

EO4NOWCAST

A project funded by the H2020 AI4Copernicus Project



Reinforcing the AI4EU Platform by Advancing Earth Observation Intelligence, Innovation and Adoption

D4.1: AI4EU NRT SM and flood nowcasting services and user documentation

Acronym	EO4NOWCAST		
Project Title	Earth Observation for Severe Weather Hazard Nowcasting		
Start Date	01/09/2022	Duration	14 Months
Project URL	TBA		
Contractual due date	31/08/2023	Actual submission date	31/08/2023
Nature	R = Prototype	Dissemination Level	PP = Restricted to other program participants (including Commission Services)
Author(s)	Emilio Calvauna, Matteo Colli and Sara Zani		
Contributor(s)			
Reviewer(s)	Matteo Colli		



This project has received funding from the AI4Copernicus project that is funded by the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 101016798.

Document Revision History *(including peer reviewing & quality control)*

Version	Date	Changes	Contributor(s)
V1.0	02/08/23	Report structure draft	Matteo Colli
V1.1	04/08/23	Contributions	Sara Zani, Emilio Calvauna
V1.2	31/08/23	Final review and QC	Matteo Colli

Executive Summary

As a specific output of the AI4Copernicus 3rd Open Call sub-project entitled “Earth Observation for Severe Weather Hazard Nowcasting - EO4NOWCAST” this document is a complementary output of the “prototype” deliverable – D4.1 AI4EU NRT SM and flood nowcastings services and user documentation – and provides ancillary information on the activity performed by the consortium and the principal results within the project Task 4.1 “AI4EU services development”.

This document includes reference information for the users of the services developed in EO4NOWCAST.

Table of Contents

1	Introduction	7
2	Report Purpose and Scope.....	8
3	Near Real-Time Soil Moisture mapping service.....	9
3.1	Microservice description.....	9
3.1.1	Service architecture	9
3.1.2	Access to the AI4Europe asset.....	11
3.2	User’s guide.....	11
3.3	Use case and sample dataset specifications	16
4	Pluvial Flood Nowcasting service.....	20
4.1	Microservice description.....	20
4.1.1	Service architecture	20
4.1.2	Access to the AI4Europe asset.....	21
4.2	User’s guide.....	22
4.3	Use case and sample dataset specifications	23

List of Figures

Figure 1 EO4NOWCAST NRT SM MS architecture.....	9
Figure 2 EO4NOWCAST NRT SM real time pipeline.	10
Figure 3 EO4NOWCAST web app login page.	12
Figure 4 EO4NOWCAST webapplication, key elements.	12
Figure 5 Time-frame selection.	13
Figure 6 Error message for missing data.	14
Figure 7 Details subsection.	14
Figure 8 EO4NOWCAST Layer Selection.	15
Figure 9 Representation of Valpolcevera basin (source: “Relazione Generale – Piano di Bacino Stralcio per la Tutela del Rischio Idrogeologico del Torrente Polcevera” - http://www.pianidibacino.ambienteinliguria.it/GE/polcevera/documenti/GE_Polcevera_RelazioneGenerale_DG435_2021.pdf)	16
Figure 10 Preview of the input domain (in green) and Valpolcevera basin (dotted grey lines).	17
Figure 11 On the left, the original NDMI satellite tile; on the right, NDMI pre-processed tile as raster image.	18
Figure 12 On the left, example of 5-minute accumulated precipitation raster file opened in QGIS. On the right, the color bar – it ranges from black (no rain) to white (maximum rain value).	19
Figure 13 EO4NOWCAST Pluvial Flood Architecture.....	20
Figure 14 PF pipeline.	21
Figure 15 Flood Nowcasting service User Interface.....	22

List of Tables

Table 1 Features of SM processing chain input geometry.	17
Table 2 Soil Moisture indicator maps for the Aol.....	18
Table 3 Real-time dataset of accumulated precipitation features.....	19

List of Terms & Abbreviations

Abbreviation	Definition
DSS	Decision Support System
CLMS	Copernicus Land Monitoring Service
S-1	Copernicus Sentinel 1 program
S-2	Copernicus Sentinel 2 program
AoI	Area of Interest
CGLS	Copernicus Global Land Service
SSM	Surface Soil Moisture
SWI	Soil Water Index
NDMI	Normalized Difference Moisture Index
NIR	Near Infrared
SWIR	Short Wave Infrared (SWIR)
S-2 L2A	Atmospherically corrected MSI data by the Copernicus Sentinel-2 mission
ARPAL OMIRL	Regional observation sensor network of the Regional Environmental Protection Agency of the Liguria Region
NRT SM	Near real-time soil moisture [-]
VM	Virtual machines
RI	Rainfall intensity [mm h^{-1}]
RA	Rainfall accumulation amount [mm]

1 Introduction

The objective of EO4NOWCAST is to realize and demonstrate an operational and replicable approach to assess severe weather events and related short-term hazards (nowcasting) built upon the synergy between EO and rainfall monitoring products. With the support of the AI4Copernicus 3rd Open Call for Experiments, Artys is experimenting an innovative Decision Support System (DSS) for the security of infrastructures and cities through reliable and timely information on the current hydro-geological conditions. The workflow is based on two machine learning products: the EO-derived near real-time (NRT) soil moisture (SM) maps and the pluvial flood now-casting. At the end of the project, they will be published in the AI4EU platform as Docker microservices that may address multiple application domains (mainly security, agriculture).

With respect to already available services, EO4NOWCAST aims to offer the following benefits:

- “ready-to-use” approach for natural hazard alerting: the EO4NOWCAST AI model exploits EO, meteorological and geographical datasets commonly available in any region;
- continuously updated pluvial flash flood mapping for operational use: EO4NOWCAST enables addressing the long-standing problem of nowcasting weather-related hazards in quick response drainage systems such as the urbanized Mediterranean mountain and coastal areas;
- EO-based EO4NOWCAST's products concern a broad range of users: updated soil moisture is a key information for multi-hazard assessment (landslides, river flood and urban floods);
- EO4NOWCAST's user-centered and ergonomic dashboard will provide the most relevant information for decision making process, particularly useful in critical and stressful situations.

In fact, the specific objectives of EO4NOWCAST are the following:

- To develop and demonstrate an AI-driven workflow based on the use of Copernicus Land Monitoring Service imagery, AI4Copernicus processing pipelines and available rainfall monitoring products to monitor the Soil Moisture (SM) at a local scale.
- To develop and demonstrate a Decision Support System (DSS) aimed at the pluvial flood now-casting in small-medium sized watersheds based on the local scale Near Real-Time (NRT) SM map product elaborated by the first EO4NOWCAST product.
- To release on the AI4EU platform two services that implement the NRT SM mapping and the Pluvial Flood Nowcasting workflows along with training and validation datasets for the operationally relevant case of the flood-prone river basins in the Genoa Metropolitan City (Italy).

This document provides the companion information necessary for the use of the AI assets and the service prototypes released as the output of WP4 and object of Task 4.2.

2 Report Purpose and Scope

The purpose of Task 4.1 is to describe the AI4EU NRT SM and flood nowcasting services and provide the user documentation.

Within this framework, the present report (D4.1) documents each microservice functioning, showing the architecture and detailing the concept behind it, the user guide and information on its sharing in the AIoD CMS platform. Finally, the use case descriptions and sample dataset specifications are provided.

3 Near Real-Time Soil Moisture mapping service

3.1 Microservice description

3.1.1 Service architecture

The NRT SM mapping service comprises a collection of distinct microservices, each tailored to fulfill a specific function within the platform. The broad architecture can be visualised through the diagram provided below (Figure 1).

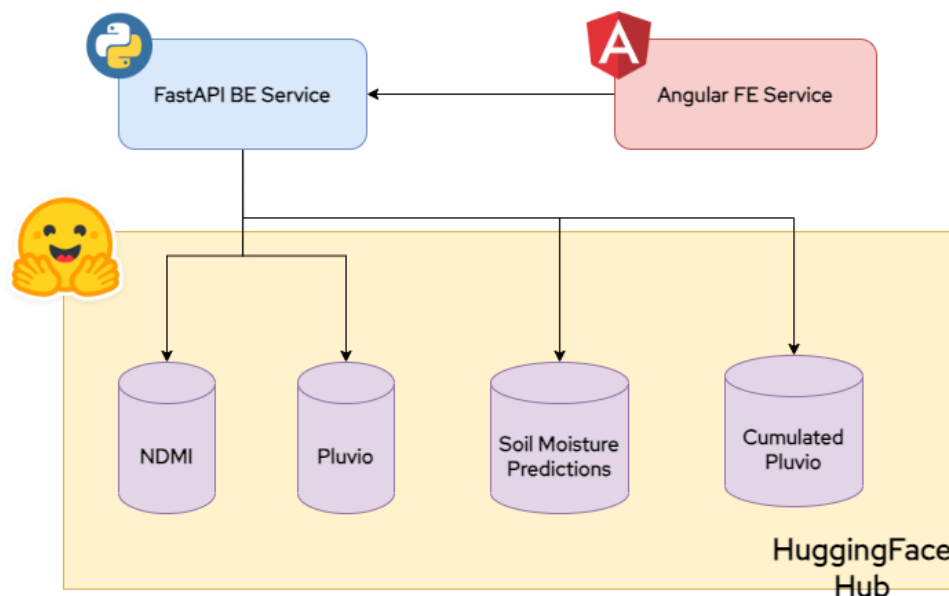


Figure 1 EO4NOWCAST NRT SM MS architecture.

The diagram in FIG.1 shows three main parts:

- the backend service, constructed using the Python FastAPI framework, facilitates the retrieval of datasets used by the machine learning model and real-time prediction pipeline. Through a collection of ReST API endpoints, clients can specify dates and acquire temporal slices of datasets or predictions. Serving as the principal gateway for external services, it manages access to platform data.
- the frontend service, which is an Angular web app globally available on cloud; it houses the primary graphical interface that lets users explore predictions and data on a map layered with geographical information. Users can define date and time parameters to visualize diverse soil moisture conditions across available timestamps. The user interface details are elaborated in section 3.2. The frontend exclusively communicates with the backend service, extracting prediction files and related data (e.g., Pluvios and NDMI), which serve as input for the model.
- the HuggingFace hub, which serves as a comprehensive repository for storing all the requisite data for model training and prediction generation. The hub offers private datasets that can be handled in a GIT-like mode, and they can be accessed using private tokens generated by the HF platform. As described on the diagram, four datasets have been created:
 - the NDMI dataset, containing the real Normalized Difference Moisture Index raster files extracted from the Copernicus hub and tailored for the E4NC use case;
 - the Pluvio dataset, containing rasterized data related to rainfalls during a particular hour of a specific day, sourced from the Copernicus platform;

D4.1: AI4EU NRT SM and flood nowcasting services and user documentation

- the Cumulated Pluvio dataset, which holds summed Pluvio data for specific timeframes, useful to the machine learning model to better learn about precipitation changes;
- the Soil Moisture Predictions dataset, which contains the actual predictions generated by the model for a specific timestamp, again in a rasterized form.

Of the parts described above, only the BE and FE services are dockerized, since the HuggingFace hub is already cloud-based.

Diving a little more on the backend APIs, the available endpoints are the following:

- ``GET /acquired-ndmis`` which allows to retrieve the real NDMI raster data, specifying a date as request body;
- ``GET /prediction`` which allows to retrieve a model-generated soil moisture prediction, again specifying a date for that prediction;
- ``GET /cumulated-pluvio`` which allows clients to acquire a raster of cumulated pluvio data for the specific date. The sum of Pluvio is based on the previous NDMI of the specified date. This endpoint is mainly used by the FE service to show prediction details.

To generate prediction in a near-real-time fashion, we leveraged the use of Apache Airflow, an open-source platform widely used for orchestrating and automating data workflows and pipelines, defining them as DAGs (Directed Acyclic Graphs) where tasks represent the nodes of the graphs and dependencies between tasks define the order in which they should be executed.

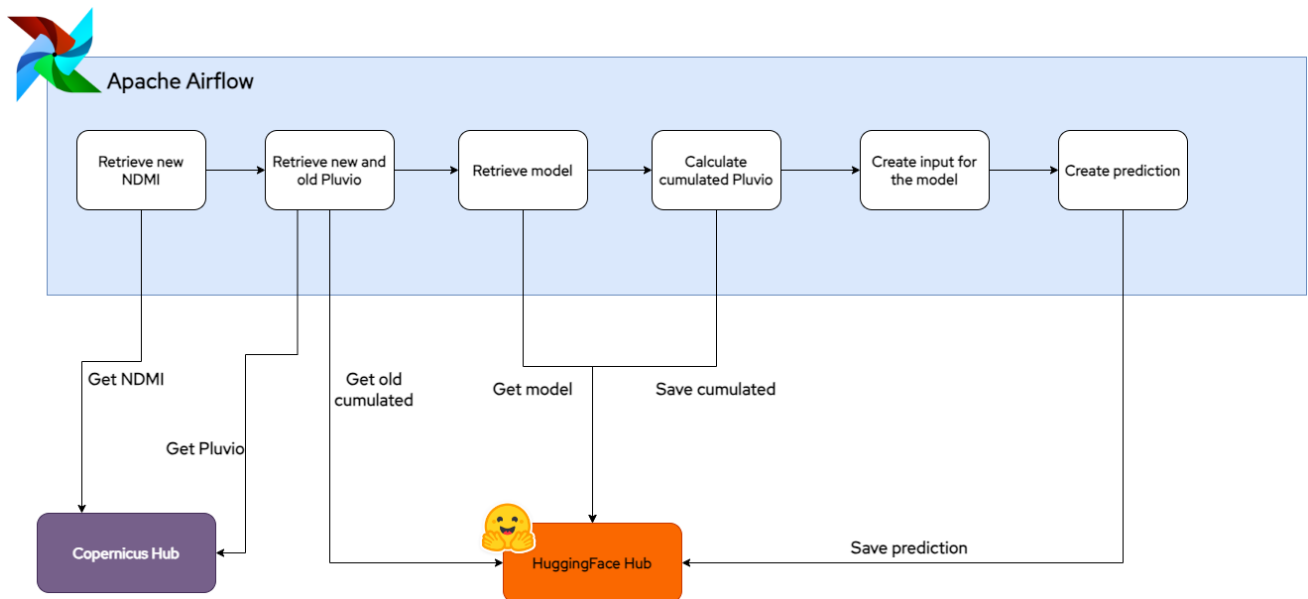


Figure 2 EO4NOWCAST NRT SM real time pipeline.

In Figure 2 we show a simplified image describing how the pipeline works. The tasks are the following:

1. retrieve the newest NDMI from the Copernicus Hub
2. retrieve new Pluvio data from the Copernicus Hub, and old Pluvio data from the HuggingFace dataset hub
3. retrieve the machine learning model from the HuggingFace model hub

4. add new Pluvio data to old Pluvio data, and save the resulting cumulated data to Hub to be retrieved in next iterations of the pipeline
5. use cumulated Pluvio and NDMI data to build the machine learning model input (called “tensor” in Deep Learning terms)
6. give the input data to the model to generate the prediction, which is then saved in the HuggingFace dataset hub keeping the timestamp as metadata.

The pipeline is implemented as a DAG in the Airflow system, which is deployed as a Kubernetes cluster since it uses different microservices (like PostgreSQL and Redis) to correctly work. The Airflow framework is not publicly available as a Dockerized service since it is a helper service, hence not useful for private clients who want to deal with already-generated predictions.

3.1.2 Access to the AI4Europe asset

The NRT SM mapping service is available into the AI Assets Catalog in the AI on Demand platform at the following link

<https://www.ai4europe.eu/research/ai-catalog/eo4nowcast-near-real-time-soil-moisture-assessment-and-pluvial-flood-nowcasting>

where it’s possible to download the docker container, which allows to start an application that includes:

- the pipeline for the hourly prediction of the NDMI
- the pipeline for the 5-min prediction of the flood risk

Pipelines are scheduled, respectively, every hour and every 5 minutes. It is possible to monitor the status of each pipeline by means of a web console available once the application is up and running.

The attachment folder contains a README.md file describing all the instructions to perform the EO4NOWCAST Near Real-Time Soil Moisture Assessment and Pluvial Flood Nowcasting services.

3.2 User’s guide

The NRT SM mapping service was developed and deployed as a single-page web application with the ambitious goal of conveying a rich set of information in a simple, intuitive way. The solution is password protected and at the moment it does not feature an open registration functionality as the foreseen delivery model is not for a fully open app but it requires a setup and configuration for each specific user.

The authentication layer was implemented using Keycloak, an open-source software product to allow single sign-on with identity and access management aimed at modern applications and services (Figure 3).

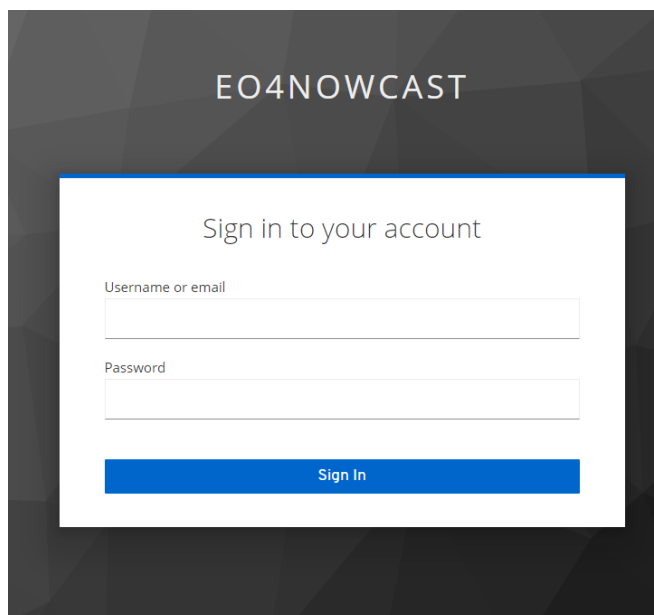


Figure 3 EO4NOWCAST web app login page.

After logging in, the EO4NOWCAST NRT SM mapping service is presented with a GIS-layer and a set of tools and elements that enable users to tune inputs and outputs of the application to obtain the desired information and results (Figure 4).

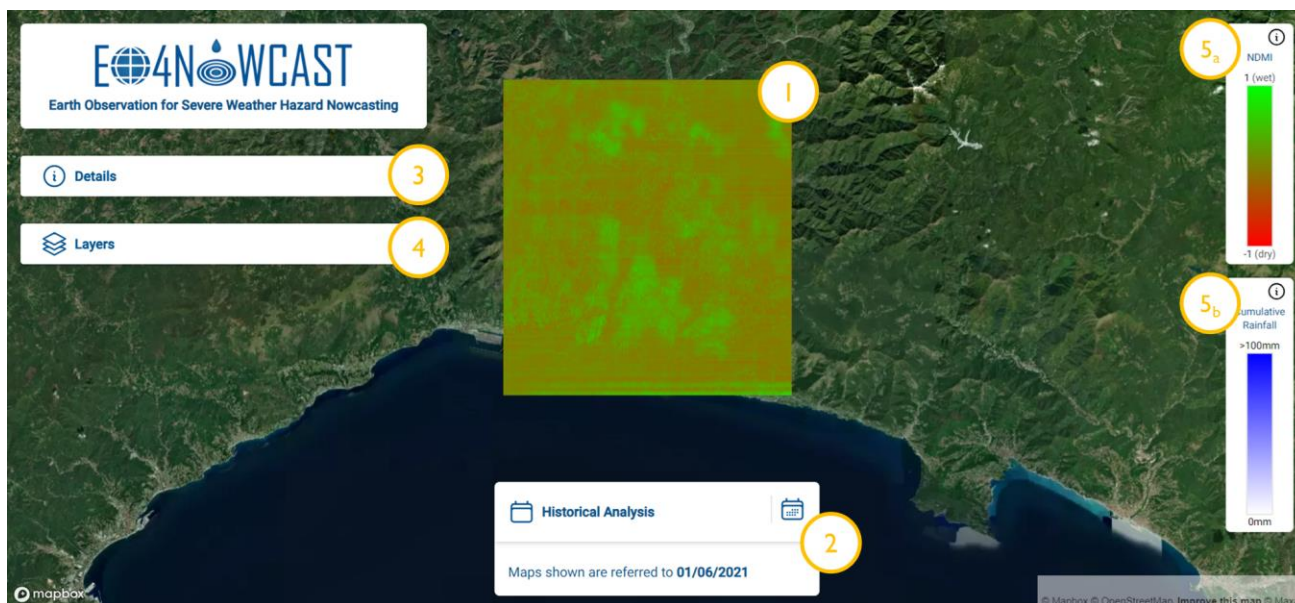


Figure 4 EO4NOWCAST webapplication, key elements.

1. Target area map

This is the target area in which data were gathered, processed and used for forecasting activities. For the sake of the project this was focused on a part of the Liguria region, with a high risk of heavy rains and floods. The Liguria region, located in the northwest of Italy, is renowned for its picturesque coastlines, historical towns, and mountainous terrain. However, over the last decade, this region has

faced an increasing threat from heavy rains and resultant floods. The unique topography, with steep mountainous slopes leading to the sea, combined with the urbanized areas, creates conditions for rapid water runoff, leading to flash floods. The small river streams, known locally as 'torrenti' or 'fiumi,' can quickly become overwhelmed. In recent years, attributed partly to climate change, Liguria has seen a notable increase in the frequency and intensity of rainfall events. The combination of heavier rainfalls and the region's urban development has made certain areas particularly vulnerable to inundation. Cities such as Genoa (the center of the target area) have witnessed catastrophic events, with loss of lives and significant damage to infrastructure. As the patterns of climate change continue to evolve, adaptive measures and better urban planning are essential to mitigate the risks associated with these natural disasters in the Liguria region.

This map view presents a set of layers containing relevant information such as Soil Moisture and Cumulative rainfall, both as historical data and AI-forecasted data.

2. Time-frame selection

By clicking on the calendar icon, the user is led to a data-picking dialog that allows her/him to tune the timeframe of analysed and presented data (Figure 5).

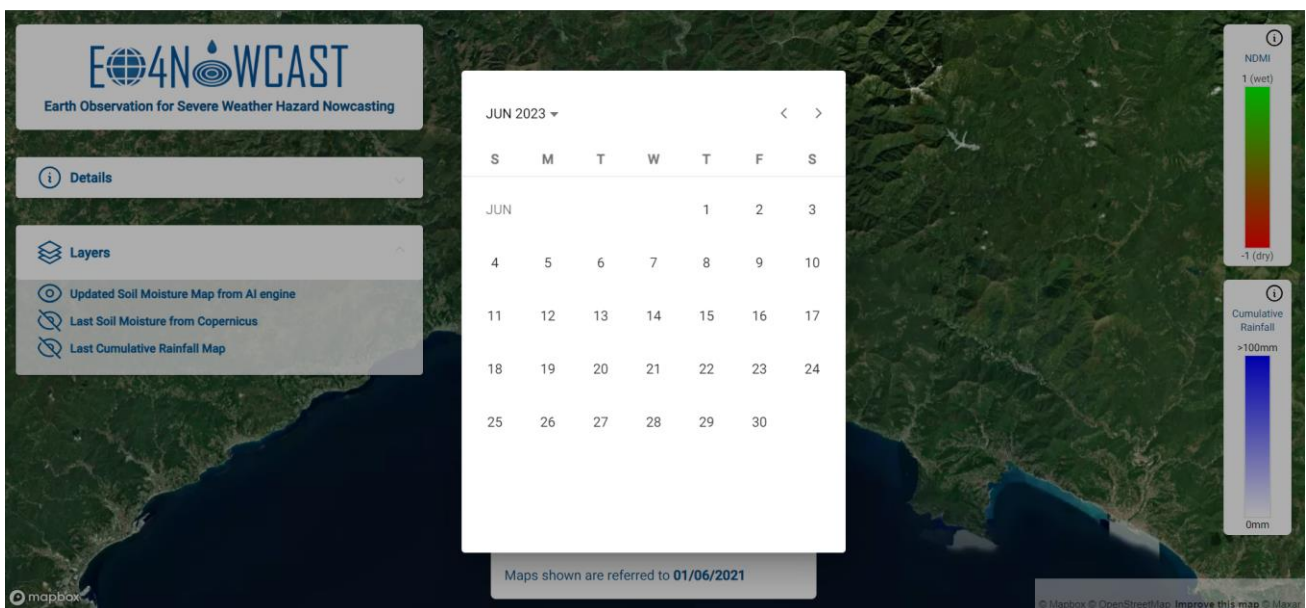


Figure 5 Time-frame selection.

The time-frame selection allows to choose a date for which to present Soil Moisture and Rainfall data as well as forecasted soil moisture. In case an unavailable date (i.e., a date not uncovered by the available dataset) is selected, the application displays an error (Figure 6).

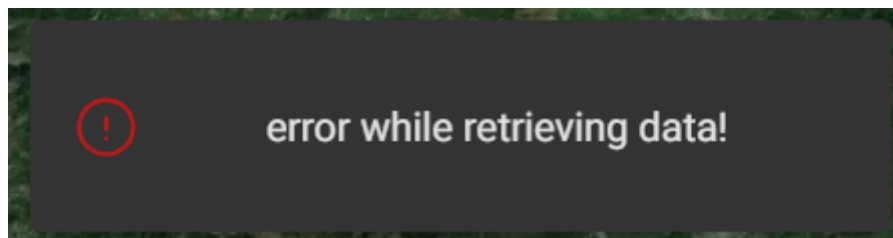


Figure 6 Error message for missing data.

3. Details section

As it was previously mentioned, the EO4NOWCAST web app aims at conveying a wide set of information in a simple interface. For this reason, a specific section of the application, accessible by clicking on the “Details” button, is devoted to explaining the presented data (Figure 7).

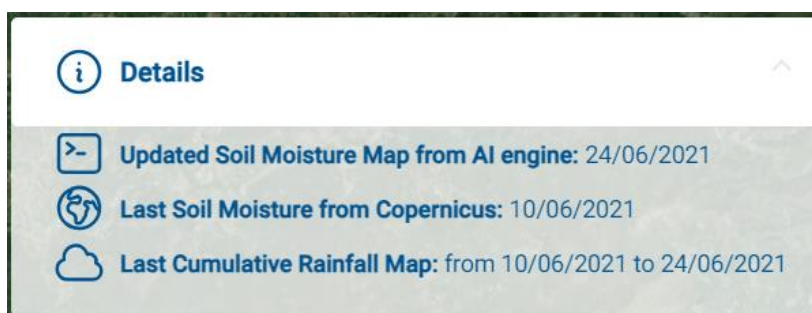


Figure 7 Details subsection.

This subsection briefly describes and presents the layers that can be seen on the map. In particular, the available historical data for each quantity is presented as well as the AI-forecast for the target date.

4. Layers Section

This section allows to select and toggle between different layers (Figure 8). In particular,

- *Updated soil moisture Map from AI engine* shows the data forecasted by the AI engine for the selected date, based on historical Soil Moisture data;
- *Last Soil Moisture from Copernicus* shows the last available Copernicus data used for the forecasting;
- *Last Cumulative Rainfall Map* shows the last historical data of cumulated Rainfall for the week prior to the selected date.

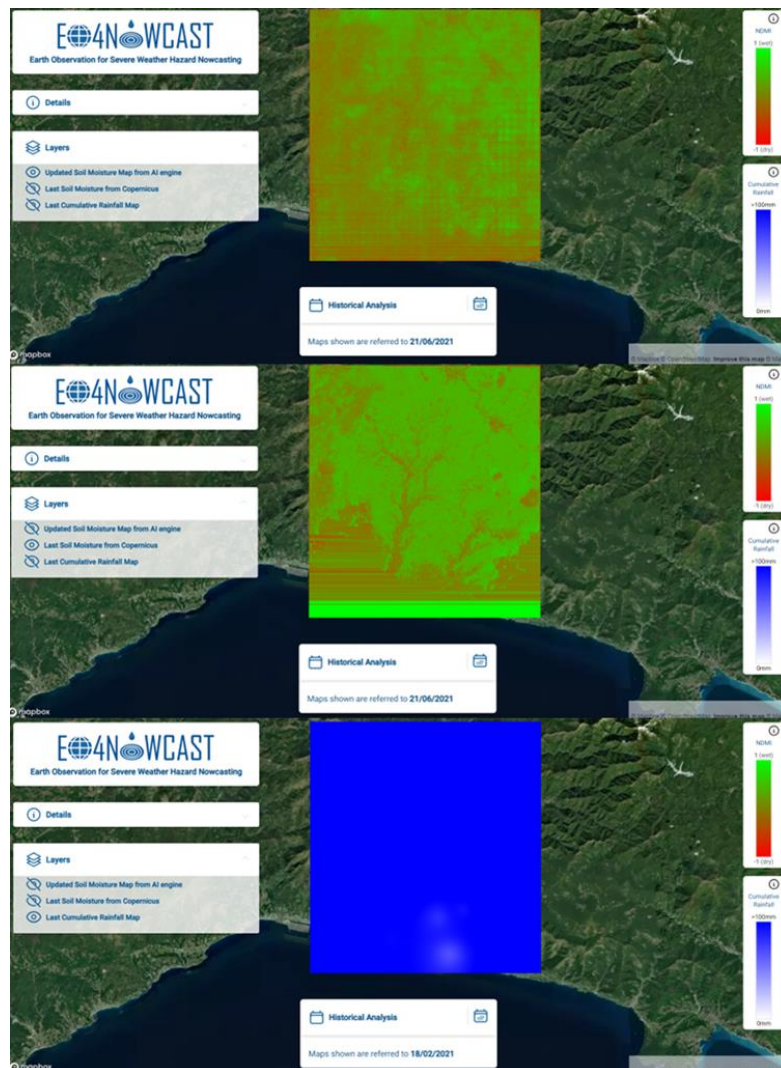


Figure 8 EO4NOWCAST Layer Selection.

5. Color Map and Legend

On the right side of the map page a color map for the displayed quantities is available for users to better understand the qualitative (i.e., color) data. Moreover, to convey clear messages and make colors and figures understandable to all users, an info (i) button with mouse-over activation was added to the legenda. By just passing the mouse over the button the following messages appear:

- **NDMI:** *NDMI is an index used in the analysis of satellite images to assess the amount of moisture present in vegetation or soil. This index is calculated using near-infrared and shortwave infrared bands in the spectral range of satellite images.*
- **Cumulative rainfall:** *Cumulative rainfall is the amount of rainfall calculated between a recorded NDMI and the date for which the NDMI forecast is requested.*

3.3 Use case and sample dataset specifications

Valpolcevera hydrographic basin has been chosen as area of EO4NOWCAST project's microservices experimentation. The basin is entirely included within the province of Genoa, Italy, and incorporates the municipalities of Genoa, S.Olcese, Campomorone, Serra Riccò, Mignanego and Ceranesi partially or completely. Its catchment area is of about 140 kmq.

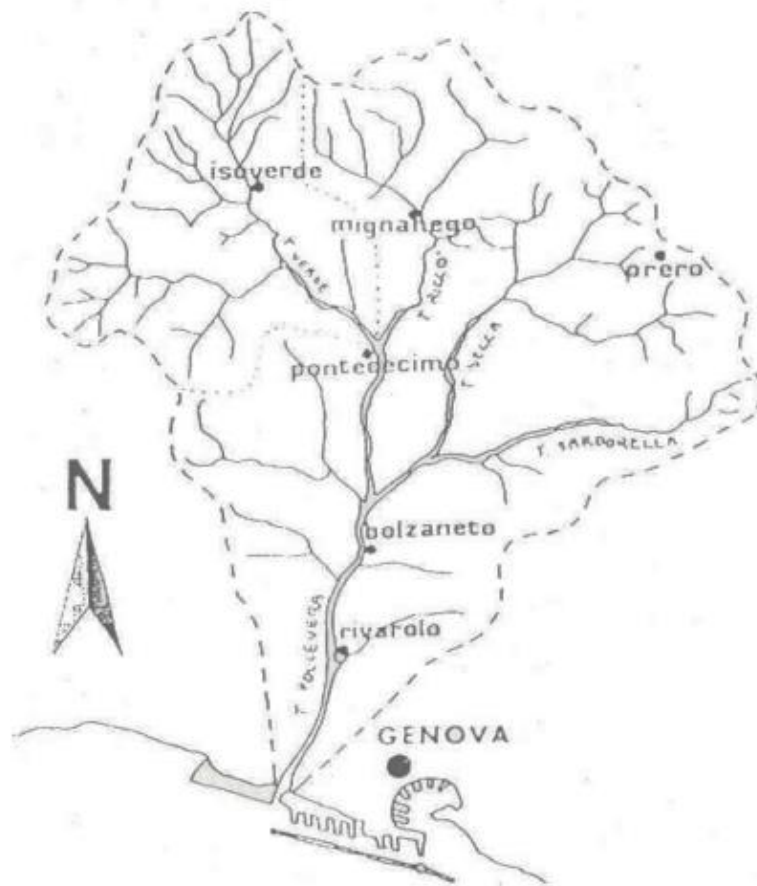


Figure 9 Representation of Valpolcevera basin (source: "Relazione Generale – Piano di Bacino Stralcio per la Tutela del Rischio Idrogeologico del Torrente Polcevera" - http://www.pianidibacino.ambienteinliguria.it/GE/polcevera/documenti/GE_Polcevera_RelazioneGenerale_DDG4_35_2021.pdf)

As fully detailed in D2.1, a polygon covering Valpolcevera basin is available to be assigned as input parameter to the NDMI pre-processing chain (Figure 10). This geometry has been used in every stage of development of all microservices for consistency purposes.

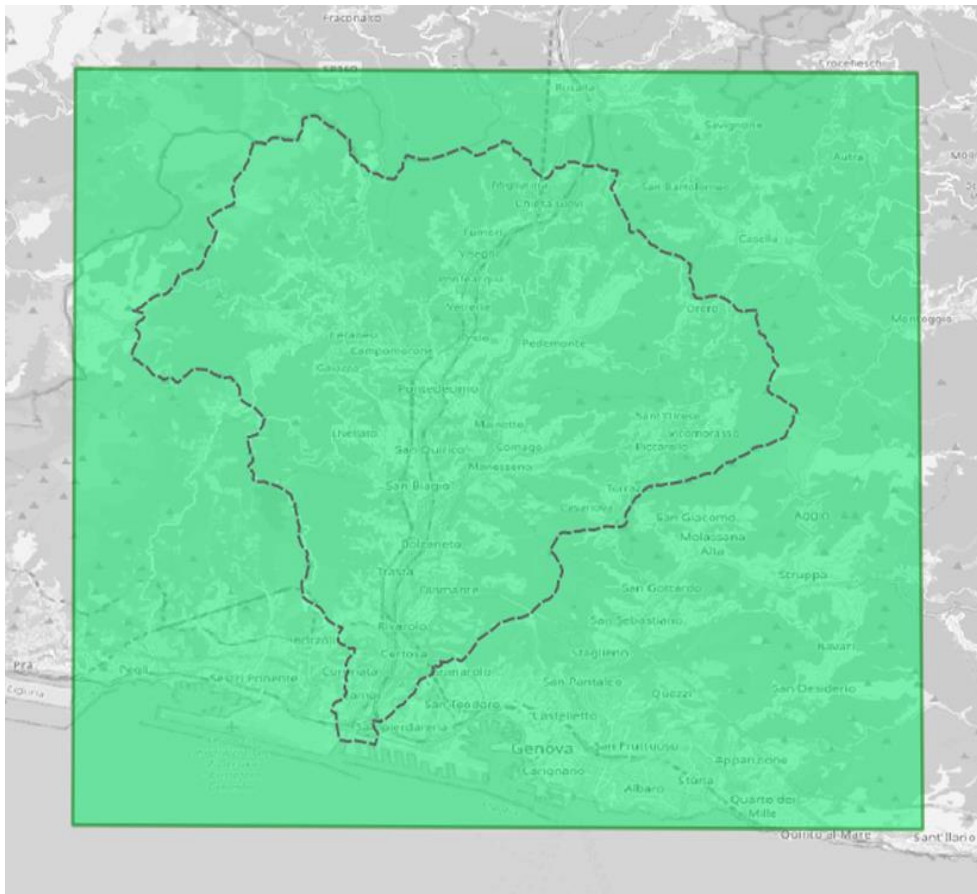


Figure 10 Preview of the input domain (in green) and Valpolcevera basin (dotted grey lines).

In Table 1, the geometry file’s features are shown.

Table 1 Features of SM processing chain input geometry.

SRID	EPSG:4326
Geometry type	Polygon (GeoJSON)
Attribute name, format, and description	The «name» attribute (<i>string</i> format) identifies the name of the zone of interest (i.e., city, basin, country, ...). Only the «underscore» character is admitted if desired.

The NRT SM service performance comprises the pre-processing of NDMI satellite tiles resulting into raster images (Figure 11).

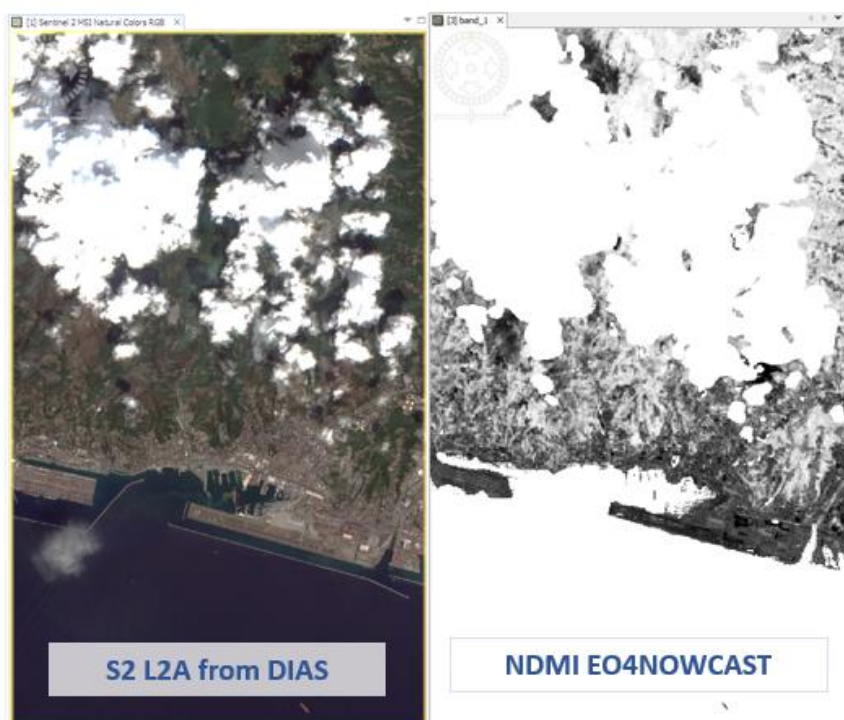


Figure 11 On the left, the original NDMI satellite tile; on the right, NDMI pre-processed tile as raster image.

In Table 2, NDMI map’s features are illustrated.

Table 2 Soil Moisture indicator maps for the Aol.

Variable	Soil Moisture indicator – NDMI processed from S2 – L2A
Units	percentage or ratio
Range	from -1 (water stress) to 1 (saturation)
Representation	The SM will be represented as a map layer with a shade of colour proportional to the relative soil moisture variable. The chromatic color scale will range from red (SM=-1) to green (SM=1)
Updating interval	the same of Copernicus S2 L2A
Spatial resolution	the same of the S2 L2A products, i.e., 20 meters.

To perform the SM predictions via NRT SM microservice, a dataset of cumulative values of rainfall within the past 5-minutes and updated every 5 minutes have to be provided. Within the project, the real-time source of precipitation is a network of sensors monitoring Valpolcevera basin based on the Smart Rainfall System (SRS) technology.

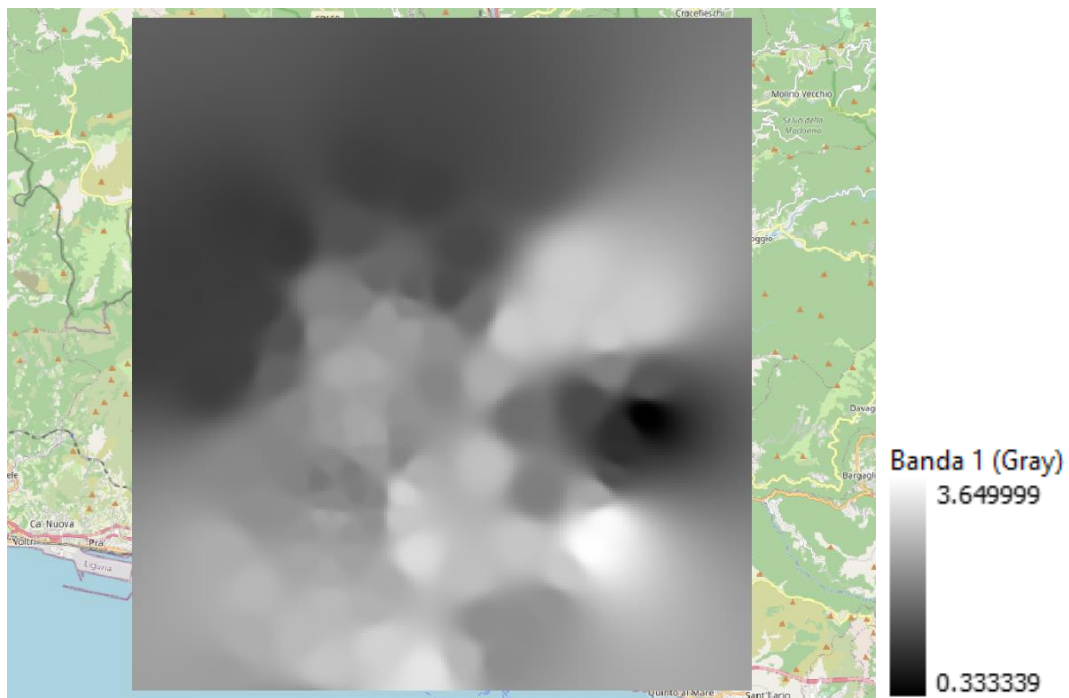


Figure 12 On the left, example of 5-minute accumulated precipitation raster file opened in QGIS. On the right, the color bar – it ranges from black (no rain) to white (maximum rain value).

In Table 3, the precipitation dataset’s features are shown.

Table 3 Real-time dataset of accumulated precipitation features.

Filename	cum5_VALPOLCEVERA_<TIMESTAMP>.tif where TIMESTAMP = YYYYMMDDhhmm (Y=year, M=month, D=day, h=hour, m=minute)
Geometry type	Raster
Attributes information	Each pixel represents the accumulated precipitation [mm] within the past 5 minutes
Timezone	UTC
SRID	EPSG:32632 - WGS 84 / UTM zone 32N
Update time	5 minutes

4 Pluvial Flood Nowcasting service

4.1 Microservice description

4.1.1 Service architecture

The PL NS has an architecture which is basically the same as the previous Soil Moisture service described in 3.1.1. The schema is the following (Figure 13):

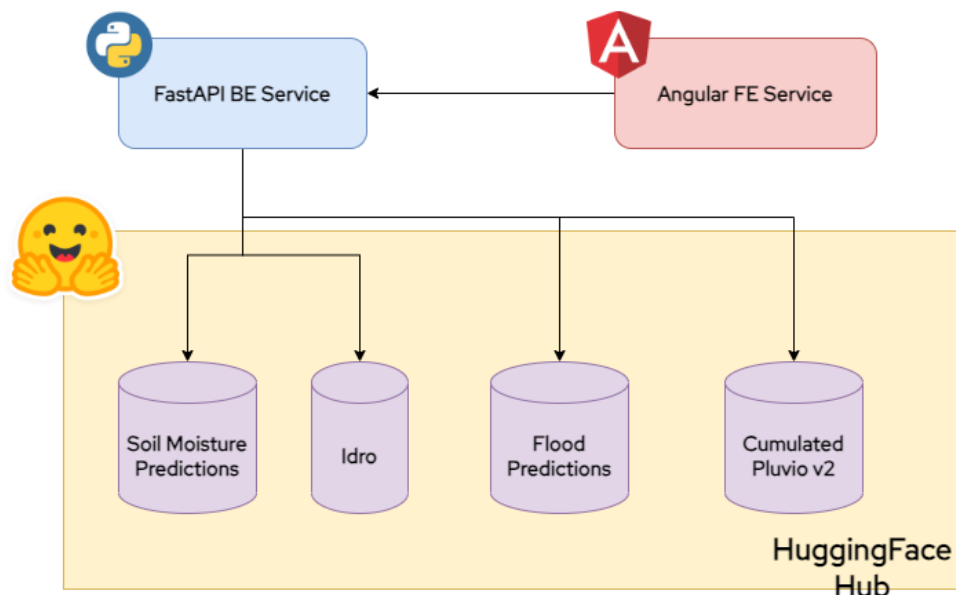


Figure 13 EO4NOWCAST Pluvial Flood Architecture.

The main microservices are the same ones used in the previous product, i.e., the BE service and the FE service. The backend service is enhanced with new endpoints:

- ``GET /flood-prediction`` which returns a model-generated flood prediction for a specific date, given as parameter in the request;
- ``GET /idro`` which gives back a JSON body containing the water level corresponding to the tracked stations in the river, again for a specific timestamp, given as input in the HTTP request.

while the FE adds to the user interface different components regarding the visualization of flood levels, danger labels and date selectors (which differ from those related to the soil moisture data). More on that on 4.1.2.

The HuggingFace repositories used in this service are also a little different, since they leverage different data to handle new model predictions, this time related, as said before, to flood dangers.

The repositories being used for this service are:

- the Soil Moisture Predictions dataset, explained in 3.1.1. We leverage this data as input for the machine learning model related to flood dangers.
- the Idro dataset, containing data related to the water level at the tracked stations with time frames of 15 minutes. This dataset is utilised during training to help the model understand levels of water related to specific combination of soil moisture and rainfall.
- the Flood Predictions dataset, here are stored the predictions generated by the PF model;

D4.1: AI4EU NRT SM and flood nowcasting services and user documentation

- the Cumulated Pluvio v2 dataset, which is different from the one described in 3.1.1 since it stores cumulated data within 15 minutes. This is the second input that the ML model used to generate its predictions.

Similarly to the NRT SM service, the dockerized microservices are again only the backend and the frontend.

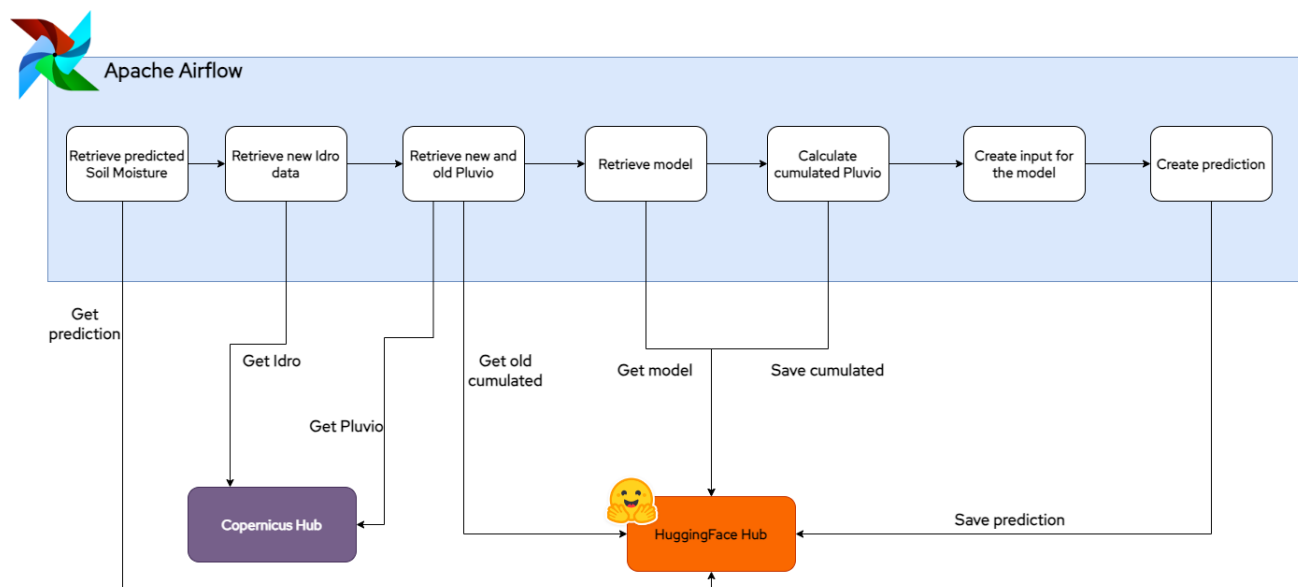


Figure 14 PF pipeline.

Regarding the real-time pipeline for this service, we deployed another DAG in the Apache Airflow system, as described in 3.1.1; this pipeline works similarly to the one used for the SM Service, having these tasks:

1. retrieve from HuggingFace dataset hub the predicted Soil Moisture for the current timestamp
2. retrieve from Copernicus hub the newly-created Idro data for the tracked stations
3. retrieve from Copernicus hub new Pluvio data, and from HuggingFace dataset hub the old cumulated data
4. retrieve the ML model through the HuggingFace model hub
5. construct the new cumulated data using the old one retrieved from HuggingFace and adding to it the new Copernicus rainfall data
6. use the predicted Soil Moisture in conjunction with the cumulated Pluvio to build the model input
7. use the model to generate the new flood prediction, which is then saved to the HuggingFace hub

Just like the SM service, the Airflow platform is not publicly displayed and therefore not included in the set of dockerized services.

4.1.2 Access to the AI4Europe asset

The Pluvial Flood Nowcasting mapping service is available into the AI Assets Catalog in the AI on Demand platform at the following link

<https://www.ai4europe.eu/research/ai-catalog/eo4nowcast-near-real-time-soil-moisture-assessment-and-pluvial-flood-nowcasting>

where it's possible to download the docker container including NRT SM mapping and Pluvial Flood Nowcasting services' pipelines.

Further details have already been described in section 3.1.2.

4.2 User's guide

For the most part the Pluvial Flood Nowcasting module overlaps in terms of functionalities with the NRT SM Map. In fact, all points and information presented in section 3.2 is still valid and can be replicated for this module as well.

However, the Pluvial Flood Nowcasting module further increases the information to be presented to the user, by adding information on the flood alerts based on the system prediction against warning thresholds (Figure 15).

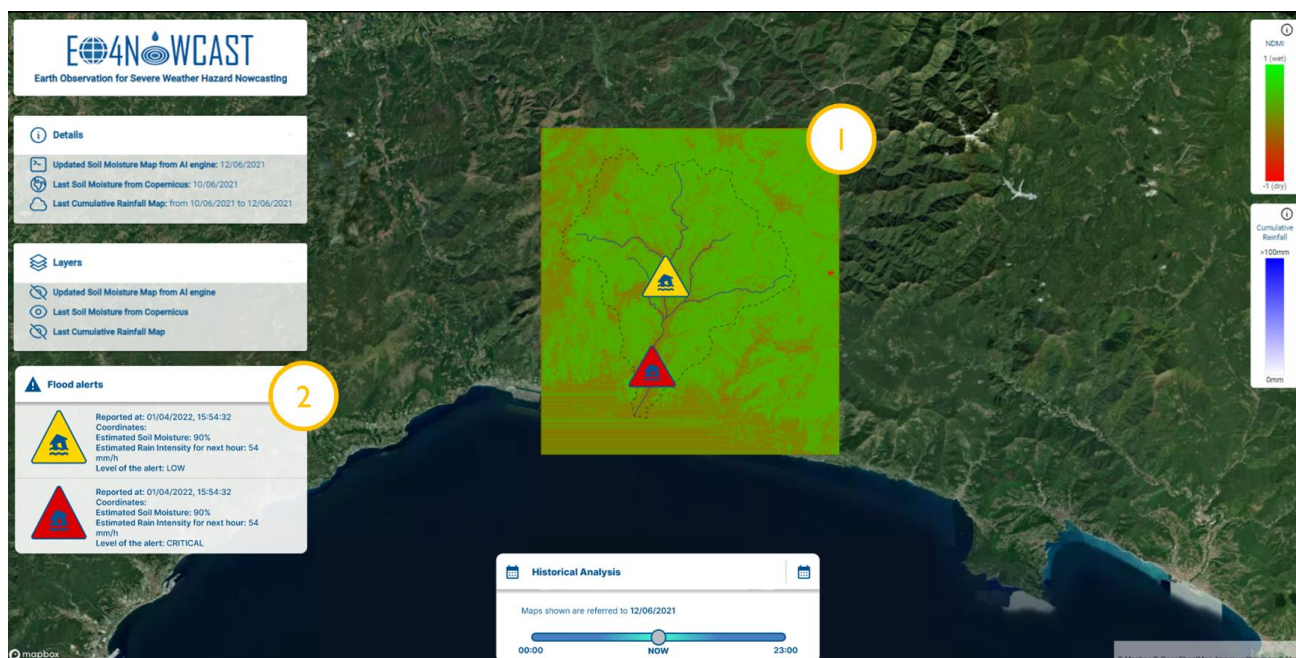


Figure 15 Flood Nowcasting service User Interface.

As it can be seen, the module leverages GIS data, specifically a dedicated layer to quickly convey qualitative information about the location of the alert (warning sign pinned on the map) and the level of the alert (pin color). This (shown in 1.) allows users to quickly identify issues and potential actions.

Moreover, on the left side of the UI (point 2.), quantitative information is displayed, allowing the user to further understand, process and react to the warning. Namely, the displayed information is:

- Reported At: warning date and time;
- Coordinates: warning location (i.e., center of the warning area);
- Estimated soil moisture: SM estimated by the AI engine for that specific moment and location;
- Estimated rain intensity for the next hour: Rain Intensity estimated by the AI engine for the warning location for the hour after the warning;

- Level of the Alert: qualitative (Low, Medium, Critical) level of the alert, based on threshold of watercourses levels.

4.3 Use case and sample dataset specifications

Valpolcevera basin has been chosen as AoI to perform the Pluvial Flood Nowcasting service. Indeed, many subareas within Valpolcevera basin are flood-prone; it may also occur that shallow landslides can be triggered frequently during flood events, when the rainfall intensity exceeds the infiltration speed of water in the soil and substrate.

The territorial context of the area and the geometry features have already been detailed in section 3.3, where the precipitation dataset's features used for both NRT SM and Pluvial Flood Nowcasting services have been described.