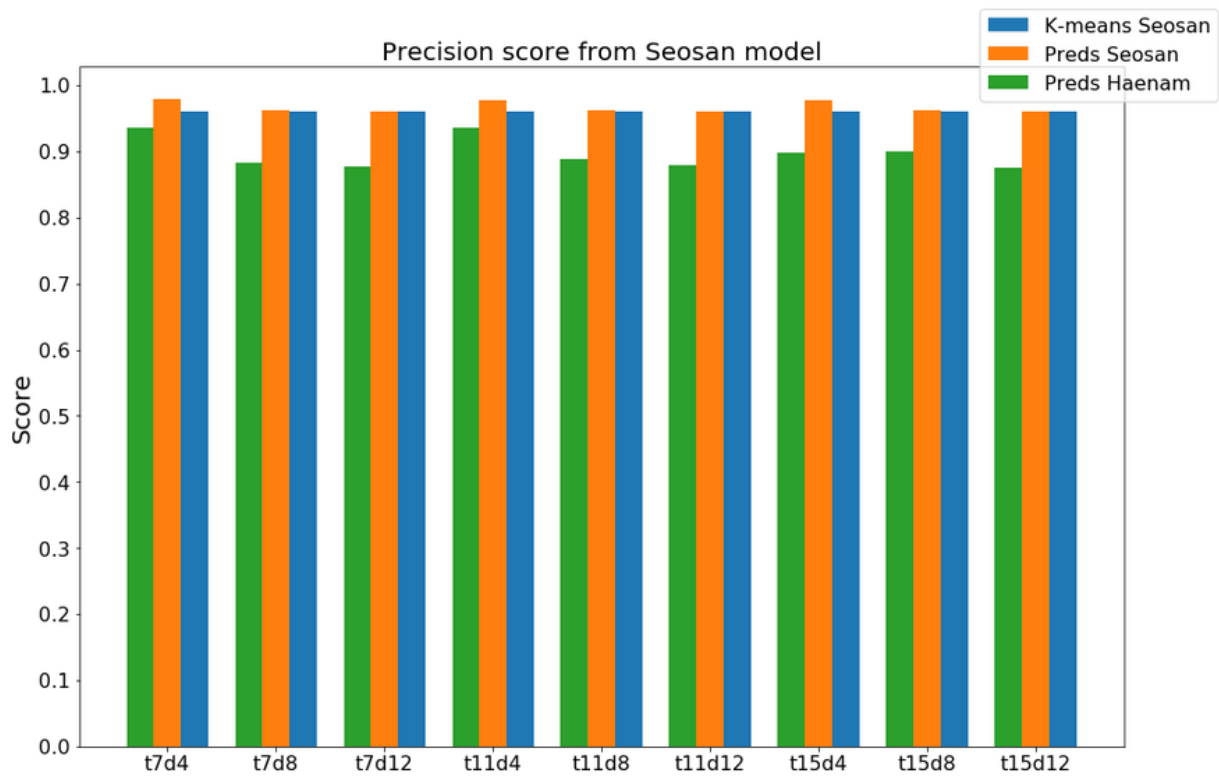
Figure 35: Precision scores obtained by predictions of random forest models trained on Seosan province using training data from k-means clustering method, both in Seosan (blue) and Haenam (orange), as well as precision score for k-means results, for different values of number of trees and depth parameters.
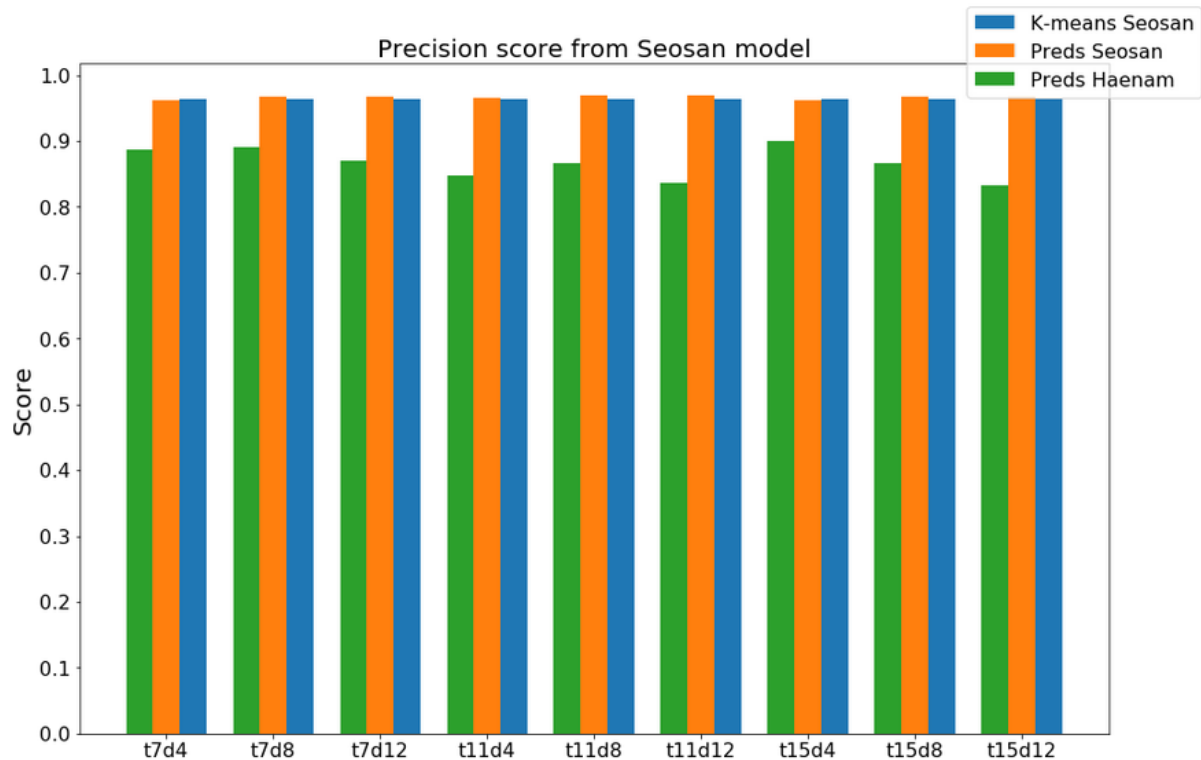
Figure 36: Recall scores obtained by predictions of random forest models trained on Seosan province using training data from k-means clustering method, both in Seosan (blue) and Haenam (orange), as well as recall score for k-means results, for different values of number of trees and depth parameters.

 The evolution of the recall score (see Figure 37 below) follows an uprising curve for both districts, reaching an almost perfect accuracy. However, while precision of predictions of Seosan is also increased through time, the precision of predictions of Haenam keeps decreasing, but even so, it always achieves scores over 87%. Since, the optimal parameterization is done based only on the Seosan city, this drop of precision score can

be   justified.
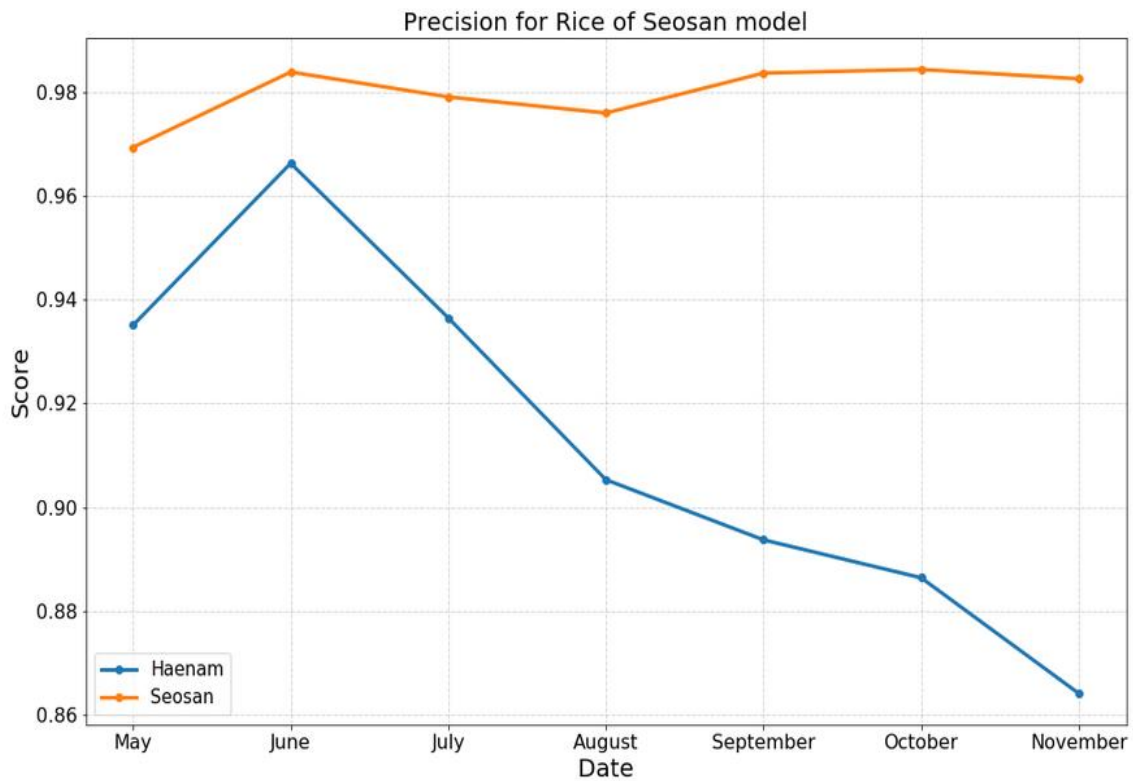

Figure 37: Precision of Rice class of predictions obtained by evolutionary Random Forest models trained on Seosan with the land cover map produced by the k-means clustering. The blue curve refers to prediction of Haenam province and the orange curve to those of Seosan province.
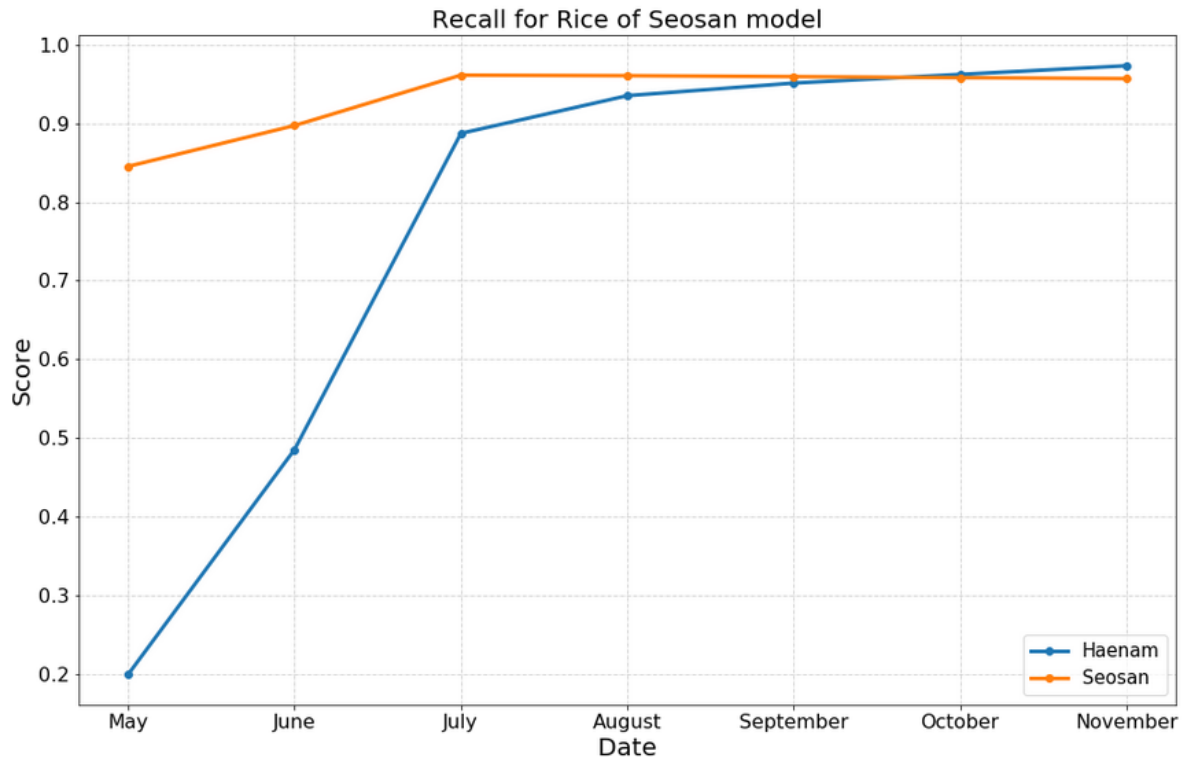
Figure 38: Recall of Rice class of predictions obtained by evolutionary Random Forest models trained on Seosan with the land cover map produced by the k-means clustering. The blue curve refers to prediction of Haenam province and the orange curve to those of Seosan province.

Finally, another Random Forest model is trained on Haenam province based on a land cover map created by the clustering method. Figure 39 and Figure 40 present a comparison between the predictions of this model, with the predictions of the model trained on Seosan. The first model performs a little better in terms of recall score, while the second in terms of precision score. The model of Haenam does not outperform the one of Seosan, but on the contrary, both of them achieve very high accuracies. Thus, and based on all the results, we can conclude that the Seosan model generalizes the problem of rice classification in South Korea, and can be applied successfully on other regions of the country.
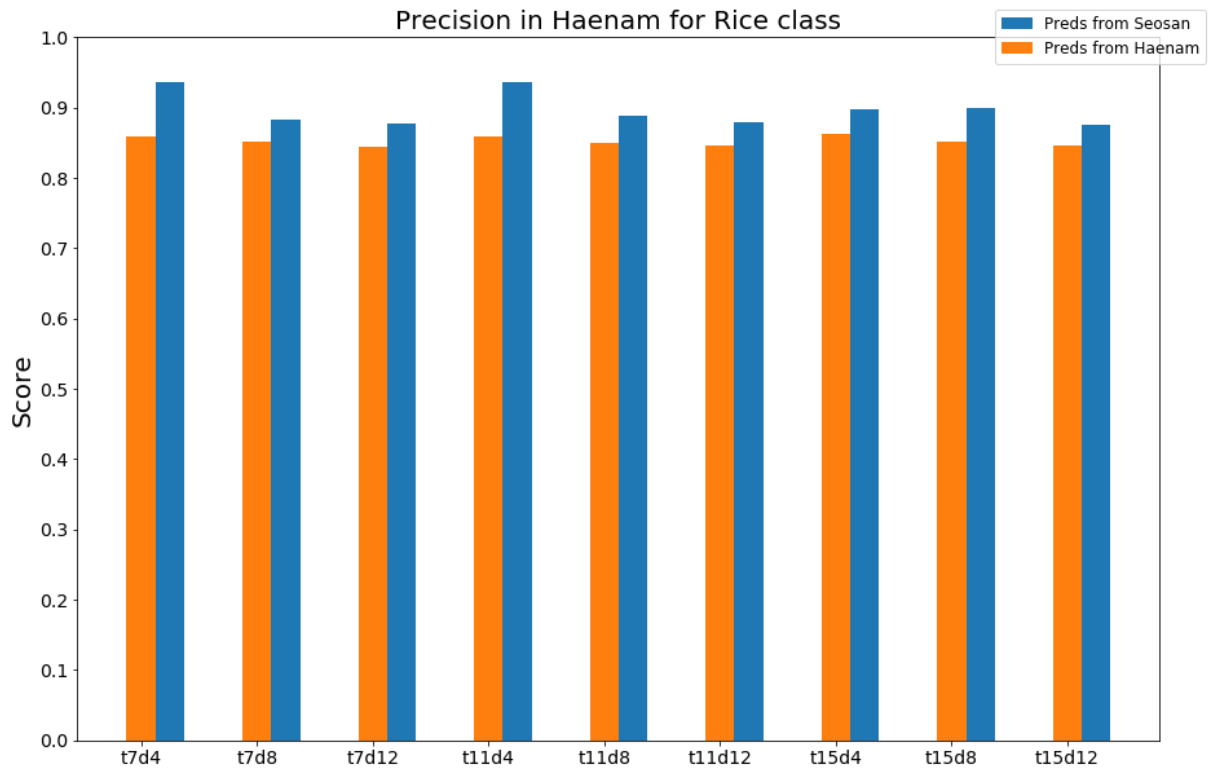
Figure 39: Precision scores obtained by predictions of random forest models trained on Haenam province, both in Seosan (blue) and Haenam (orange), for different values of number of trees and depth parameters.
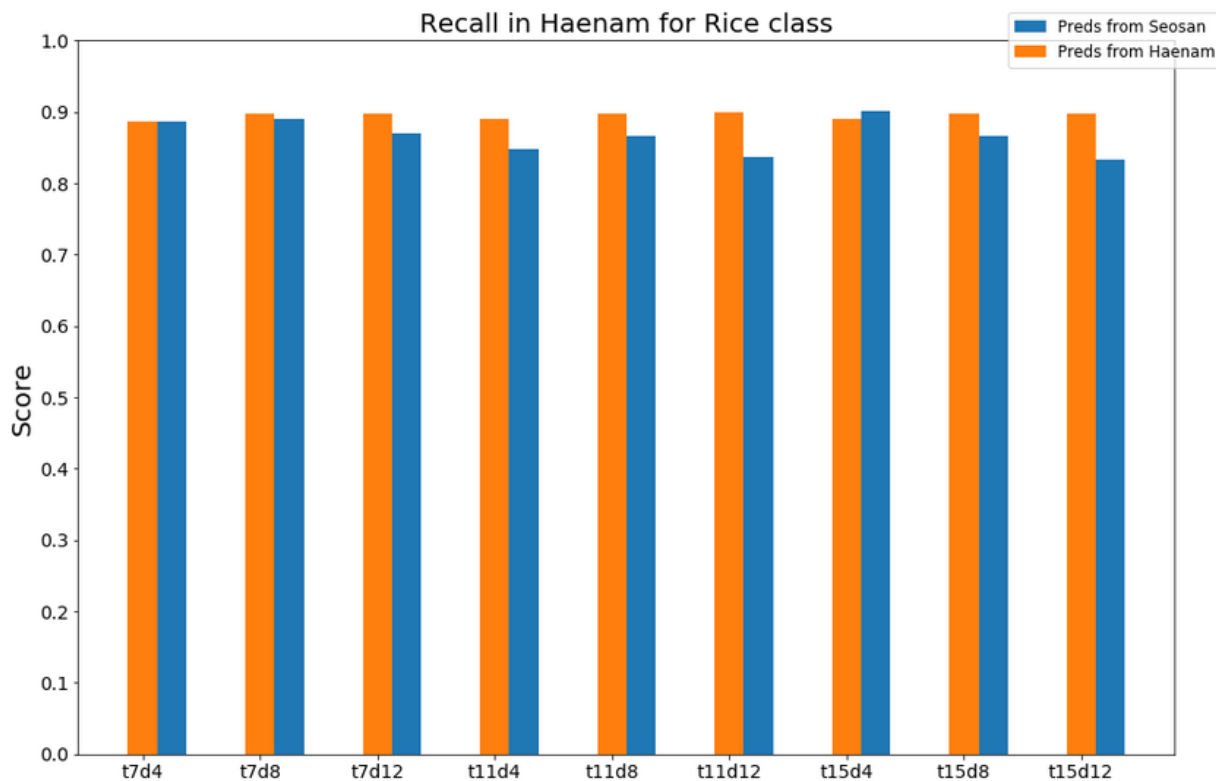
Figure 40: Recall scores obtained by predictions of random forest models trained on Haenam province, both in Seosan (blue) and Haenam (orange), for different values of number of trees and depth parameters.

Scenario 4

Following the rice classification of Scenario 2 and 3, in this scenario will investigate how the rice status indicator, which was thoroughly described in the methodology section, can extract valuable information regarding the biomass and the yield of the rice crop. The primary identification of start, peak and end day of year at the NDVI time-series, apart from indicating crucial crop cultivation phases, reveal the behavior of the crop over time, which can be quantified from the calculation of integrals, in order to provide biophysical meanings to the NDVI curve(Boschetti et al., 2009).

Using the TIMESAT software the start, peak and end day of the year of the crop season are defined. In addition, based on these extracted days, a biomass and a yield indicator are calculated from the NDVI time-series. The biomass indicator can be derived at the peak of the season (early in the year - timely decision making is possible), while the yield indicator at the end of the season.

The study of NDVI profile can identify the determination of the three main crop growth phases. After transplanting, which in our case can specify the start of season, the rice crop is at the late vegetative phase, as the germination at the seedbeds has proceeded, where the tillering stage begins until the stem elongation of the crops. Subsequently, the reproductive phase follows, where the peak of season is noted with the highest value of the NDVI time-series, when the crop is at the heading stage. As regards the end of season

of the rice crop status indicator, is referred to the day of year where the crop is mature enough to be harvested and as a result is at the ripening phase.

The study of the NDVI time-series was implemented at the object level, by creating objects of 7x7 pixels. The table that follows, contains figures illustrating the development of rice crop status in parallel with the crop phase. The colour pallet illustrates the differences of NDVI values with the peak (highest) NDVI value for every object.

Table 35: Development of crop status and crop phase

	Start Crop Status Indicator: **Vegetative Phase**

	Peak Crop Status Indicator: **Reproductive Phase**
	End Crop Status Indicator: **Ripening Phase**
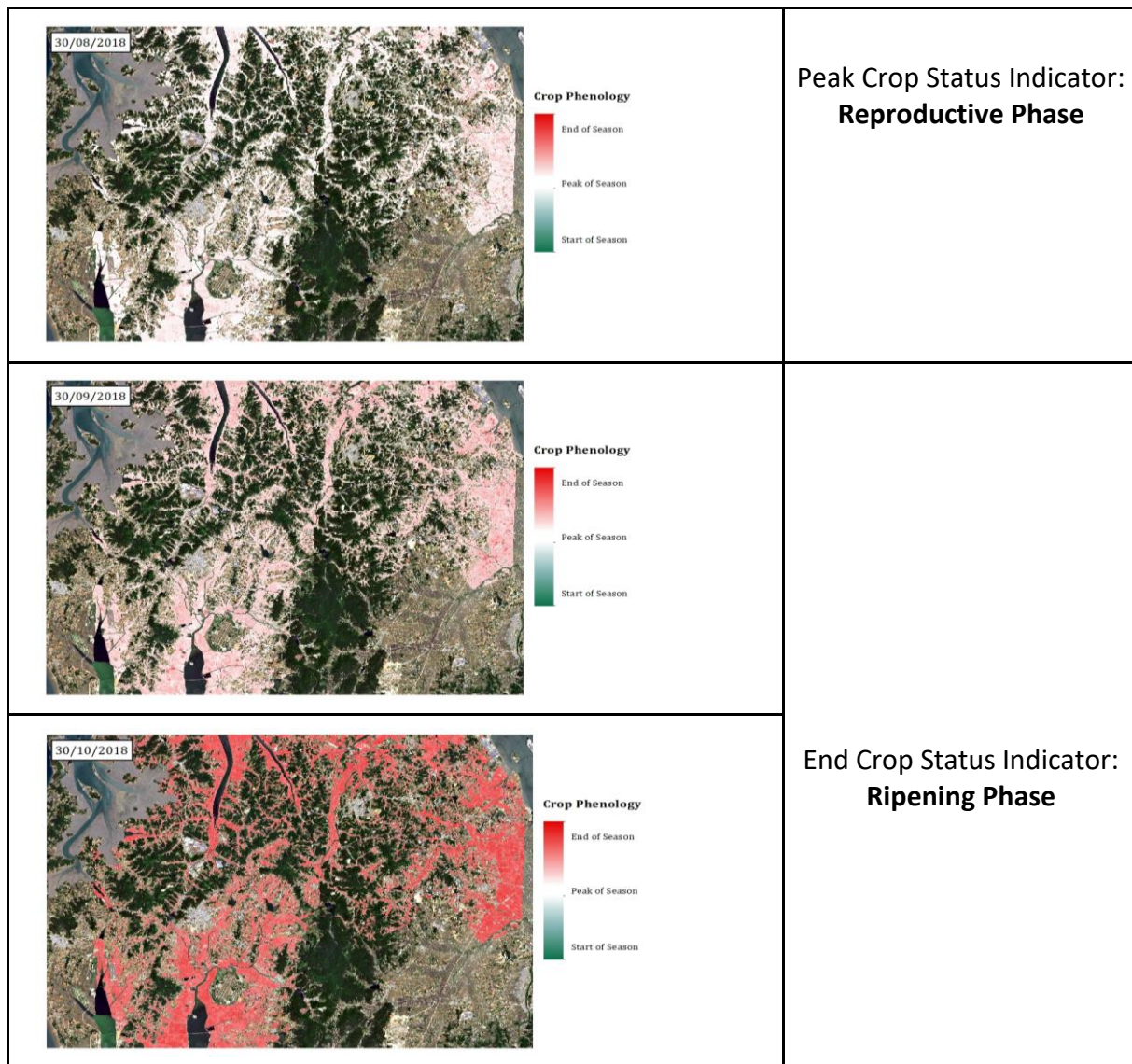
The above figures reveal the development of the rice crop, highlighting with green color the start of season, with white color when passing to the peak of season and with red color when the crop reaches the end of season. The different colors at the above figures, point out that the rice crops can be at different growth phases at a specific day of year. As mentioned before two essential proxies are obtained from the calculation of start, peak and end of season, which are a biomass and a yield indicator. These two are delivered from the calculation of the left and right integral, between start-peak and peak-end with the base level respectively. A high value of the biomass indicator, pinpoints a high accumulated biomass until the peak of season, which can be used as an early indicator of the final production.
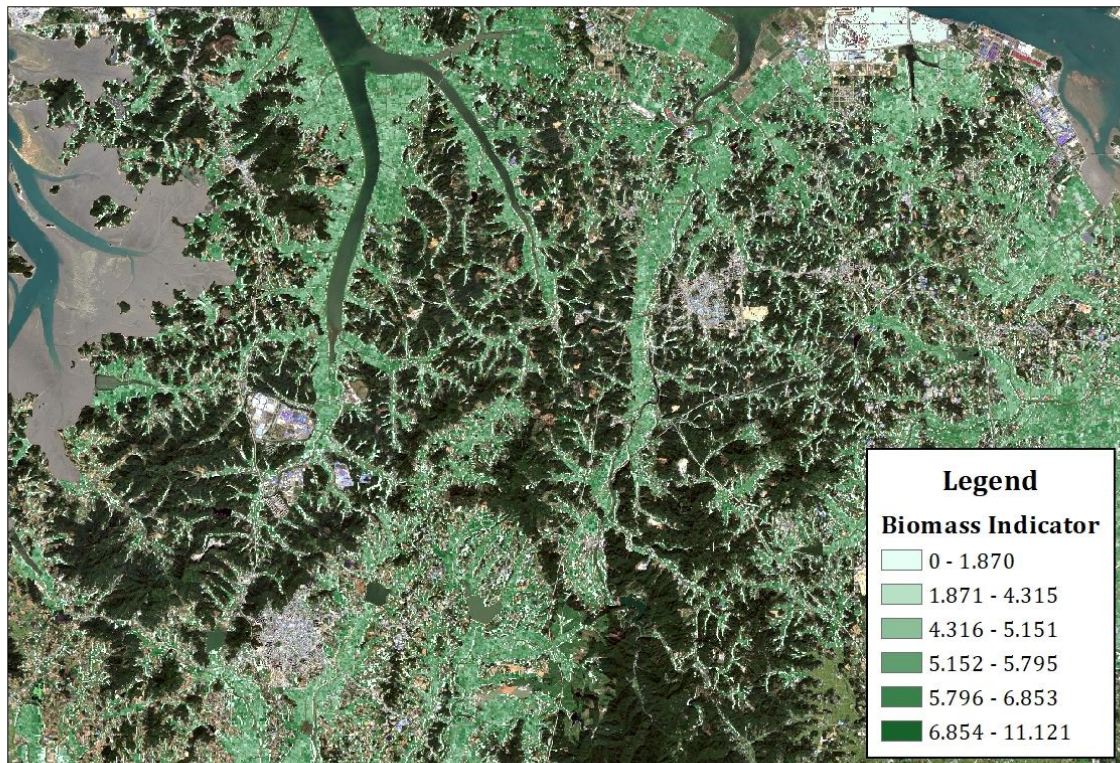
Figure 41: The Biomass Indicator results for the district of Dangjin.

Moreover, at the end of season, the yield indicator is calculated, as a proxy of the final production. The figure of the yield indicator, illustrates with deep red color the areas with expectedly high final production. This indicator can be utilized only as a proxy, since it is derived exclusively from the study of the NDVI curve, without considering any other parameters.
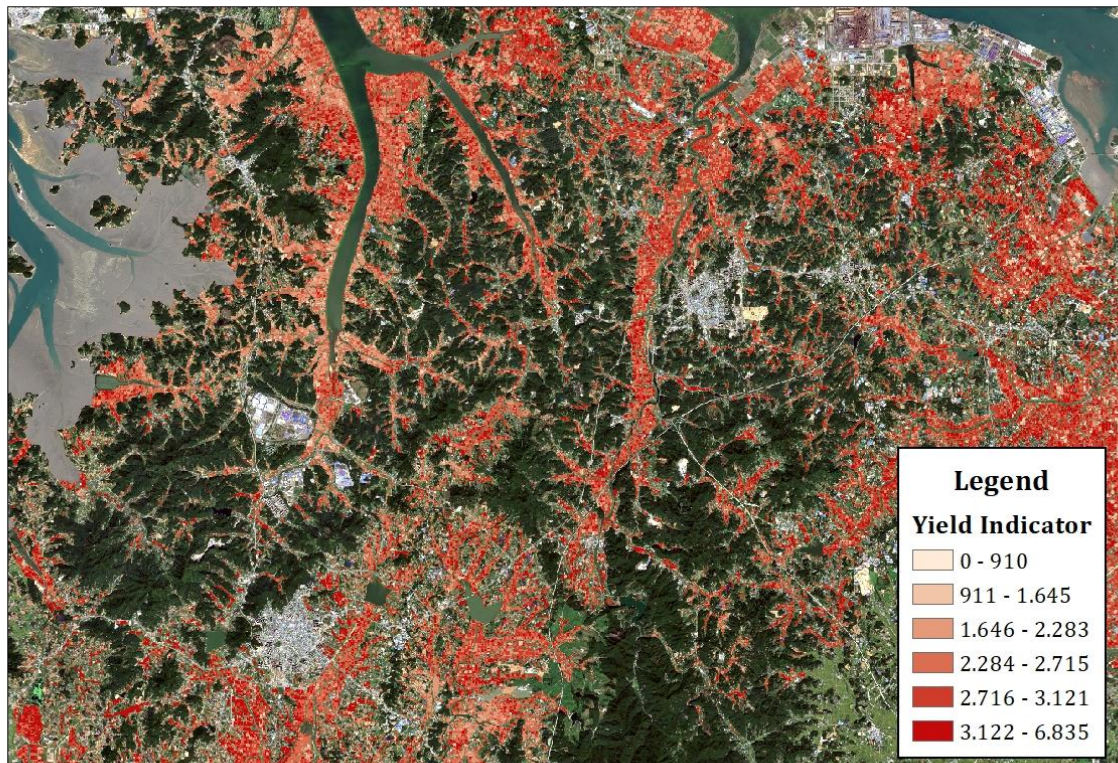
Figure 42: The Yield Indicator results for the district of Dangjin.

The values of the biomass and yield indicators, as shown above, should be considered as quality values, instead of quantity, since they derived from the calculation of integrals and cannot be used for quantifying with accurate numbers the accumulated biomass and the final production.

Validation of the rice status indicators was a challenging part. The absence of validation crop-calendars for the rice crop at the area of interest, led to the investigation of other means of evaluation. The analysis of the related bibliography provided a benchmark for the primary results of the crop indicators, and that is SOS and EOS. As mentioned Park et al.(2014), based on the heading period or the growing days, the rice (Oryza sativa L.), which is cultivated widely in South Korea, can be grouped into early maturing, medium maturing and medium-late-maturing cultivars, which result in different transplanting periods, since the transplanting period is determined by the relative growth difference of rice plants at a certain time. Besides, Gutierrez et al.(2013) highlight that late-maturing plots are flooded twenty four days more than the early maturing. Specifically, as mentioned Jang et al. (2012), at the district of Dangjin, where mainly early and medium-late maturing rice crops are cultivated, transplanting starts at the middle of May until the early of June for the medium-late maturing and middle until end of June for the early maturing cultivars. As regards the harvesting procedure, as Kim et al. (2015) underline, occurs in October for the medium-late maturing and starts in middle of September for the early maturing cultivars .
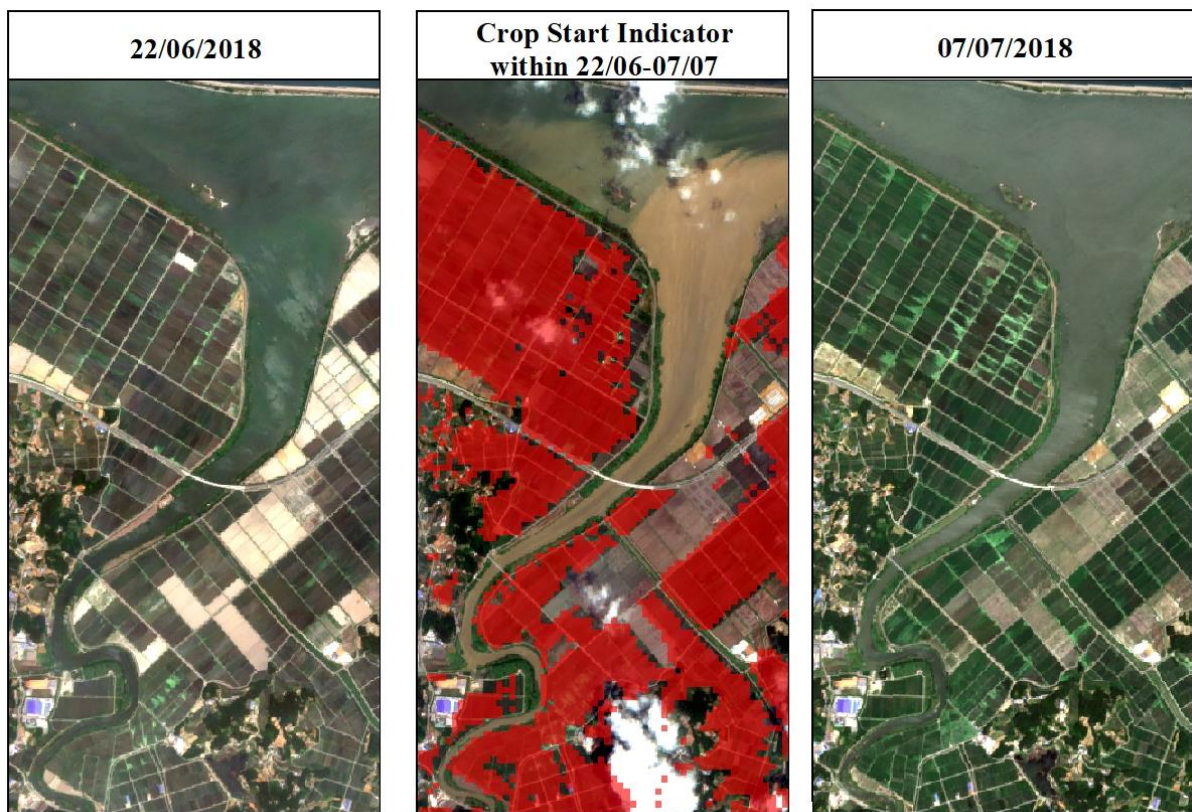
Figure 43: Validation of Crop Start Indicator results through the available cloud free imagery.

Moreover, rice status maps are created regarding the start and end of crop development, which will be also presented at the User Portal. The validation of the resulting rice status maps was achieved through photointerpretation at the before mentioned time periods of transplanting and harvesting. Furthermore, the RSI methodology returns a specific Day of Year (DOY), indicating the start and the end of season and in order to validate the output results the pre and the post available acquired images were examined. The results succeed in an acceptable accuracy through photointerpretation of the previous and the following cloud free acquisition and the resulted DOYs are inside the referenced time periods of transplanting and harvesting processes.
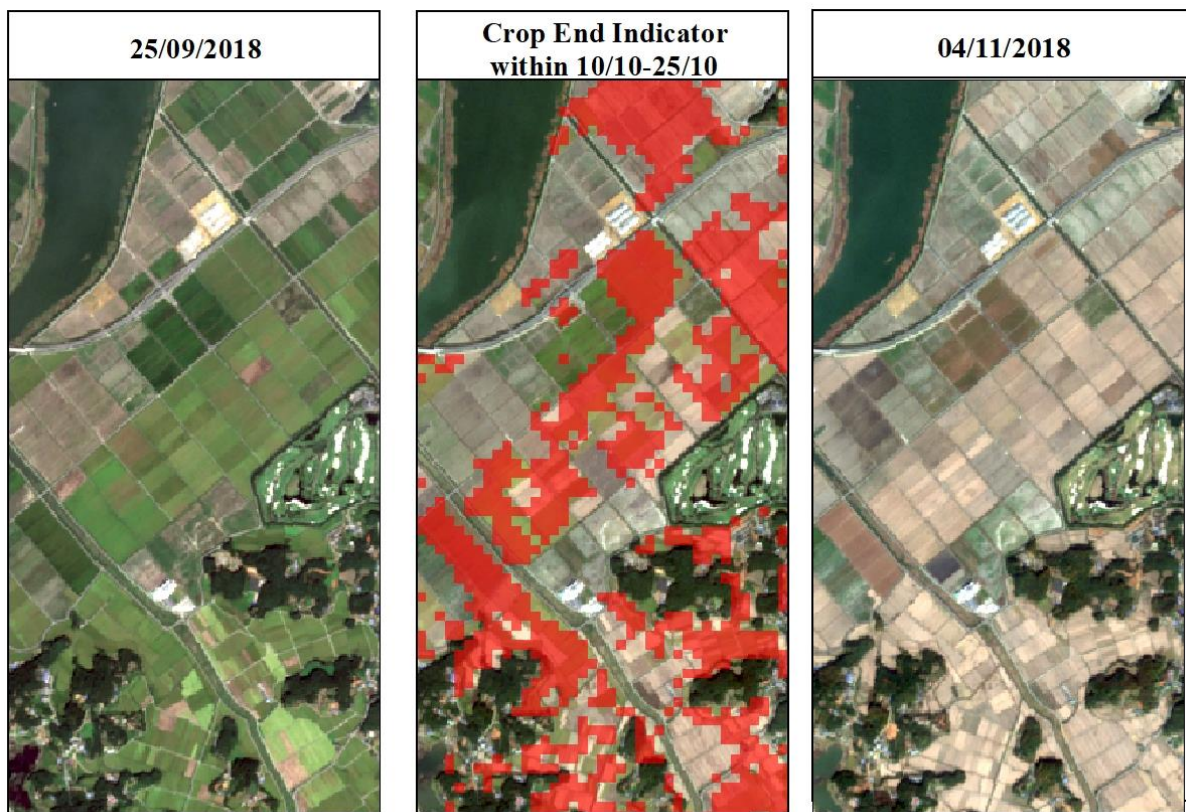
Figure 44: Validation of Crop End Indicator results through the available cloud free imagery.

6.3.7 Overall Assessment, Lessons Learnt and Future Plans

PUC2 can be thought as a successful use case, addressing the majority of the call requirements, using a plethora of the generic EOPEN scientific and technological KRs and addressing effectively the majority of the collected user requirements.

More specifically, PUC2 addresses Call Requirements I, by developing value adding services (i.e. land cover map update, clustering for pseudo-labelling, distributed rice classification, rice growth monitoring). The majority of the aforementioned pipelines have been deployed on the interoperable EOPEN platform, in the form of modular processors that can be reused and parameterized by users. Call Requirements X is also addressed, by effectively exploiting Copernicus data to tackle food security issues, targeting non-traditional users, such as food companies (CJ), economic institutes (KREI), ecological planning and policy researchers.

PUC2 has effectively used a number of scientific and technological KRs to address its user requirements, showcasing the value of the EOPEN platform, in its interoperability, modularity and generality. Specifically, PUC2 has used KR01 to update past land cover maps and KR05 to perform unsupervised rice classification to create training datasets. Additionally, KR02 is utilized to identify events related to food and agriculture in South Korea. KR10 is used to automatically search and download the Sentinel data time-series in an efficient and coherent manner. KR13 is also used to access the korean databases of meteo data. Maps of precipitation, temperature and solar radiations will be available in

gridded images on the EOPEN User Portal. KR12 continuously harvests relevant food security tweets in both Korean and English, which are annotated as relevant or not by KU. KR14 is exploited to allow the user to visualize the PUC2 outputs (i.e timeseries of rice status indicators, evolution of rice classification product, meteo datasets etc.). KR15 is extensively used, with PUC2 having effectively utilized the HPDA module to perform distributed clustering and supervised classification for the large-scale application of the outputs. Finally, KR17 - the EOPEN platform, has fostered the integration of the various PUC2 pipelines in an interoperable platform that will allow the user to reuse the processors and alleviate any infrastructure and deployment challenges.

PUC2 has developed solutions targeting South Korean stakeholders. The distance, different time zones and language barriers have brought up certain challenges with respect to the active engagement of those stakeholders, in providing timely feedback and evaluation for the PUC2 outputs. Nevertheless, over the next months PUC2 partners will intensify the stakeholder engagement activities, seeking 3rd party evaluation of the use case outputs. In that respect, the list of stakeholders (see table above) has been revised and pertinent communications with all of them have been started.

In conclusion, PUC2 has demonstrated the value and applicability of the developed products and services through a series of test scenarios that have been thoroughly reported in the previous subsection.

6.4 PUC 3 – Monitoring Climate Change through Earth Observation

6.4.1 Use Case story

In the present and future climate change environment, the average temperature in Finland is rising / will rise more (2°C by 2040), and precipitation will increase faster (5– 10% by 2040) than the global average. The changes are affecting winters more than summers with the largest changes in the northern part of the country: Finnish Lapland. Since the early 2000s, Finland has taken a proactive role in managing the Climate Change situation nationally, with mitigation and adaptation plans. The activities of the Transportation sector runs deeply through the sectors: sustainable industry, land use planning and construction, tourism and recreation as summers become warmer, wetter and longer and snow packed regions shift northward. Our use case begins with historical snow and temperature data, supplemented by EO data, which support Finnish Transportation Infrastructure Agency (FTIA)'s current and future road maintenance for the Finnish drivers and riders. Our Use Case continues with temperature and snow data support for the Finnish Lapland communities who are, and will be, experiencing the greatest climate change consequences. The reindeer herders are the frontline people experiencing Climate Change.

Table 36: PUC 3 main stakeholders

Stakeholder	Description
Reindeer Herders	Small-scale livelihood in mostly Northern Finland
Reindeer Researchers	Researchers in the University of Lapland

Finnish Transportation Infrastructure Agency (FTIA)	Authority for transportation and road maintenance in Finland

The PUC3 Stakeholders provide the on-the-ground view and the national view for mitigating Climate Change in Finland.

Reindeer Herders: Herding is small-scale livelihood in Northern Finland with about 5000 herders. Directly exposed to changes in land use and climate and have an intuitive, high fidelity knowledge of climate conditions to support reindeer care. Reindeer activities drive tourism in Northern Finland → 90% of tourists are from foreign countries.

Reindeer Researchers: Due to reindeer herding's large Finnish social, cultural and economic importance → multiple research institutes provide financial support to studies. The Reindeer Researchers study herding's economic development and sustainability, the herders' adaptation to / mitigation of Climate Change, and the herders' effects on land-use and environment.

Finnish Transportation Infrastructure Agency (FTIA): FTIA is responsible for transportation and road maintenance in Finland. Finland has a total of 78 000 km of public roads, many in sparsely populated regions. Transportation in general accounts for 20% of all greenhouse gas emissions in Finland, out of which road transport in particular accounts for 90%. Climate change is expected to impact all facets of the transport system: the infrastructure, modes of transport and operations and, especially, the maintenance. The FTIA must make decisions based on yearly trends in weather and climate.

6.4.2 Timeline of PUC 3 implementation

Table 36: PUC3 implementation timeline

Milestone	Completion month	Status	Comments
Questionnaires	M8	Done	
Requirements gathering	M8	Done	Both user and data requirements are defined. These may be refined during the project based on the feedback of the end users.
Phone interviews with stakeholder representatives	M10, M18	Done	FMI conducted phone interviews in order to refine the user requirements
Road passability product	M33	In progress	Part of KR01
Integration of change detection algorithm	M33	In progress	Part of KR01
Data gathering	M21	Done	

Data integration with platform	M29	In progress	
Prototype demonstration	M24	Done	Held in South-Korea.
User workshop	M30	To be done	Event will be a webinar.

6.4.3 **Monitoring of PUC 3 implementation against the User Requirements**

Table 37: Monitoring of PUC3 implementation against the User and Data Requirements

Requirement ID	Description	Current status
PUC3_GC1	Should be provided with a user friendly interface.	Done
PUC3_GC2	Should be provided with easy access and management of datasets.	Done
PUC3_GC3	Must be provided with data integration capabilities.	Done
PUC3_GC4	Must be enabled to overlay satellite and ground observation data over a map background.	Done
PUC3_GC5	Should be provided with the capability to select a user-defined area of interest.	Done
PUC3_GC6	Should be enabled to visualize real time situation (most recent data).	Done
PUC3_GC7	Must be enabled to browse historical observations.	Done
PUC3_GC8	Should be provided with tools for statistical analyses of selected area, time period and dataset(s).	In progress
PUC3_GC9	Should be provided with tools for visualizing time series of datasets and statistical analyses.	In progress
PUC3_GC10	Should be provided with time series trend analysis and future projections capabilities.	In progress
PUC3_GC11	Must be enabled to compare multi-platform data (e.g. satellite vs satellite, satellite vs. ground observation).	Done
PUC3_GC12	Should be provided with access to relevant tweets with specific key words through the platform.	Done

Dataset ID	Dataset Name	Current status	Description
PUC3_DC1	GlobSnow SWE	Done	Gridded daily, weekly or monthly data for snow cover with 25 km resolution

			starting at 1995. Based on satellite and ground weather station data.
PUC3_DC2	SMOS Level 3 Freeze/Thaw	Done	Gridded daily soil state (Thaw/Frozen/Partially frozen) with 25 km resolution starting at 2010. Based on satellite and ground weather station data.
	Sentinel LST Product	To be done	Land Surface Temperature data from Sentinel satellites.
PUC3_DC3	Air Temperature FMI ClimGrid	Done	Grid-interpolated daily averages of air temperature with 10 km resolution starting at 1961 until 2018. Based on ground station data. Pre-computed maps (no API).
PUC3_DC4	Snow depth FMI ClimGrid	Done	Grid-interpolated daily averages of snow cover maps with 10 km resolution starting at 1961 until 2018. Based on ground station data. Pre-computed maps (no API).
PUC3_DC5	FMI Open Data, Climatological values (Climatological data for meteo observations (ref. D1.4)	In progress	Long-term (30-year) statistics of meteorological variables
PUC3_DC6	Social media data (twitter)	Done	Annotated set of tweets of snow. FMI provides annotations, CERTH handles the rest.
PUC3_DC7	FMI Open Data, Climate model output	In progress	Estimated changes in climatological values over decades, based on multiple climate models and future climate scenarios.
PUC3_DC8	FMI Open Data, Weather observations	In progress	Instantaneous weather observations and daily, monthly, seasonal, and yearly values.
PUC3_DC9	FMI Open Data, Weather model output (HIRLAM wather	Done	Numerical weather prediction forecasts and model reanalysis, e.g. ERA5 from Copernicus Climate Change Service

	forecast ref. D1.4)		
PUC3_DC10	Regions of Finland	Done	Region and municipality area borders in Finland. Pre-computed maps.
PUC3_DC11	Reindeer herding areas	Done	Herding area borders in Lapland. Pre-computed map.
PUC3_DC12	FTIA Road maintenance classes	Done	Road maintenance classification in Finland for snow removal and sanding.

User requirements PUC3_GC1 – PUC3_GC7 and PUC3_GC9 are user interface (UI) requirements, and thus are provided by the user interface capabilities. FMI provides the Use Case with the meteorological and climatological data (datasets PUC3_DC1 – PUC3_DC5, PUC3_DC7 – PUC3_DC9), which is offered through the EOPEN UI. The UI implementation is currently in progress by SPACEAPPS.

Of these requirements, PUC3_GC5 is given special consideration. In addition to giving the user an option to select their own area of interest, EOPEN will also provide predefined areas. The planned areas are Finnish region and municipality borders (PUC3_DC10), and reindeer herding districts (PUC3_DC11). Moreover, our stakeholders have indicated that road maintenance classification dataset would be useful, and so we plan to offer the data in EOPEN (PUC3_DC12).

Lastly, FMI provides the human annotations for Finnish tweets about snow coverage (see section Social media crawlers). The annotation work is required in order to satisfy user requirement PUC3_GC12 and to generate the dataset PUC3_DC6.

6.4.4 Implementation

Examples of the visualizations for the GlobSnow and SMOS Level 3 Freeze/Thaw products in Finland on the 27th of January 2019 are shown in Figure 45. The combined snow depth (left panel) and ground condition (right panel) information is crucial for reindeer herders for estimating the winter feeding requirements and potential grazing areas for the reindeer.
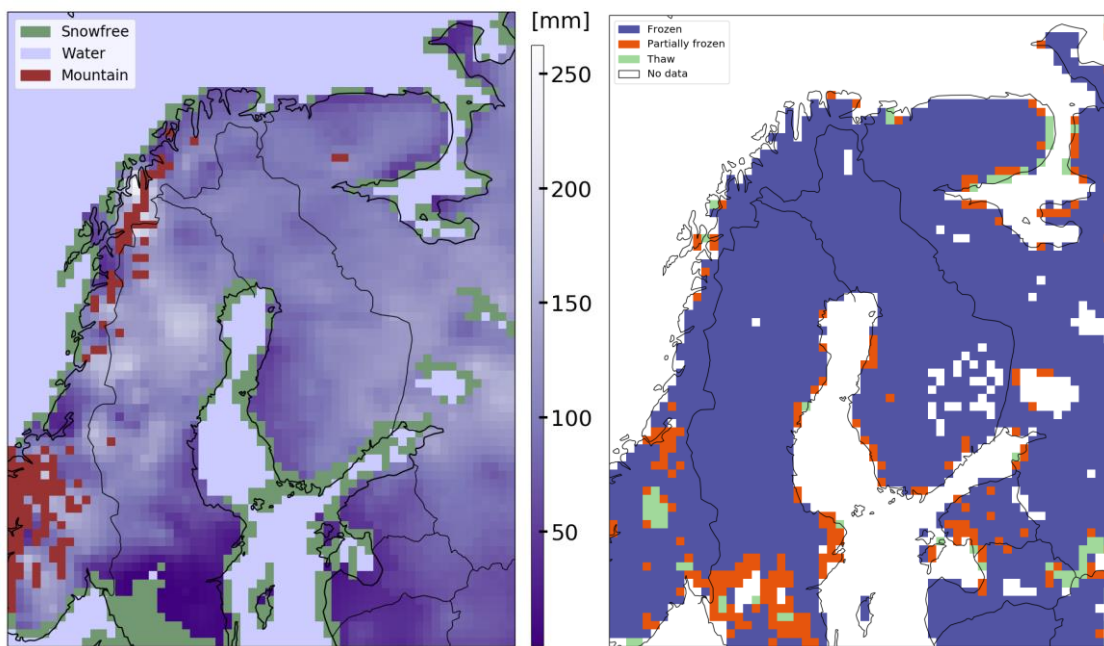
Figure 45: GlobSnow and SMOS Level 3 Freeze/Thaw products over Finland on the 27th of January 2019.

The same as in Figure 45, but for the late spring (13th of April 2019) is shown in Figure 46. In this example, some of the areas in Lapland are already melted (Thaw, right panel), but still have some snow on top of them (left panel). This allows the reindeer better access to food on ground.
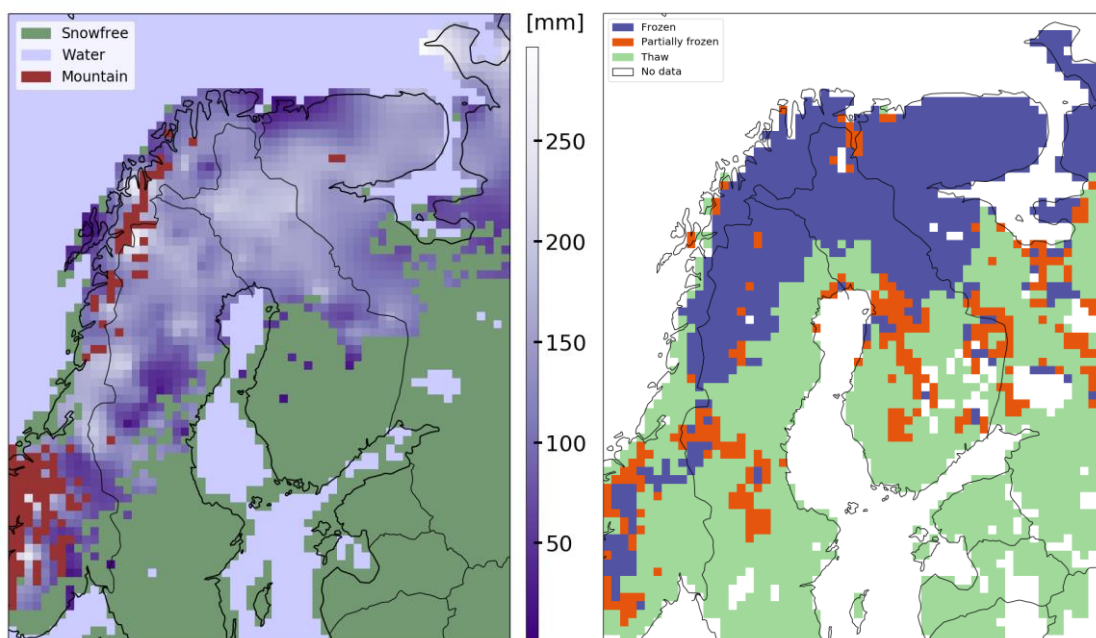

Figure 46: Same as in Fig. 45, but for the 13th of April 2019.

The ClimGrid data shows the snow depth and air temperature using ground-based instruments. Example of the data and tool for displaying statistical properties of it is

shown in Figure 47. In the example, a specific reindeer herding area have been selected and the tool shows statistical information over the area.

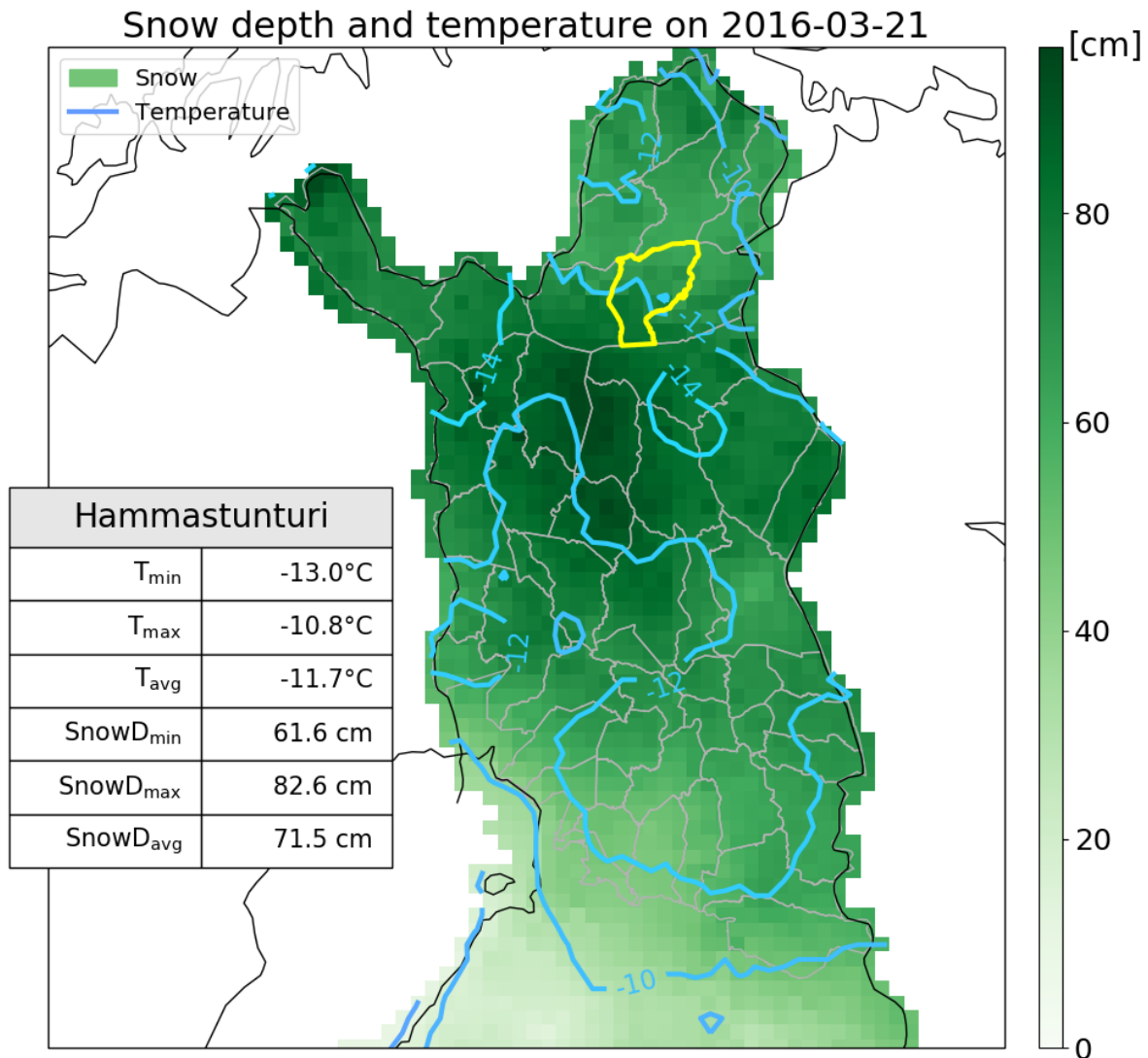| Hammastunturi	
T_{min}	-13.0°C
T_{max}	-10.8°C
T_{avg}	-11.7°C
$SnowD_{min}$	61.6 cm
$SnowD_{max}$	82.6 cm
$SnowD_{avg}$	71.5 cm

Figure 47: ClimGrid data (snow depth and air temperature) over Lapland on 21st March 2016. Separate window displaying statistical information from a reindeer herding area.

The HIRLAM daily rainfall accumulation forecast over Lapland is shown in Figure 48. Also, the road maintenance classification form FTIA is also shown for the road network. This information can be used by the FTIA to estimate their daily road maintenance activities and provide priorities for remote parts of Lapland.
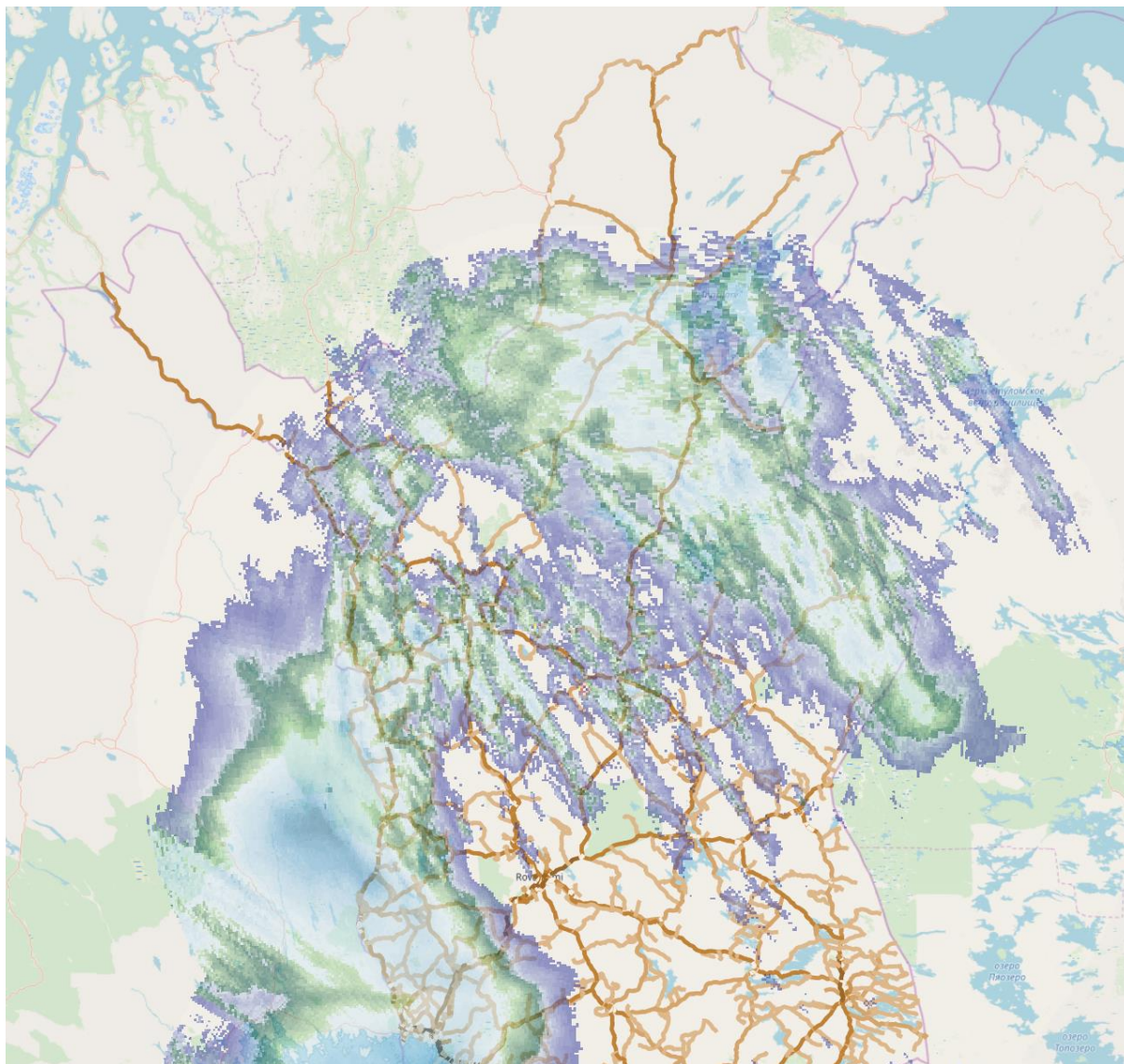
Figure 48: HIRLAM 24h rain accumulation forecast over Lapland. The road maintenance classification from FTIA is also shown.
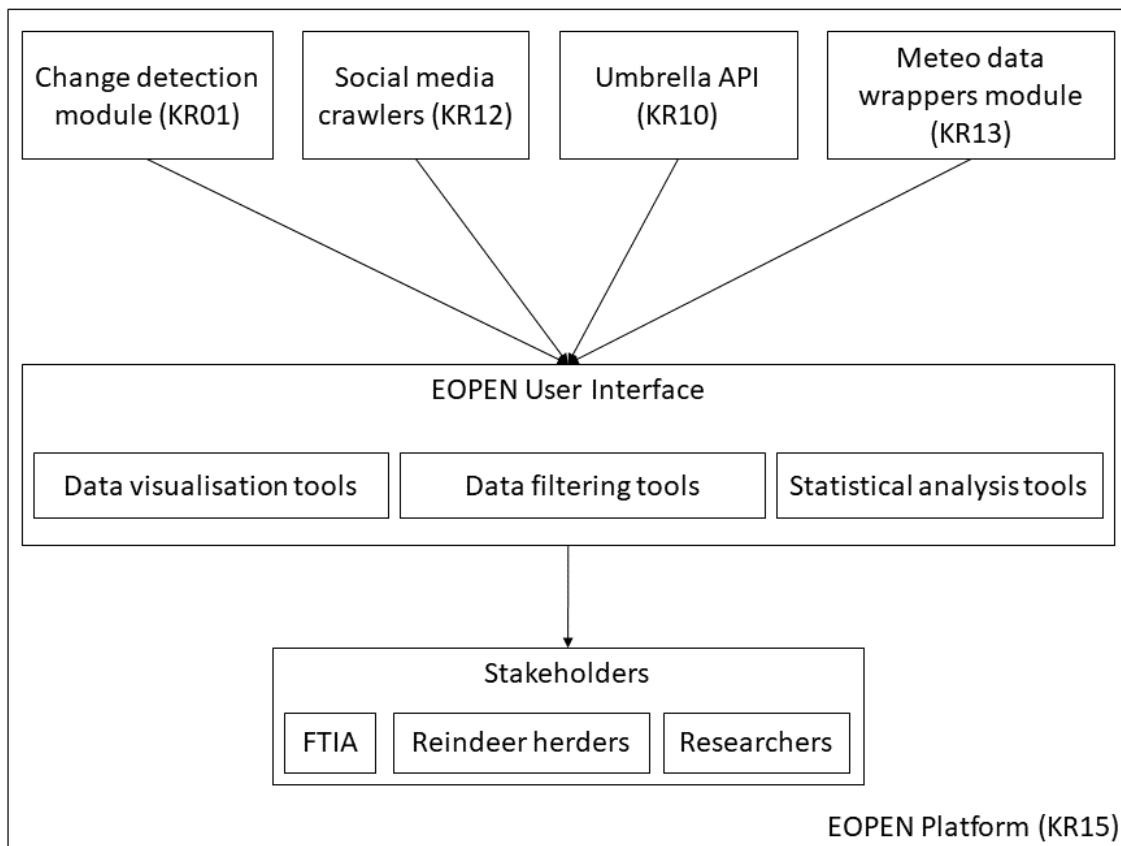
6.4.5 **PUC 3 Implementation workflow**

Figure 49: PUC3 Implementation workflow

6.4.6 **Test scenarios**

Scenario 1

Description: The winter 1972-1973 was exceptionally bad for reindeer herders due to 'deep icy snow' and 'deep snow' as reported by them (Vuojala-Magga et al., 2011). Due to these difficult snow conditions, more herding and additional feeding were required. Visualise the trends for snow depth over past winters.
KPI: quantitative difference in snow depth, length of time to get the results
Data: PUC3_DC3, PUC3_DC4, PUC3_DC11
Tools: Statistical measures, data filters
Criteria: User is able to compare snow depth maps between the good and the bad winters for particular reindeer herding areas in Lapland, and potentially find a statistically significant difference in snow depth.

Scenario 2

Description: A recent paper (Anttila et al., 2018) showed that in the European Arctic zone the long-term trends for the onset of melting snow is getting earlier and the melt season is getting longer. Visualize the same trend in EOPEN. Compare the trends between satellite-based and ground-based data.

KPI: qualitative agreement with the trends, number of data sources/tools used, length of time to get the results

Data: PUC3_DC1, PUC3_DC3, PUC2_DC4, PUC3_DC5

Tools: Statistical and time series analysis tools, data filters

Criteria: User is able to show the long-term trend over time for the beginning of snow melting and its length in Northern Finland.

Scenario 3

Description: Climate change will increase flooding and cause permafrost to thaw in localized areas. FTIA wants to minimize gaps in road maintenance at low-priority regions and to improve their road maintenance planning. Visualize the trend of thawing on remote roads in Northern Finland

KPI: quantitative difference in thawing trends, number of data sources/tools used, length of time to get the results

Data: PUC3_DC2, PUC3_DC12

Tools: Statistical and time series analysis tools, data filters, map filters

Criteria: User is able to show the trend in thawing over the winters in remote road areas in Northern Finland.

As shown in Table *37*, most of the data products have been integrated in the platform. The FMI Open Data API is still being implemented and will be ready between M28-M29. The data products PUC3_DC3, PUC3_DC4, PUC_DC6, and PUC3_DC9-PUC3_DC12 have been tested on the platform and will require only minor, mainly visual improvements. The other products are not yet implemented.

The test scenarios will be implemented once the tools for statistical and time-series analysis are ready and all the data products have been tested on the platform.

6.4.7 Overall Assessment, Lessons Learnt and Future Plans

This use case focuses on non-traditional user communities (reindeer herders, reindeer researchers and traffic authorities) and provides access to data sets that were not previously available on the same platform. This enables the users to not only compare multiple data sources together, but also alleviates the need for computer resources (High level Call Requirement I). The statistical and time-series tools that will be implemented helps the users to quickly estimate the importance of detected events both locally and globally (High level call requirement IV).

As FMI is responsible for the meteorological data acquisition (Technological KR13), it will be utilized also in the use case and demonstrated in the test scenarios. Weather-related data interface is crucial for the use case as it provides easy access and can be shown in a user-friendly way on the platform. The users are also notified about snow-related tweets on the platform by utilizing the social media crawlers (Technological KR12). This will be enhanced further by the event detection module (Scientific KR02), which can alert the users for possible snow-related events. Finally, the change detection module (Scientific KR01) provides the road passability product that can be utilized by the traffic authorities in Finland together with PUC3_DC12.

The status of the user requirements that have been implemented on the platform is shown in Table *37*. Currently, the platform provides an easy-to-use interface for the implemented data sets, allows multiple data sets to be displayed simultaneously, and lets

the user to select arbitrary areas in the map (PUC3_GC1-PUC3_GC5, and PUC3_GC11). The HIRLAM NWP data (PUC3_DC9) is also implemented and shows updated forecasts (PUC3_GC6). It also allows users to select past time periods if historical data is present (PUC3_GC7), such as the ClimGrid snow-depth and air temperature maps (PUC3_DC3 and PUC3_DC4). It shows snow-related tweets in the Social media dashboard (PUC3_GC12). The requirements for the statistical and time-series tools (PUC3_GC8-PUC3_GC10) will be implemented between M28-M29. This internal evaluation of the data products and the tools will continue until the test scenarios have been implemented and verified. It should be noted that this is just internal evaluation and will not be final until feedback from stakeholders and external users.

The test scenarios will be evaluated based on the achieved KPIs, which are both qualitative and quantitative depending on the scenario, but usually in relation with the other products. The implementation of Test scenario 1 is scheduled between M28-M29. The other scenarios are scheduled after Test scenario 1 between M29-M30 as they require time-series tools and also utilize the FMI Open Data products, which are scheduled to be implemented between M28-M29.

After the internal evaluation is complete (M30), stakeholders will be contacted for their evaluation. The evaluation will start by a webinar (M30), where both the general functionalities and use-case specific data and tools are presented. Stakeholders are encouraged to experiment with the platform and provide feedback on the user-friendliness of the UI, usefulness of the tools and data, convenience of the platform in general, and performance of the system. Feedback is collected directly and through a questionnaire and email. The feedback is used to assess the fulfilment of user requirements and to improve specific parts of the platform. During the stakeholder evaluation, there is also an opportunity to utilize peers as additional users and gather feedback from them through email.

For the external evaluation, there will also be a hands-on exercise organized for several persons working in FMI (M31). The user feedback is gathered directly and through email. As the exercises will be less structured, the evaluation will be mostly qualitative. Final improvement of the use-case data and tools are scheduled between M31-M32.

7 CONCLUSIONS

D7.2 is the second deliverable iteration of T7.1 Pilot Implementation and Evaluation. The present document was an update of D7.1, in which the then implementation progress was reported and the evaluation framework for PUCs and the EOPEN platform components was established. In this second iteration, we reported all progress with respect to the implementation of individual KRs and PUCs, but also filled in the relevant evaluation and performance metrics, as have been measured thus far. Any outstanding implementation and evaluation activities have been described, along with a timeplan for their future fulfilment. All further progress will be reported in D7.4.

8 REFERENCES

Anttila K., Manninen T., Jääskeläinen E., Riihelä A., and Lahtinen P. (2018). "The Role of Climate and Land Use in the Changes in Surface Albedo Prior to Snow Melt and the Timing of Melt Season of Seasonal Snow in Northern Land Areas of 40◦ N–80◦ N during 1982–2015", Remote Sens. 10, 1619, DOI: 10.3390/rs10101619

Araya, S., Ostendorf, B., Lyle, G., & Lewis, M. (2018). CropPhenology: An R package for extracting crop phenology from time series remotely sensed vegetation index imagery. Ecological Informatics, 46, 45–56. doi: 10.1016/j.ecoinf.2018.05.006

Baevski, A., Edunov, S., Liu, Y., Zettlemoyer, L., & Auli, M. (2019). Cloze-driven Pretraining of Self-attention Networks. arXiv preprint arXiv:1903.07785.

Blondel, V., Jean-Loup, G., Renaud, L., Etienne, L., Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P10008 (12pp) doi: 10.1088/1742-5468/2008/10/P10008. ArXiv: http://arxiv.org/abs/0803.0476

Eklundh, L., & Jönsson, P. (2015). TIMESAT: A Software Package for Time-Series Processing and Assessment of Vegetation Dynamics. Remote Sensing Time Series Remote Sensing and Digital Image Processing, 141–158. doi: 10.1007/978-3-319-15967-6_7

Florian, R., Ittycheriah, A., Jing, H., & Zhang, T. (2003, May). Named entity recognition through classifier combination. In Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4 (pp. 168-171). Association for Computational Linguistics.

Gutierrez, J., Kim, S. Y., & Kim, P. J. (2013). Effect of rice cultivar on CH4 emissions and productivity in Korean paddy soil. Field Crops Research, 146, 16–24. doi: 10.1016/j.fcr.2013.03.003

Jang, M.-W., Kim, Y.-H., Park, N.-W., & Hong, S.-Y. (2012). Mapping Paddy Rice Varieties Using Multi-temporal RADARSAT SAR Images. Korean Journal of Remote Sensing, 28(6), 653–660. doi: 10.7780/kjrs.2012.28.6.5

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S. and Darrell, T., 2014. "Caffe: Convolutional architecture for fast feature embedding", In Proceedings of the 22nd ACM international conference on Multimedia (pp. 675-678). ACM.

Jönsson, P., & Eklundh, L. (2004). TIMESAT—a program for analyzing time-series of satellite sensor data. Computers & Geosciences, 30(8), 833–845. doi: 10.1016/j.cageo.2004.05.006

Kim, H.-O., & Yeom, J.-M. (2014). Effect of red-edge and texture features for object-based paddy rice crop classification using RapidEye multi-spectral satellite image data. International Journal of Remote Sensing, 1–23. doi: 10.1080/01431161.2014.965285

Kim, H.-O., & Yeom, J.-M. (2015). Sensitivity of vegetation indices to spatial degradation of RapidEye imagery for paddy rice detection: a case study of South Korea. GIScience & Remote Sensing, 52(1), 1–17. doi: 10.1080/15481603.2014.1001666

Nguyen, T. V. T., & Moschitti, A. (2012). Structural reranking models for named entity recognition. Intelligenza Artificiale, 6(2), 177-190.

Vuojala-Magga T., Turunen M., Ryyppö T., and Tennberg M. (2011). "Resonance strategies of Sami Reindeer Herders in Northernmost Finland during Climatically Extreme Years", Arctic 64 (2), 227-241

Zanoli, R., Pianta, E., & Giuliano, C. (2009). Named entity recognition through redundancy driven classifiers. Proc. of Evalita, 9.