Project 780118

Horizon 2020

Innovation Action

ICT programme

Project: 780118 FANFAR

Full project title:
Reinforced cooperation to provide operational flood forecasting and alerts in West Africa

Deliverable: D3.4
Demonstrator of distribution channels deployed and pre-operational

FINAL VERSION

Due date of deliverable: 30/06/2018
Date accepted for publication/submission to EU: 24/08/2018
Actual submission date: 24/08/2018

Project 780118

Title	Demonstrator of distribution channels deployed and pre-operational
Creator	Fabrizio Pacini
Editor	Fabrizio Pacini, Emilie Breviere, Berit Arheimer
Brief Description	This document provides a description of the prototype built within the Hydrology-TEP to demonstrate the support for the generation on the platform, on-demand and systematically, of flood forecast information together with its distribution via the already supported distribution channels.
Publisher	FANFAR Consortium
Contributors	F. Pacini, E. Mathot, E. Boissier
Type (Deliverable/Milestone)	Deliverable
Format	Prototype
Creation date	18/06/2018
Version number	V2.0
Version date	13/07/2018
Last modified by	F. Pacini
Rights	Copyright "FANFAR Consortium". During the drafting process, access is generally limited to the FANFAR Partners.
Audience	☐internal ☐public ☐restricted, access granted to: EU Commission
Action requested	☐ to be revised by Partners involved in the preparation of the deliverable ☐for approval of the WP Manager ☐for approval of the Internal Reviewer (if required) ☐for approval of the Project Co-ordinator
Deadline for approval	30/06/2018

Version	Date	Modified by	Comments

Contents

List of Figures

List of tables

Executive Summary

The overarching concept of FANFAR is a multi-level, interdisciplinary cooperation built around user involvement. Two types of users will be engaged:

A. forecast producers, operating the hydrological forecasting and alert system, and
B. information end-users, utilizing the forecast and alert information for productive purposes in society (e.g. civil protection agencies, emergency response aid organizations, farming cooperatives, and reservoir managers).

The FANFAR cooperation is focused on providing an Operational Hydrological Forecasting and Alert (OHFA) system tailored for West Africa. Figure 1 below illustrates the essential components and data flows of an OHFA system whose core is a hydrological model, aimed at predicting the effects of meteorological dynamics (e.g. rainfall and temperature) on river flow, water level, soil moisture in rivers, lakes, wetlands, and all land surface areas. In FANFAR, we will use the HYPE model applied to the Niger River basin (Niger-HYPE, http://hypeweb.smhi.se/nigerhype), and adapt it to the entire West African domain. The forecasted hydrologic conditions generated by the model are compared against pre-defined thresholds to determine the severity of the situation and the potential alert level. This information is finally delivered to information end-users through various distribution channels (e.g. email with alert text sent in case of an alert, web visualization of the latest forecast time series, or FTP site with the current forecast and alert data).

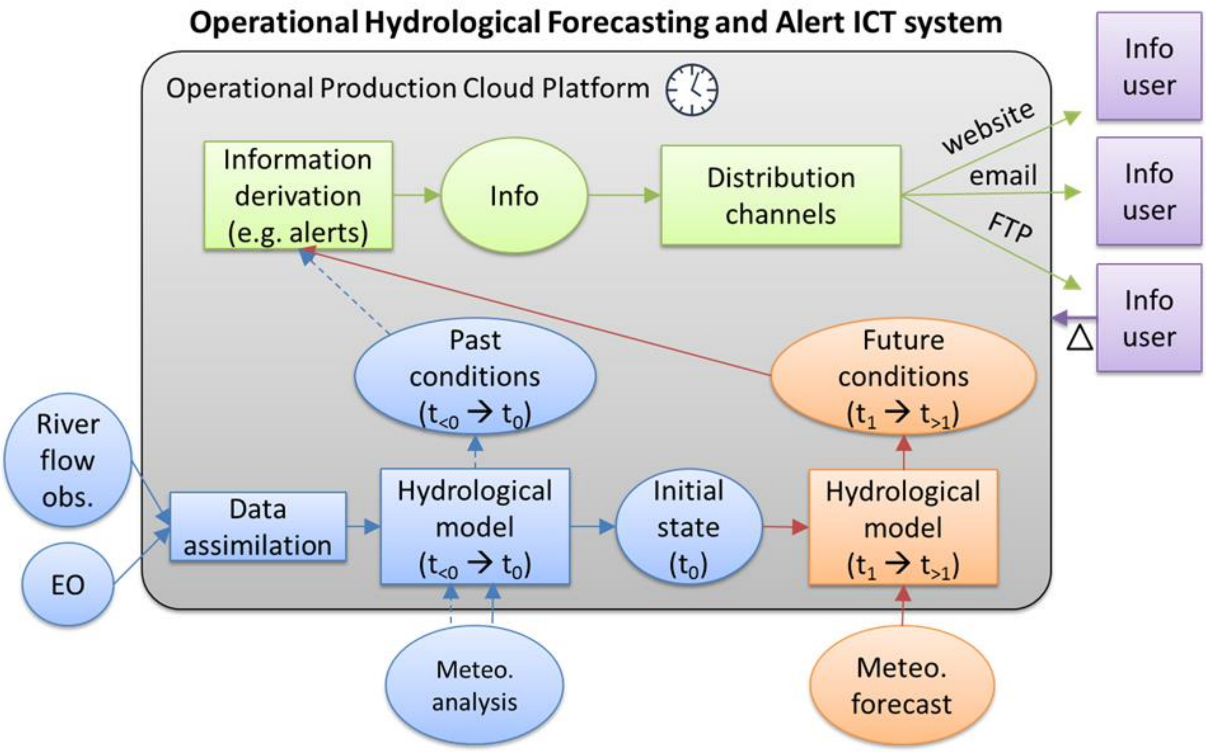


Figure 1: Schematic overview of the essential components and data flow of an Operational Hydrological Forecasting and Alert system (OHFA). Circles represent data, and boxes represent processors (computational models, scripts, and human interpreters)

Operational forecasting heavily relies on scheduled, automated execution and monitoring of the data flows and processing tasks on an Information and Communication Technology (ICT) environment that is always up and running, i.e. it requires an Operational Production Cloud Platform (OPCP). In FANFAR, this will be provided by the Hydrology Thematic Exploitation Platform (H-TEP) that already integrates the Niger-HYPE forecasting model. The platform will provide a dedicated workspace and a cloud-based processing infrastructure representing the FANFAR forecast production system in support of forecast producers (see bullet A. above). In addition it will support various distribution channels and will be complemented by a newly developed Early Warning portal to reach end-users (see bullet B. above) with the expected information.

Starting from the services and tools currently made available by the H-TEP, a prototype has been setup and deployed in pre-operations to demonstrate the FANFAR forecast production system capabilities to perform on-demand and automated/scheduled execution of the forecasting model (Niger-HYPE for the moment) as well as to push modelling outputs to selected distributions channels.

1. Introduction

This document provides a description of the prototype built within the H-TEP to demonstrate the support for the generation on the platform, on-demand and systematically, of flood forecast information together with its distribution via the already supported distribution channels.
It includes

- an introduction of the FANFAR forecast production system (the producers workspace) configured within the H-TEP
- a brief description of the solutions adopted to support the generation and distribution of the forecast model results
- a description of the scenario covered by the demonstrator with an example of execution of the scenario with related screenshots.

It is important to note that at this stage only a subset of the identified distribution channels is supported. Additional channels will be analysed and prioritised after the 1st project workshop with users planned on Q3 2018.

2. The H-TEP and the FANFAR forecast production system

FANFAR will use the H-TEP[1] as its operational production cloud platform. H-TEP is part of a larger TEP ecosystem developed and validated with the support of the European Space Agency (ESA).

The cloud-based H-TEP aims to facilitate knowledge creation, data processing and data handling for specific water-related communities, starting in the Niger River basin (West Africa) and the Red River (Vietnam). H-TEP has been developed to simplify the exploitation of satellite data for hydrology e.g.

[1] H-TEP - https://hydrology-tep.eo.esa.int/

through integration with hydrological models to improve flood forecasts. H-TEP is designed to provide a user-friendly cloud platform to discover, access and process Earth Observation (EO)-data, integrate results with hydrological models, visualize and save outputs. A H-TEP user is therefore able to access and generate computationally-demanding EO products and run hydrological models, while at the same time avoiding dependence on reliable internet connection for massive data download and investment in large computational resources.

The H-TEP is organized in users Communities and Thematic Apps (i.e. dedicated, customised working environments). A dedicated thematic app, representing the entry point to the FANFAR forecast production system, has been created within the H-TEP West Africa Hydrological Modelling community, and will be accessible only to the users of this community.

2.1 Community

A community in H-TEP simplifies the management of the users and help focus the user objective for thematic business within those groups. It is an important concept in the TEP to gather users around a specific aspect or an organization of the thematic. For instance, an institution may create its own thematic group with its selected members and experts. It may also have business relationship with data or ICT providers that are represented as a role within the group.

The aim of FANFAR is to reinforce the cooperation between West African and European hydrologists, ICT experts, decision analysts, and end-user communities to provide a co-designed, co-adapted, integrated, and co-operated streamflow forecasting and alert pilot system for West Africa. The West Africa Hydrological Modelling community aims at gathering all users interested in hydrological modelling in the West Africa area.

2.2 Thematic app

A thematic app brings a simple way to define an application of a specific aspect of the thematic. It specifies together the form of the application, its features such as the map and the layers, its data and services.
It aims at putting directly the user into the subject he seeks to. Instead of having one common application with all the collections, all the services, all the map layers, all the functions, the thematic application targets the scope to a product, a service, a collection or a function.

The FANFAR thematic app aims at providing to users the functionalities to operate the FANFAR forecast production system for West Africa.
For that purpose, it provides access to the specific Web Processing Services (WPS), data collections and tools.

2.2.1 Web processing services

These are the processing services that will be made available in the thematic app:

Forecasting model

- Niger-HYPE forecast
- Niger-HYPE historical simulation

Input data preparation

- Niger-HYPE EO data pre-processing
- IsardSAT water level service

Warning levels definition

- HYPE return period analysis

Additional services (to be added at a later stage)

- IsardSAT water extent service
- Flood mapping service (TBC)

2.2.2 Data collections

These are the data collections that will be made available in the thematic app and which can be used as possible input of the previous services:

EO data (altimetry)

- Sentinel-3
- Jason-2, Jason-3
- SARAL/ALtika
- Water Masks

Meteo/in-situ data

- Hydro-GFD from SMHI
  - already fetched regularly by the platform and made available to the processing services)
- Hydro-GFD and in-situ data from AGRHYMET
  - systematic data fetching to be setup upon reception of information from AGRHYMET about the data repository and the preferred ingestion mode (pull/push)
  - this will also include Hydro-GFD data provided to AGRHYMET by SMHI every day and adjusted with meteo observations (10 days old)

Additional data for water extent services (to be added at a later stage)

- Sentinel-1
- Sentinel-2

2.2.3 Tools

These are the tools that will be made available in the thematic app:

Store upload:

- Allows users to upload data in either a private or a community repository on the platform for direct use by processing services or visualization

Publication service:

- Allows to publish in the platform catalogue previously uploaded data enabling sharing, searches and referencing of the data

Background layers:

- currently the Niger River background is provided, a similar one from the whole West Africa will be added as soon as it made available by SMHI
- additional layers that will be needed in the future can be easily configured provided they are served by a OGC WMS service

2.2.4 Authorization

The H-TEP portal authorisation mechanism allows a great flexibility for managing users and groups and their permissions with the items in the system.
In order to enable all the requirements specific to the TEP, the community uses the security components as the following:

- **Tep.UserTep**, user registered via the authentication mechanism integrated in the portal (e.g. EO-SSO)
- **Portal.Group,** regrouping a set of Tep.UserTep put together for organisational purpose. For instance, all Terradue staff users are grouped in the Terradue Group.
- **Portal.Domain** is named "Thematic Group" and englobes all users, groups and objects having a thematic scope in common. For instance, there could be a "Volcanoes" thematic group that would have expert users in volcanoes monitoring, the data collections used for monitoring them (e.g. Sentinel-2 and 3), the features related to this domain (e.g. latest most important eruptions) and the all the processing services relative to volcanoes.
- The initial roles are defined as Tep.RoleTep

Objects identified and used in TEP authorization mechanism are

- Data collections or series
- Data packages
- Web processing services
- WPS job representing an instance of a processing service execution
- Cloud provider
- Thematic application that combines at user level the previous objects

3. Distribution channels for modelling outputs

One of the last fundamental steps in the FANFAR forecast production system supported by the H-TEP, is the distribution of information to users. A range of potential distribution channels exist (e.g. web visualisation, FTP, SMS, email, social media etc.), each requiring different types of data, and several of these channels will be operated in order to reach different type of users with tailored information as well as to provide redundancy, and thereby increase the chances that the information actually reaches the users.

The selection and prioritization of the FANFAR distribution channels will be done after the first project workshop in Q3 2018 and the related needed capabilities will be implemented later in the project.
At this stage two channels are already supported by the platform, being the distribution of forecast model outputs via HTTP and FTP to external infrastructures for further processing and visualisation. These channels can be configured in both the foreseen operation modes of the OPCP:

- Forecast-on-demand, in which users specify settings (or load previously saved settings) and launch the processing manually,
- Scheduled automatic execution of the forecasting chain based on saved settings

The solutions adopted to support the above mentioned functionalities are briefly described in the following subsections.

3.1 Distribution channels support in H-TEP

Distribution channels are organized in the TEP infrastructure according to the **model version** and the **processing run initiator**, identified as the delivery ID.
Indeed, the production service in the H-TEP platform, shall produce results for a given model version initiated by a data producer either on demand or with a configured coordinator. In this latter case of coordinated runs, the channels shall distribute results with a delivery ID including the coordination parameters for which the data is produced (e.g. production period, AOI...).

There is a 2-layered system for implementing the distribution channel. This system is directly supported in H-TEP.

1. Synchronous results storage.
This first layer is used to store persistently all results produced on the H-TEP platform. It uses two services: storage and catalogue: After a successful processing,

- the results are systematically sent to the persistent storage identified by the delivery ID;
- the results metadata are indexed in the catalogue.

At this stage, all results are directly accessible from the H-TEP platform for searching and downloading the data using the delivery ID among other filters.

2. Asynchronous results synchronization
The second layer offers the possibility to export the results to a remote system using several protocols. The synchronization is triggered by the arrival of new data in a synchronized folder on store system in H-TEP. Since the results folders in H-TEP are organized by delivery ID, configuring a synchronization scheduler for a distribution channel with a remote system is pretty straightforward.

For instance, in the case of the export of the results to NIHSA infrastructure, a synchronization agent shall be watching for new results in the root folder identified by the scheduler triggering daily run and producing new daily results in children folders (v3, v4, v5…). Upon arrival of new data in that folder, the system shall schedule a synchronization task that shall copy the new data to the FTP server in NIHSA infrastructure. All the transfer details (e.g. delivery ID target, protocol, credentials…) are configured within the synchronization agent.
The following protocols are planned for the initial deployment:

- FTP transfer
- HTTP POST (REST protocol)

3.2 Scheduled forecast modelling

The H-TEP component in charge of the orchestration of the processing workflows and their coordination in time is named coordinator.
It is based on Oozie which is a workflow scheduler system for Hadoop. Within Oozie, users application are structured and contained from bottom to top:

- in workflows, that execute jobs on Hadoop;
- in coordinators, that manage workflows based on date/data triggers;
- in bundles, that batch the coordinators.

In the same container described above, many programs can be called in a specific order with different processing settings such as task concurrency. Within a workflow, Oozie coordinates those programs calls using directed acyclic graphs (DAG). It also allows exporting user's workflow as a single application for an easy deployment of the application on a multi-tenant cluster.
In the proposed solution for orchestration, the Oozie coordination layer is the most important part, indeed it acts as a scheduler either driven by date or by data.
The coordinators are composed of configurable triggers that execute a specific action on a workflow (mostly submission). When a user exports his application from the cloud developer sandbox and deploys it on the H-TEP multi-tenant cluster, it can be contained in any of the three layers. As a simple workflow, the user's application is seen as a single service with job submission on demand via WPS. As a coordinator, the application is seen as a schedulable service with job scheduling settings (data source or dates); in this case the WPS interface exposes new jobs input parameters for the data or date-driven settings to be set at job submission. Finally, as a bundle, user may export his application as a set of coordinators with several advanced scheduling options.

In the context of FANFAR, scheduled processing (also referred to as 'data pipeline' in the following) is supported via the use of the so-called triggers that are the date-driven workflows that trigger the data transformation application (i.e. the processing service). The trigger is a workflow deployed on the H-TEP Production Center. The coordinator associated to the triggers instantiates it according to a date-driven approach. Every time the coordinator creates an instance of a trigger for a given time interval (defined as the coordinator configuration) it can pass as input catalogue entries that have their update (dct:modified) metadata value in the time interval so that newly available data can fed the processing as needed.

The scheduled run of the Niger-HYPE forecast services, but more in general any systematic processing in H-TEP, is supported by a tracking index and its associated tracking series.

This concept applied to our data processor dcs_hypeapps_forecast_workflow-0001 gives e.g. the index: https://catalog.terradue.com/hypeapps_forecast_workflow-0001.

This index has six series allowing full monitoring of the running pipeline:

- source-queue: a series for tracking the identifiers in the queue
- source-in: a series for tracking the systematic processing where the identifiers being processed are inserted
- source-out: a series for tracking the systematic processing where the identifiers successfully processed are inserted
- source-err: a series for tracking the systematic processing where the identifiers not successfully processed are inserted
- source-limbo: a series for tracking the systematic processing where the identifiers not successfully processed are inserted after a TBD criteria (see dedicated section in this document)
- results: a series for the thematic data where the generated products metadata and enclosures are stored

An entry in the tracking index (and associated series) is an OWS Context document that contains at least:

- An identifier (the identifier of the input product)
- A title
- A date/time coverage
- A spatial coverage
- A link pointing to the input resource (required for recovery scenarios)
- A published date (the date of the source-in stage)
- A category expressing the processing stage (source-queue, source-in, source-out, source-err or source-limbo)
- The generator as version of the processing
- The output opensearch offerings limited by the geo:uid element
- The processing offering
    - The URLs for the WPS GetCapabilities and Describe Process GET requests
    - The URL for the WPS GetStatus GET request
    - The URL for the WPS Execute POST request
    - The WPS POST request

The definition and setup of a data pipeline is supported by the module reported below just as an example, that allows the detailed exchange of information between the hydrologist and platform support team that is in charge of preparing the environment for starting the scheduled processing.

Table 1 Systematic processing information template

Data pipeline definition	**Data pipeline**	
	Identifier	hypeapps_forecast_workflow-0001
	Trigger	lp-hypeapps_forecast_workflow-0001

	Trigger repository	https://gitlab.com/ellip/triggers/ec-fanfar/hypeapps_forecast_workflow-0001
	Catalogue index (same as identifier)	hypeapps_forecast_workflow-0001
	Catalogue series in	https://catalog.terradue.com/hypeapps_forecast_workflow-0001/source-in
	Catalogue series out	https://catalog.terradue.com/hypeapps_forecast_workflow-0001/source-out
	Catalogue series err	https://catalog.terradue.com/hypeapps_forecast_workflow-0001/source-err
	Result series	https://catalog.terradue.com/hypeapps_forecast_workflow-0001/thematic
	Build (CI)	https://build.terradue.com/job/triggers/job/ec-fanfar/job/hypeapps_forecast_workflow-0001/
	Reference documents	
	Applicable documents	
	Thematic application	
	Application	hep_hypeapps_forecast_workflow-0001
	Application repository	https://github.com/ec-fanfar/hypeapps_forecast_workflow-0001/
	Build (CI)	https://build.terradue.com/view/Communities/job/communities/job/ec-fanfar/job/hypeapps_forecast_workflow-0001/
Coordinator parameters	**Coordinator parameters**	
	quotation	no
	_T2username	
	t2_coordinator_date_start	2018-06-01T00:00Z
	t2_coordinator_date_stop	2018-12-31T23:59Z
	t2_coordinator_period	0 0,12 * * *

Trigger parameters	**User mapping**	
	api_key	
	username	
	Data pipeline parameters	
	data_pipeline	hypeapps_forecast_workflow-0001
	wps_url	https://ec-fanfar-deployer.terradue.com/zoo/
	process_id	ec_fanfar_dcs_hypeapps_forecast_workflow-0001_1_2_5
	OpenSearch parameters	
	update (dct:modified)	${coord:formatTime(coord:dateOffset(coord:nominalTime(), -24, 'HOUR'), "yyyy-MM-dd'T'HH:mm:ss")}/${coord:formatTime(coord:dateOffset(coord:nominalTime(), -12, 'HOUR'), "yyyy-MM-dd'T'HH:mm:ss")}
	geom (geo:geometry)	POLYGON((-2.549 3.075,-2.549 22.025,16.787 22.025,16.787 3.075,-2.549 3.075))
	Production center parameters	
	lp_quotation	no
Application parameters	**Thematic parameters**	
	Sub Basin	
	Output variables	
	Assimilation	
	...	

4. Demonstration scenario

The demonstrator has been built around the use case described below.
An hydrologist/forecast producer starts processing of a hydrological model on the H-TEP

- either on-demand
- or systematic, (e.g. scheduled, daily run)

The model is actually composed by a chain of services as follows:

Water Level ⇒Input data preparation ⇒Niger-HYPE EO data pre-processing ⇒Niger-HYPE Forecast

The hydrologist/forecast producer wants that the generated results are:

- published in the H-TEP catalogue
- stored persistently on the H-TEP
- pushed to a set of external sites (distribution channels). At this stage these are
    - a HTTP server at SMHI for consumption by a dedicated web application
    - a FTP server at NIHSA for further processing and dissemination
- files related to each run must be distributed in folders whose naming is unique and linked to the production date
- specification of the enabled distribution channels and their characteristics must be configurable by the hydrologist.

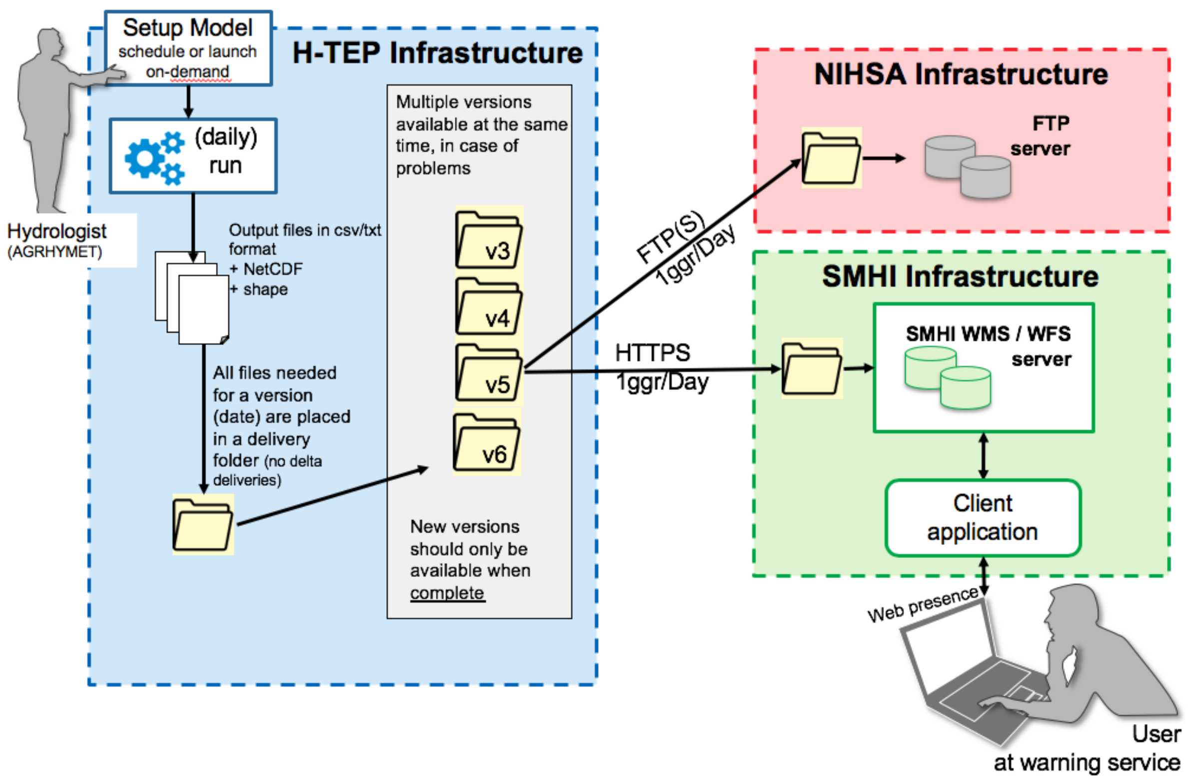


Figure 2 Demonstrator scenario: hydrological forecast generation and distribution

The deployed demonstrator fully supporting the introduced scenario is available here: https://hydrology-tep.eo.esa.int/geobrowser/?id=fanfar#.
The following section recounts the detailed steps involved in the execution of the on-demand branch of the introduced scenario.

4.1 Access to the FANFAR forecast production system

The FANFAR forecast production system is accessed from the home page of the H-TEP community portal:

1. Enter the **H-TEP portal** and **sign in** with your H-TEP community user account https://hydrology-tep.eo.esa.int
2. Open the list of **Thematic applications** with the "view apps" link below the Discover Thematic Apps icon.
3. Find the link to the FANFAR application (Open App):

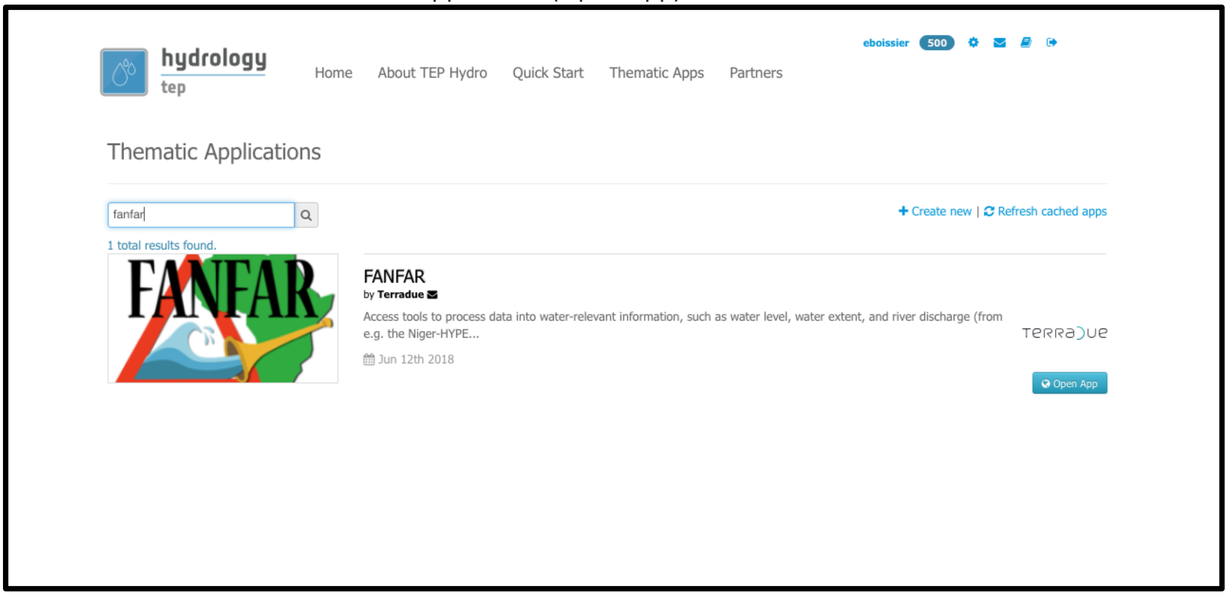


Figure 3: FANFAR forecast production system entry page

4. **Processing services** are found on the **Processing Services** tab on the right side of the geobrowser:
    a. Niger-Hype simulation of historical period
    b. Niger-HYPE forecast
    c. EO-data pre-processing
    d. Return period magnitude analysis
    e. Water level service

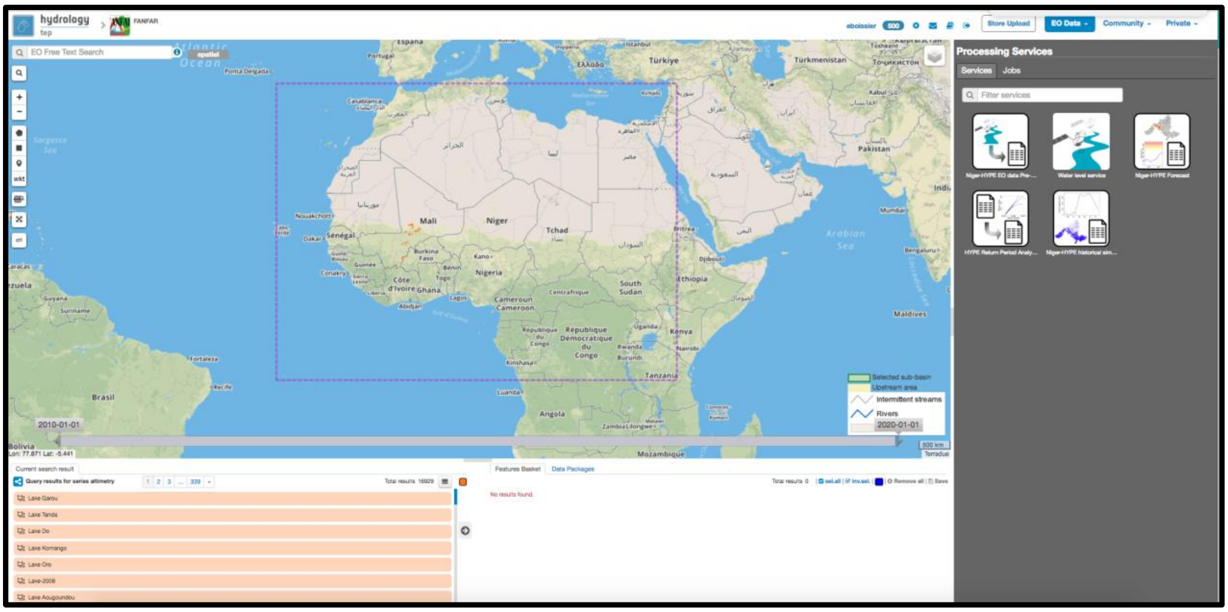


Figure 4: FANFAR forecast production system processing services

4.2 Input data preparation (from water level service)

Water level services provide water level time series based on altimetry for lakes and rivers.

4.2.1 View altimetry tracks on the Geobrowser

Optionally, nominal tracks from Jason, Altika, and Sentinel-3 can be activated as layers, in order to perform a visual inspection of the water bodies that are crossed by tracks, in other words, where altimetric measurements could be retrieved from satellite data.

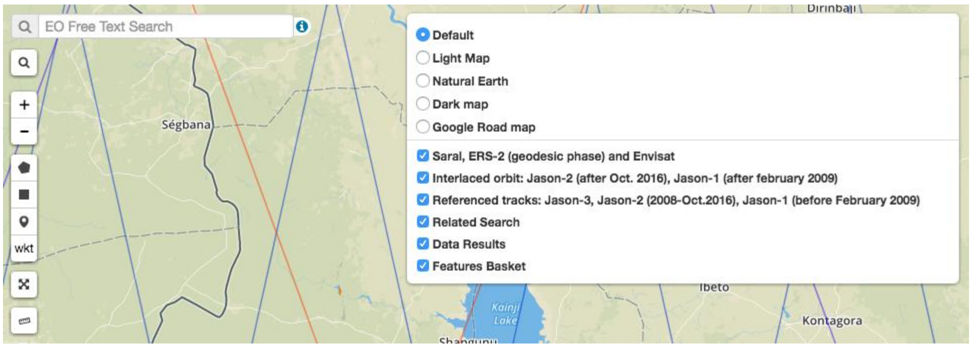


Figure 5: Water level service - altimetry tracks

To discover the actual tracks from each mission, that cross a specific water body during a certain period of time, define an area of interest with the spatial filter tool.

Click on the water body (*EO Data/Altimetry data from Lakes and Rivers* must be selected), in the related search menu, select 'Altimetry over feature'

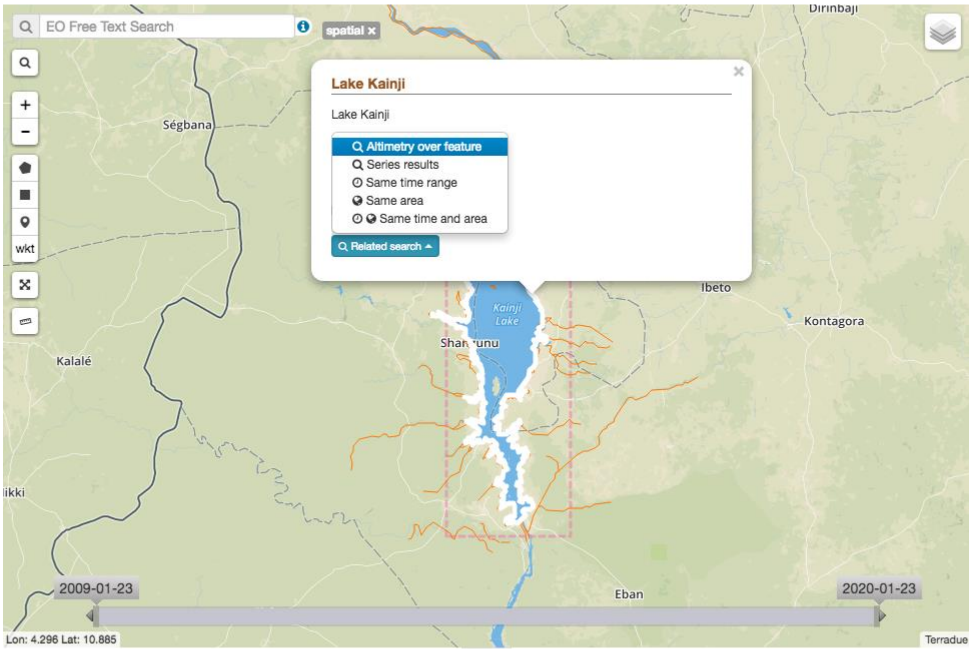


Figure 6: Water level service - data selection

4.2.2 Execute the water level service

In the geobrowser, define your area of interest drawing a rectangle with the spatial filter tool, click on the water body and select 'Altimetry over feature' from the related search menu. Apply your desired temporal filter using the time bar. Resulting tracks and existing water bodies will be displayed in the query results (paginated).

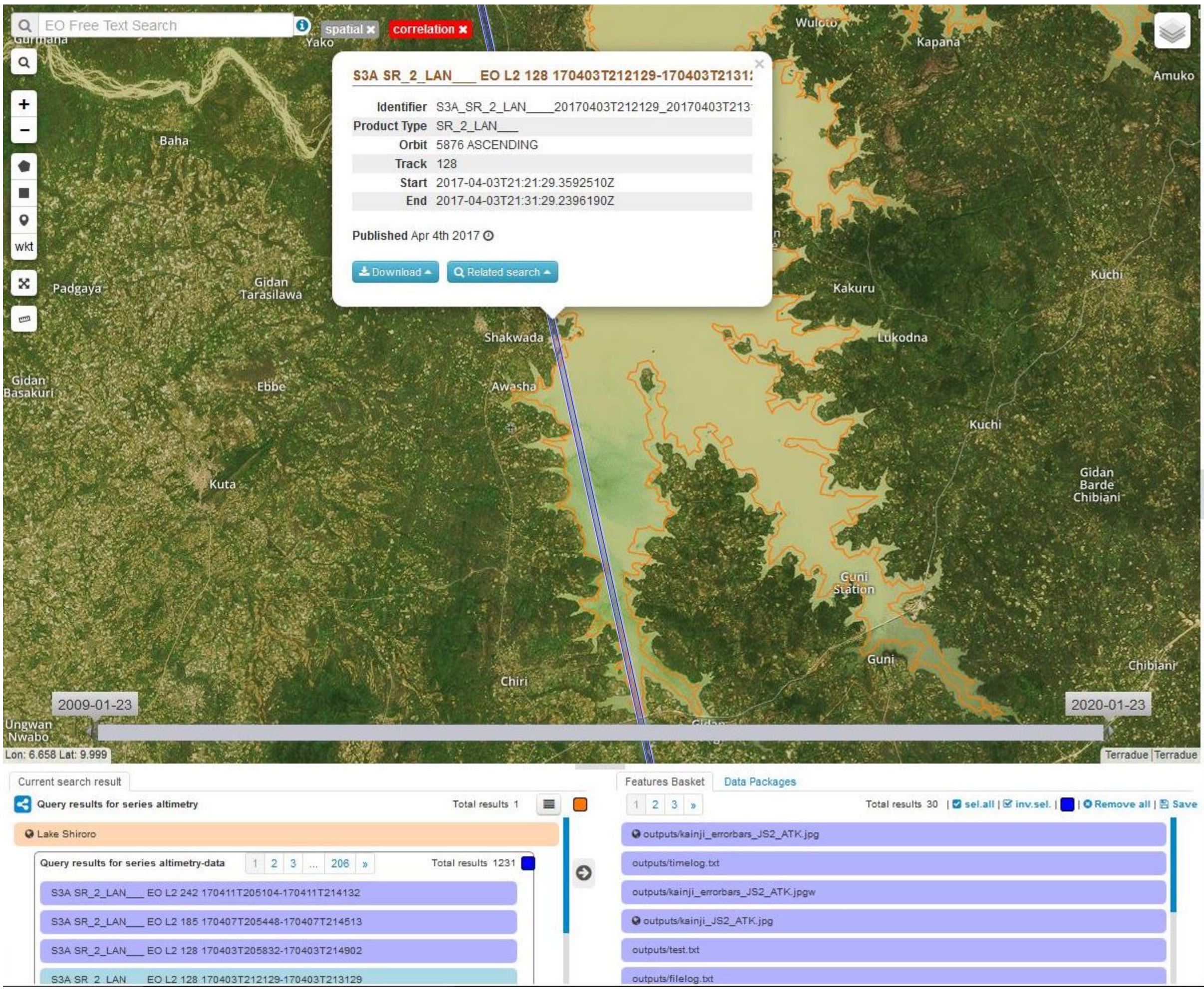
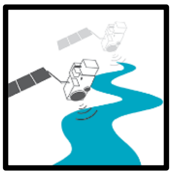


Figure 7: Water level service – AoI definition

1. Open the Water level service 

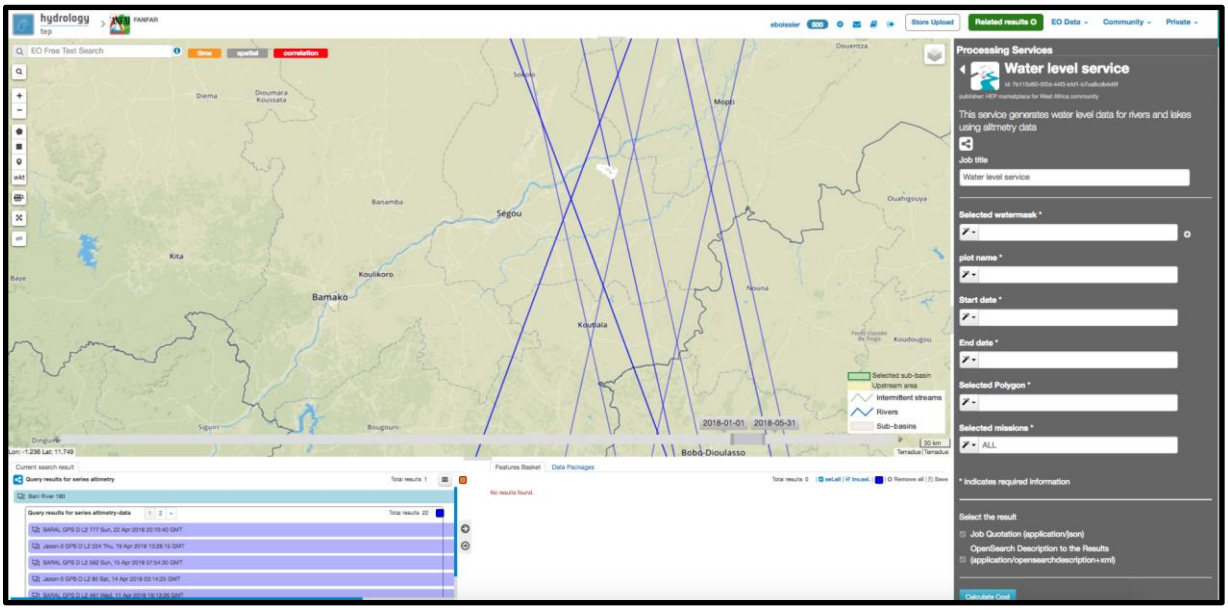


Figure 8: Water level service user interface

2. Set the different inputs
    a. **Job title:** name your water level job as it will appear in your jobs list.
    b. **plot name:** name your plot as it will appear in the resulting time series plot from your job.
    c. **Start date/End date:** use the geobrowser clipboard button to select a start date and an end date according to the temporal filter you have applied (you can also type the dates).

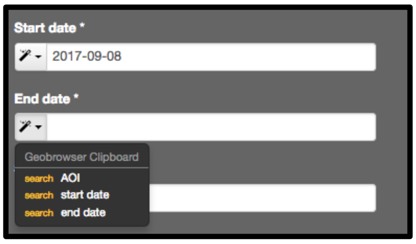


Figure 9: Start/End date selection

    d. **Selected Polygon:** use the geobrowser clipboard button to select the **AOI** you want to process (you can also type a valid WKT).
    e. **Selected missions:** select a mission: **S3**, **ATK**, **JS**. Use **ALL** as default if you do not have a preference for a specific mission.
    f. **Selected watermask:** in the geobrowser, select the water body of interest, you will identify it in the results list as it will be shown in blue, drag and drop the water body from the query results lists into the field. More than one water mask can be processed within the same job, drag and drop all the water masks you wish to process.

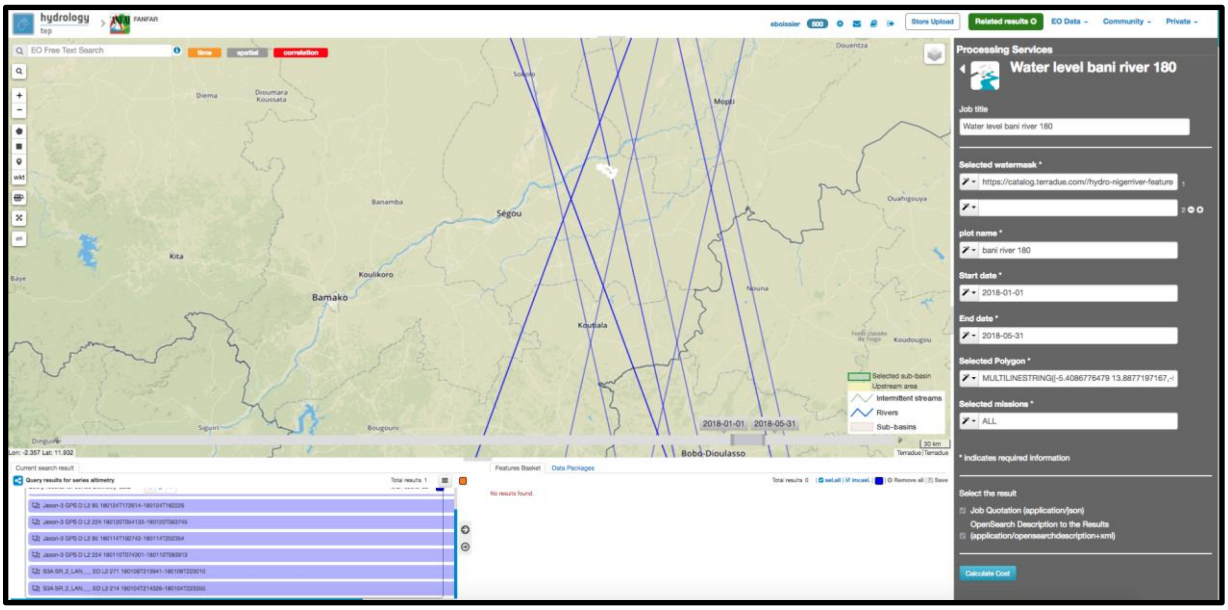


Figure 10: Water mask selection

3. Click on **Calculate Cost** and then **Run job**
4. The job will start running

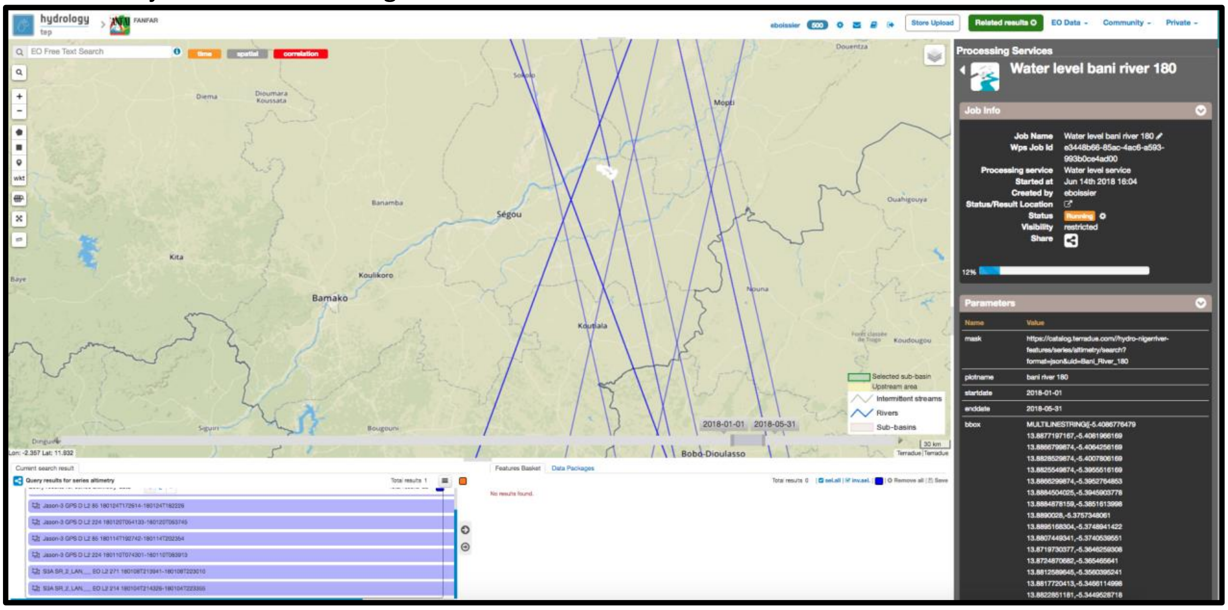


Figure 11: Water level service execution

5. Wait for the job to be finished
6. You can view the results on the map by clicking on **Show results**

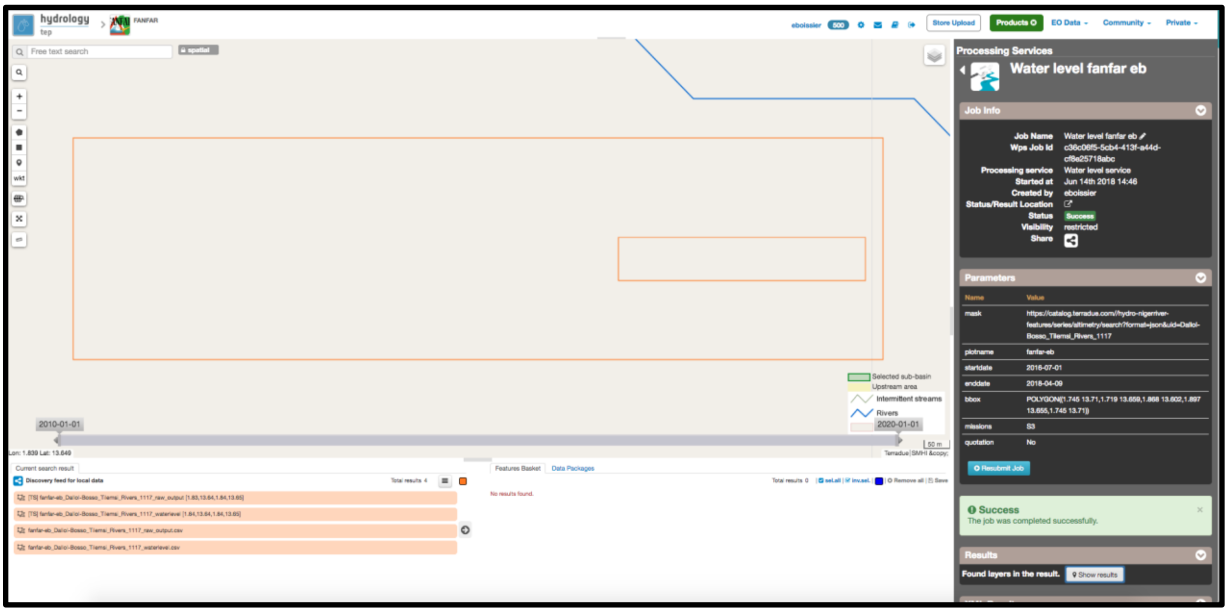


Figure 12: Water level service - show results

7. The list of results appears on the bottom left of the map, and should contain some time series **[TS] <filename>**
8. Drag and drop the results into your **Feature basket** to be able to use them later (optionally save as a new data package if you wish to reuse it at a later stage)

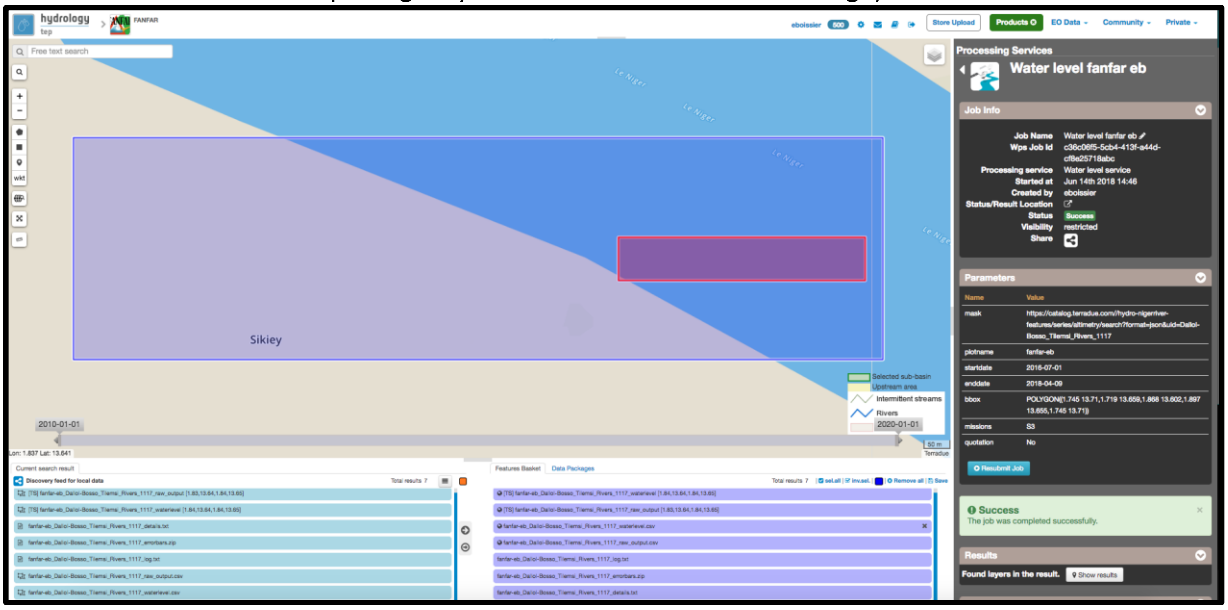


Figure 13: Water level service - save results

4.3 Niger-HYPE EO data pre-processing

The purpose of the service is to transform data sets generated by the H-TEP EO data services to the time-series format required by the hydrological model. This includes both spatial and temporal aggregation of the EO data sets. Currently, the service is configured to do temporal aggregation of data from the Water Level service, to provide time series with data representative for a selected

Niger-HYPE sub-basin. The output is a text file in the specific Xobs text format required to be assimilated in the HYPE model (see further on the HYPE wiki pages).

4.3.1 Execute the Niger-HYPE EO data pre-processing service

1. Open the Niger-HYPE EO data pre-processing service 

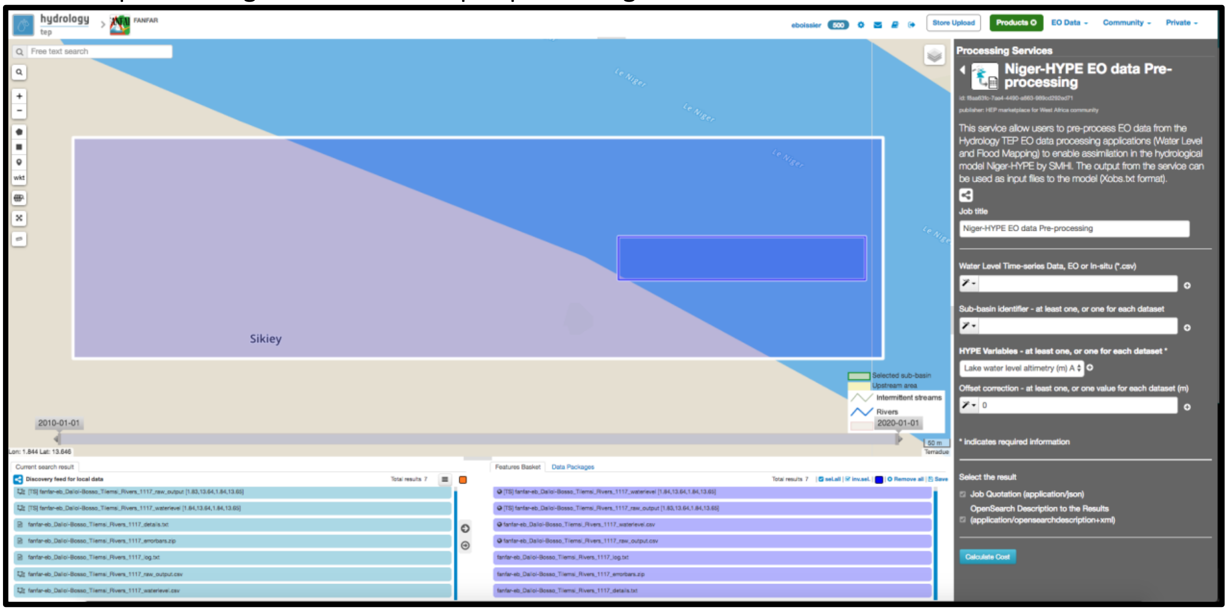


Figure 14: Niger-HYPE EO data pre-processing - service interface

2. Set the different inputs
   a. **Job title:** name your job as it will appear in your jobs list.
   b. **Water Level Time-series Data, EO or In-situ (*.csv):** drag the Water level *.csv files created from Feature basket into the input field.
   c. **Sub-basin identifier:** set the sub-basin identifier corresponding to the EO data set in the input field.
   d. **HYPE Variables:** select the HYPE model variable corresponding to the EO data in the input field:
      i. Lake water level altimetry (m) AOWL for altimetry data, or
      ii. Lake water level in-situ (m) WSTR for in-situ data
   e. **Offset correction:** set the vertical Offset correction of the EO (or in-situ) data set:
      i. This parameter is useful for reducing the mean bias between the model and the observations - which is important if the data is intended for assimilation in the model. The data assimilation method can only adjust the bias to a certain extent, and is more efficient to adjust temporal errors if the mean bias is small.

ii. The offset is an additive correction of the EO data, and can be a negative (EO data is corrected downwards) or positive number (EO data is corrected upwards).

iii. It is recommended to make an initial comparison with the simulated lake water level using offset=0, and then redo the pre-processing with a correction factor estimated from this initial comparison before assimilating the EO data in the model.

3. Click on **Calculate Cost** and then **Run job**
4. The job will start running

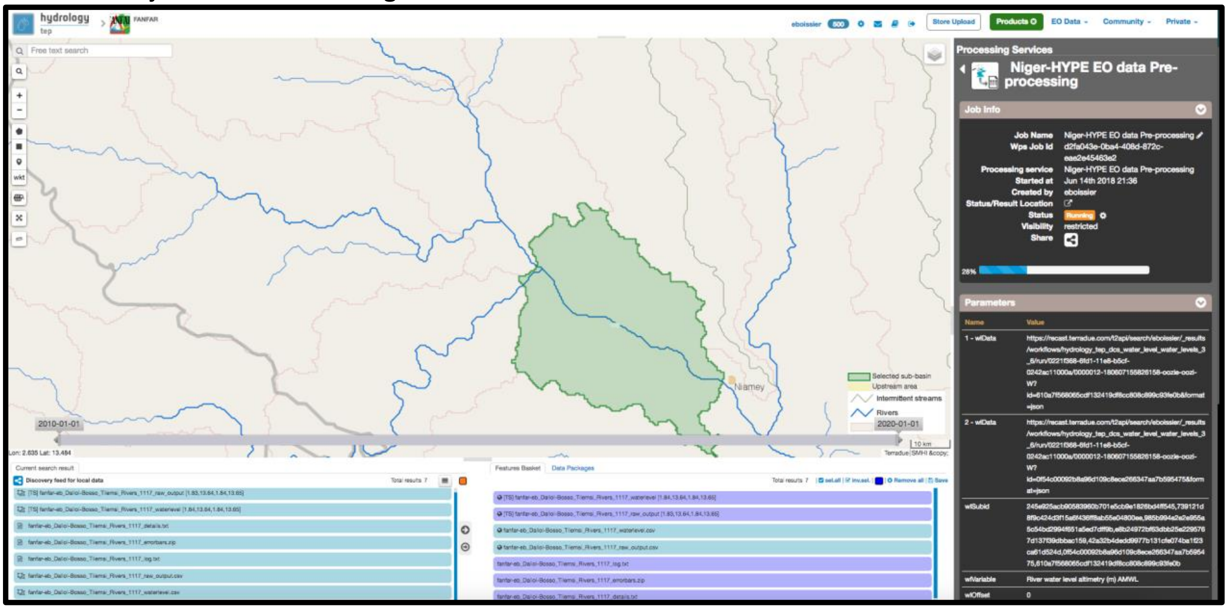


Figure 15: Niger-HYPE EO data pre-processing - service execution

5. Wait for the job to be finished
6. You can view the results on the map by clicking on **Show results**
7. The list of results appears on the bottom left of the map
8. Clear your **Feature basket**, and Drag and drop the results into **it** to be able to use them later (optionally save as a new data package if you wish to reuse it at a later stage)

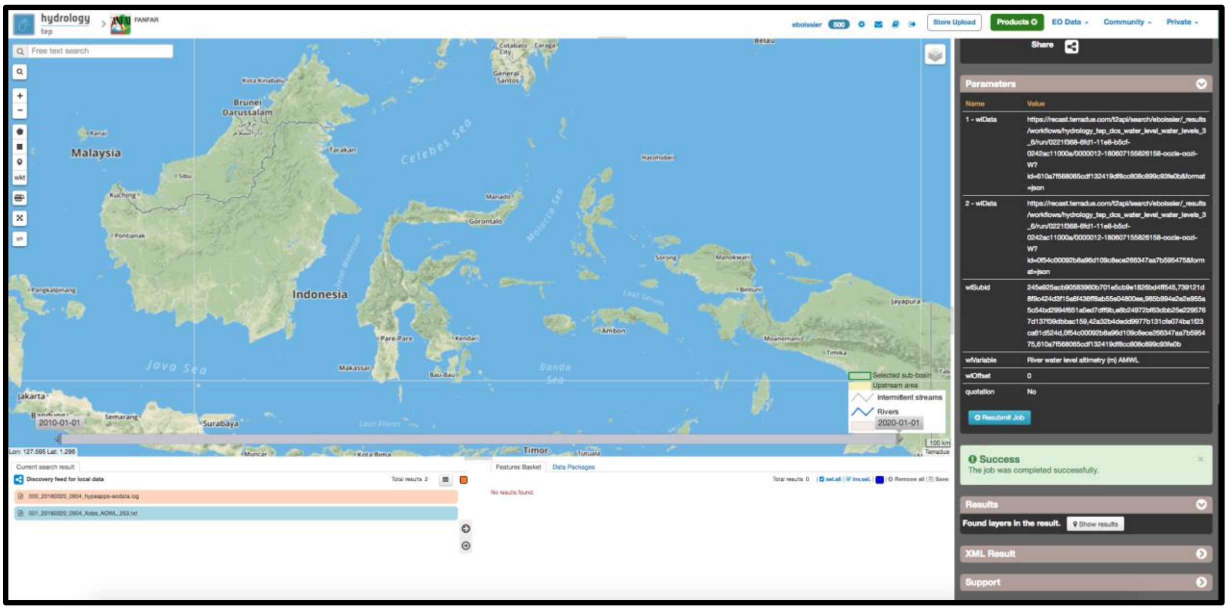
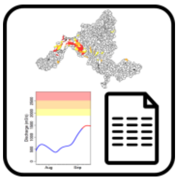


Figure 16: Niger-HYPE EO data pre-processing - service results

4.4 Niger-HYPE forecast

The forecast service always makes two simulations - first a 3-month warm-up simulation (the hindcast) ending on the day before the start of the 10-day forecast simulation (the forecast). The outputs from hindcast and forecast simulations are published separately in the application results. The first day of the forecast simulation is called the forecast issue date, and is one of the input parameters to the applications. Data is saved in the system to enable forecasts for issue dates between 2016-06-01 until the day before the current date. The application may automatically adjust the forecast issue date to an earlier date if the update of the meteorological forcing data for some reason is lagging behind (update is usually made around noon every day).

4.4.1 Execute the Niger-HYPE EO forecast service

1. Open the Niger-HYPE forecast service 

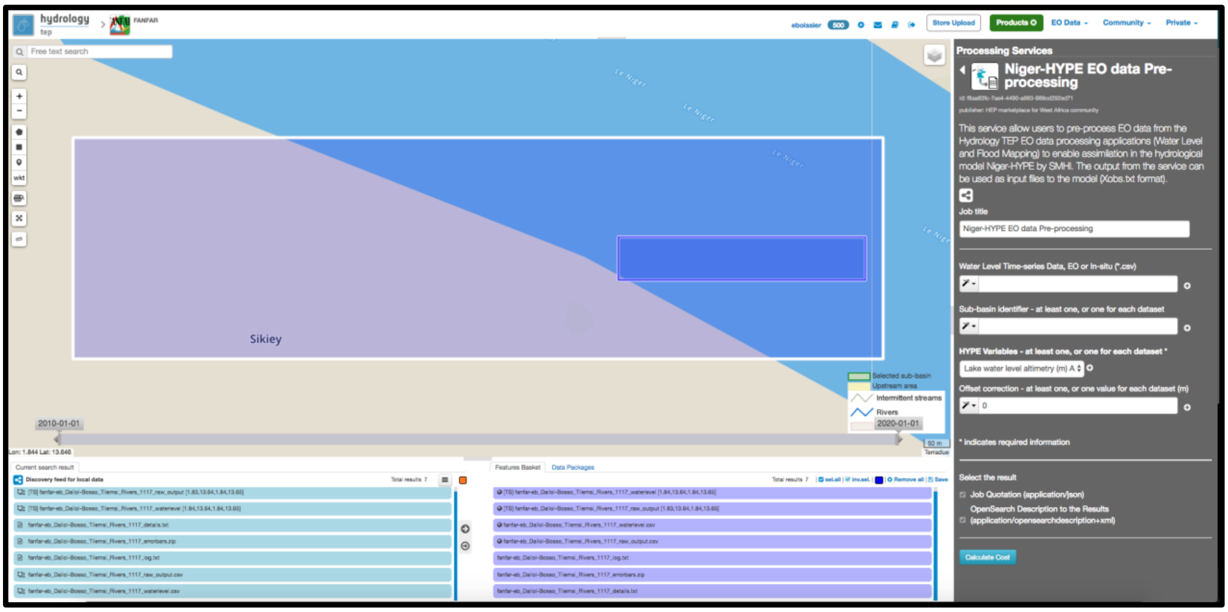


Figure 17: Niger-HYPE EO forecast – service interface

2. Set the different inputs
    a. **Job title:** name your job as it will appear in your jobs list.
    b. **Forecast issue date:** set the start of the 10 day forecast simulation in the input field with format yyyy-MM-dd.
    c. **Output variables:** select outputs from the fixed list of available simulated variables:
        i. The output variable list is build up of three components on the format Variable Name (unit) [4-letter code used internally in the HYPE model]
        ii. Click the + sign to enable selection of more variables.
    d. **Output sub-basins:** select output locations (sub-basins in the Niger-HYPE model) from a fixed list of named locations:
        i. The output location list includes all sub-basins in the Niger-HYPE model where there is a discharge station (Qstn) or a outlet lake (Lake). The name of the discharge station/lake and the sub-basin identifier are included in the drop down list.
    e. **Output sub-basins (SUBID):** *(optional)* select additional output locations that are not listed in the Output sub-basins list by entering the sub-basin identifier in the input field:
        i. The sub-basin identifier can be found for any basin in the model by clicking on the basin in the map browser. It is also possible to set the sub-basin identifier in the input field by using the magic tool in the left side of the input field and select “subbasin id”

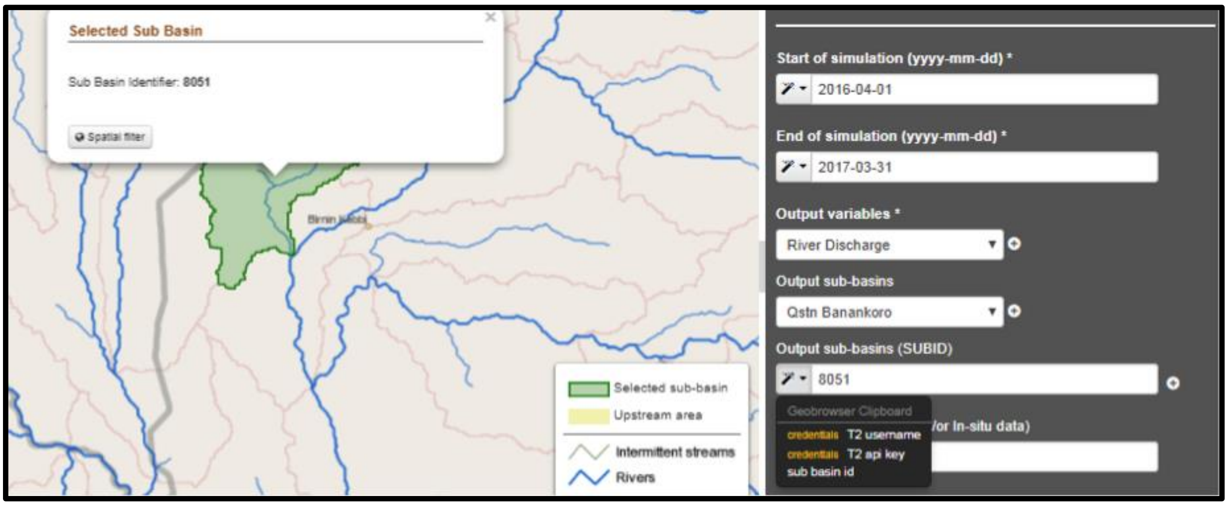


Figure 18: Niger-HYPE EO forecast – sub-basin selection

f. **Return Period Warning Level File:** *(optional)* enter a custom made file for the calculation of river discharge forecast warning levels (or leave **"default"** to use the default data embedded in the application, representing 2,5, and 30 years return period for warning levels 1, 2, and 3, respectively, based on a historical simulation 1979-2013).

g. **Xobs-file with pre-processed EO or in-situ data**: *(optional)* add a so-called Xobs-file to the model simulation by dragging an available data set from your Features basket:
   i. An example can be found in a public data package called Niger-HYPE RPM check data set, which includes a data set called Xobs-eodata.txt with a time-series of lake water level from Lake Shiroro generated with the Water Level application.

h. **Data assimilation:** (*optional)* switch Data Assimilation on or off in the input field:
   i. Switching assimilation on will trigger an Ensemble Kalman filter assimilation run, where the data inserted in the previous step is used to adjust the model simulation towards the assimilated data.

i. **Assimilated observations:** *(mandatory of Assimilation is On, otherwise optional)* select which variables to be assimilated in the input field. Currently it is only possible to select one of two options:
   i. Lake Water Level Altimetry AOWL WCOM (altimetry based lake water level is assimilated if a variable called AOWL is present in the Xobs-files entered in the previous input field)
   ii. Openloop ensemble simulation without assimilation OPEN LOOP (the same type of ensemble simulation is generated as for an assimilation run, however no observations are assimilated).

Please note that the ensemble simulations generated when switching assimilation on requires much more processing time (at least 10 times more) compared to assimilation off. This is due to two reasons:

- The Ensemble Kalman filter method is based on ensemble simulations. Currently, the Niger-HYPE application is configured to include 10 members in the model ensemble (which is actually already a rather small number, 50 or more would be better).

- The model ensemble must also have enough variability to be able to adjust to the observations in a realistic way. Currently, the only method used here to generate ensemble spread is to add random perturbations to the meteorological forcing data (precipitation and temperature). Thus, it becomes important to include at least one rainy season in a warm-up period before the assimilation period. Consequently, the selected start of the simulation period is automatically adjusted to meet this criteria.

It is advisable to first make a simulation without assimilation to check results, and also the bias between simulated variable (WCOM in the example) and the observations to assimilate (AOWL in the example) and possibly correct the offset input parameter in the EO data pre-processing.

3. Click on **Calculate Cost** and then **Run job**
4. The job will start running

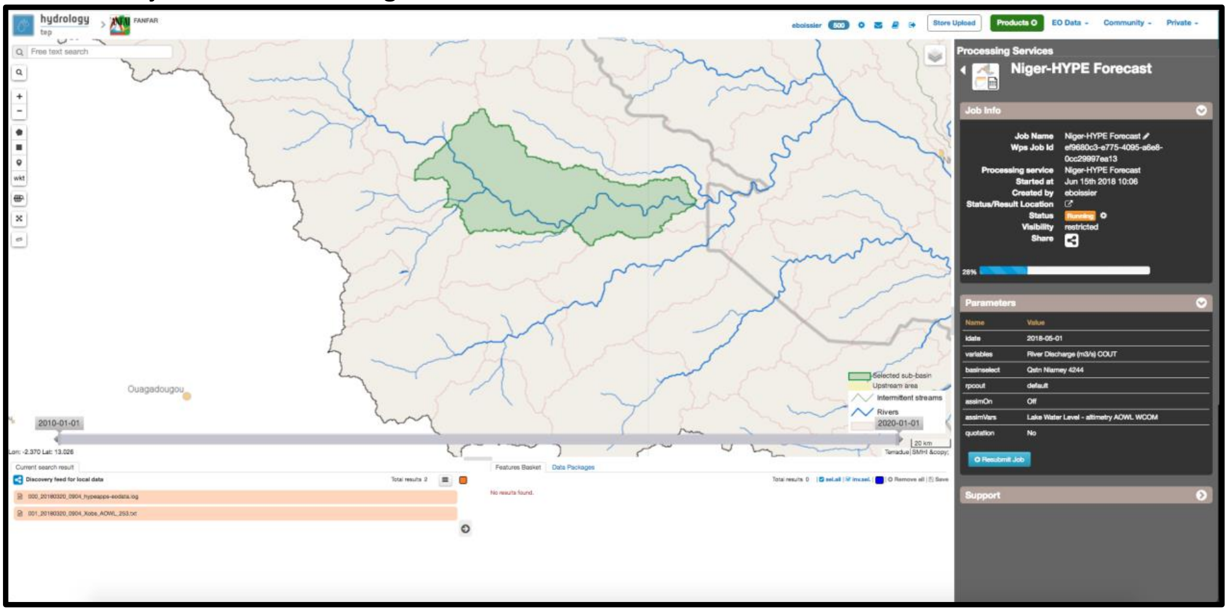


Figure 19: Niger-HYPE EO forecast – service execution

5. Wait for the job to be finished
6. You can view the results on the map by clicking on **Show results**
7. The list of results appears on the bottom left of the map

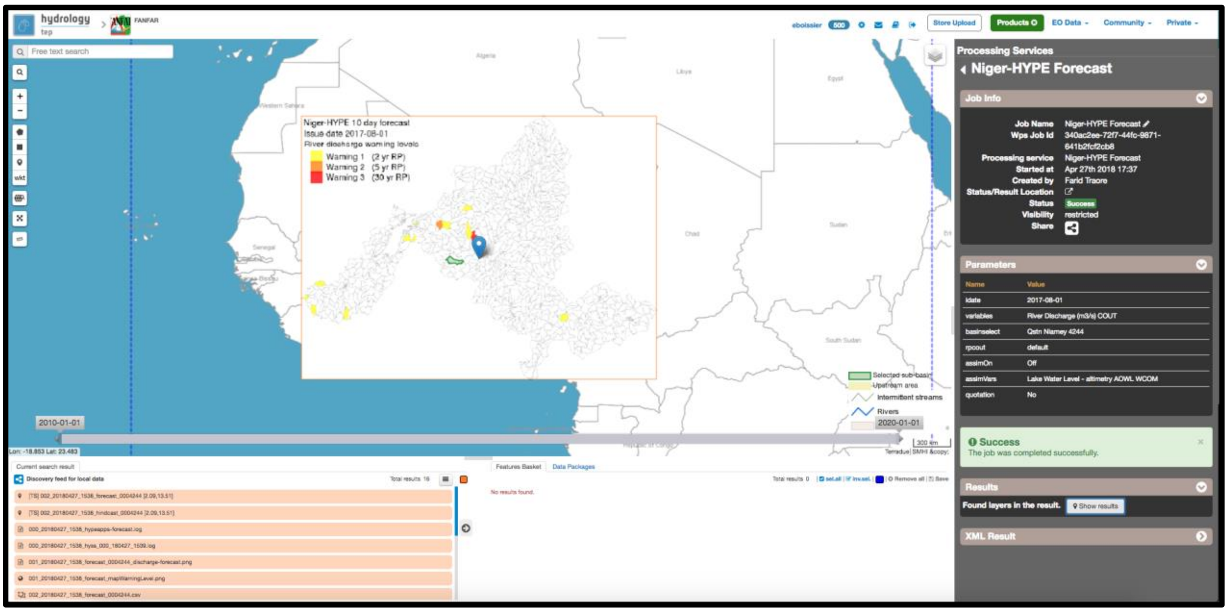


Figure 20: Niger-HYPE EO forecast – service results

8. The results from the processing service are described in the table below. The prefix numbers are used as a trick to order the output files in a certain order. Results are sorted in subfolders called forecast and hindcast, respectively.

Table 2 Hype forecast results

Output files	**Description**
001_[subid]_[name].png 001_[subid]_[name].pngw	Quicklooks (PNG format with associated word file PNGW for visualization in the map browser) with time-series plots of simulated variable [name] for sub-basin [subid], where [name] corresponds to the 4-letter code for HYPE model variables and [subid] to the Niger-HYPE sub-basin identifier.
001_map[name].png 001_map[name].pngw	Quicklook (PNG format) with maps of variable [name] with mean simulated variable during the simulation period, where [name] corresponds to the 4-letter code for HYPE model variables.
002_[subid].csv 003_[subid].txt	Text files with time series output for selected sub-basins: - *.csv provide data in the H-TEP standard csv format *.txt provide data in the HYPE basin-output format. - Each text file includes data for the entire simulation period (daily values) for all selected output variables for one sub-basin.
004_map[NAME].txt	Text file with average (full simulation period)

	simulated value for a selected variable specified by the file name (4-letter variable names, see HYPE wiki pages). Each row represents one subbasin of the model.
005_time[NAME].txt	Text file with average (full simulation period) simulated value for a selected variable specified by the file name (4-letter variable names, see HYPE wiki pages). Each row represents one subbasin of the model.
006_simass.txt 006_subass.txt	Text file with assessment of the agreement between simulated and observed data (depending on which observed variables added to the simulation through the Xobs-files).
009_hyss_000_YYMMDD_hhmm.log	Log-file for the HYPE model simulation
hypeapps-historical-logfile.txt	Log-file from the processing service.
001_[subid]_discharge-forecast.png 001_[subid]_discharge-forecast.pngw	River discharge forecast time-series plots (png quicklooks) for the selected output sub-basins [subid], with the return-period levels based warning levels plotted in the background. The quicklooks are visualized in the map browser located with the lower left corner in the centre-coordinate of the corresponding basin.
001_mapWarningLevel.png 001_mapWarningLevel.pngw	River discharge forecast warning level maps (png quicklooks), showing the maximum forecasted river discharge warning level in each sub-basin of the Niger-HYPE model. The quicklook is displayed in the map browser scaled to the current map scale and centered on the Niger River basin.
004_mapWarningLevel.txt	Maximum forecasted river discharge warning level in the HYPE map output text format (same format as 004_map[NAME].txt)
hypeapps-forecast-logfile.txt	Log-file from the processing service.

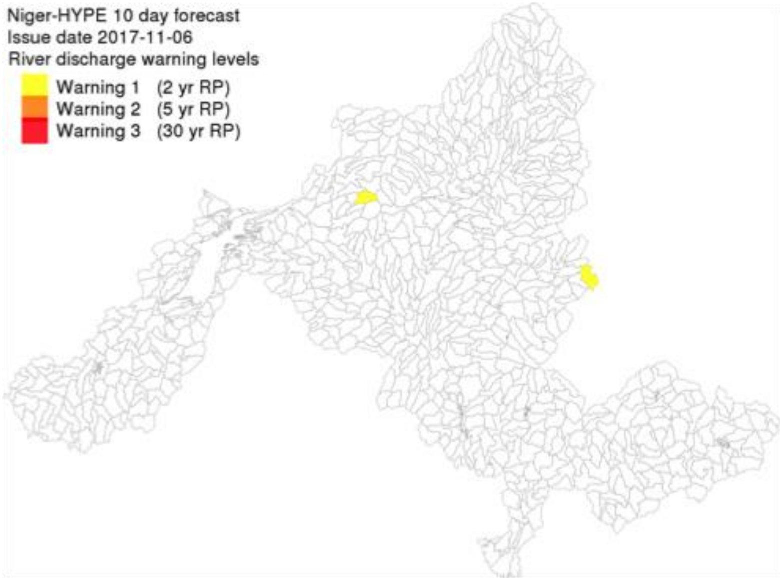


Figure 21 Example of River discharge forecast warning level map output (left) and River discharge forecast time-series output (right)

4.5 Results visualization

- Image files can be visualized directly on the map. WMS layers are automatically created from the image files in order to give a best visualization.
- *.csv files can be visualized as time series by clicking on the "get Time Series" button on the associated [TS] file. The H-TEP comes with a dedicated time series visualization tool.

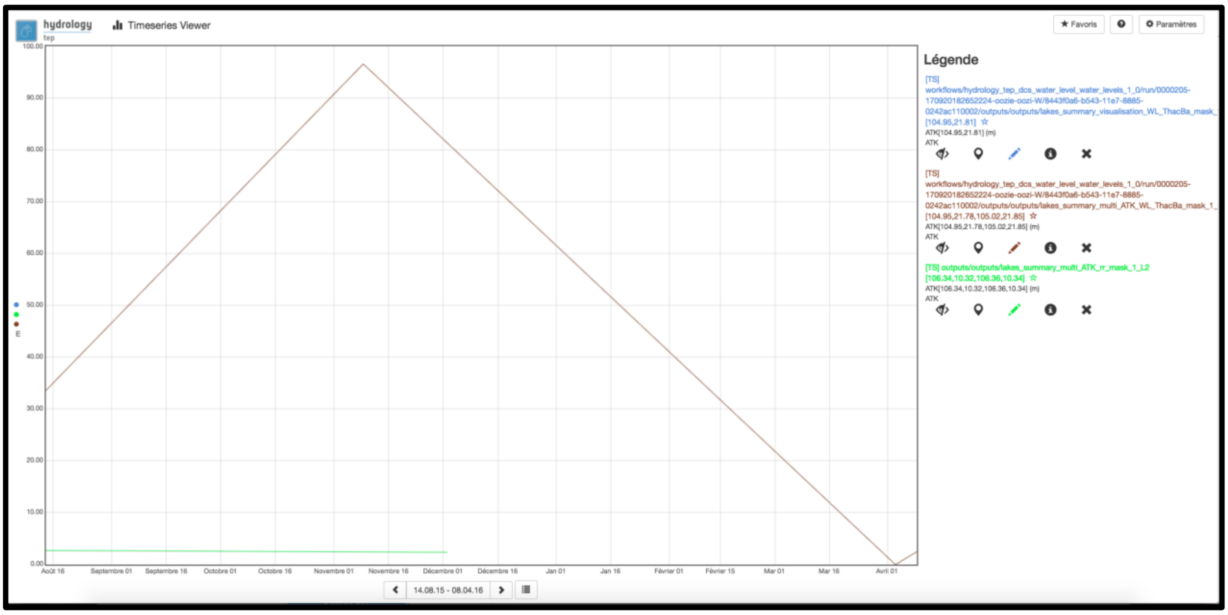


Figure 22: Niger-HYPE EO forecast - result visualization

- all files can be downloaded, by clicking on the "Download" button of the metadata popup associated to the file (appears when double clicking on the item representing the file in the results list).

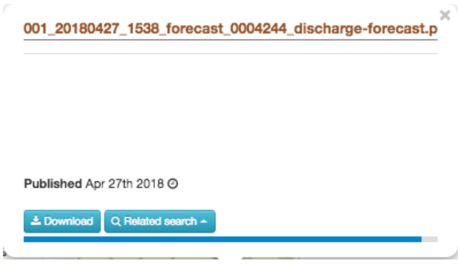


Figure 23: Niger-HYPE EO forecast - results download

4.6 Results storage and distribution

Results files are stored in the user's H-TEP private and persistent storage, in a dedicated repository created for this job run, e.g. *https://store.terradue.com/<username>/_results/workflows/hydrology_tep_hypeapps_forecast_workflow_1_5/run/e620fe86-4a30-11e8-a5d8-0242ac11000a/0000082-180319142931691-oozie-oozi-W/001_20180427_1538_forecast_0004244_discharge-forecast.png*

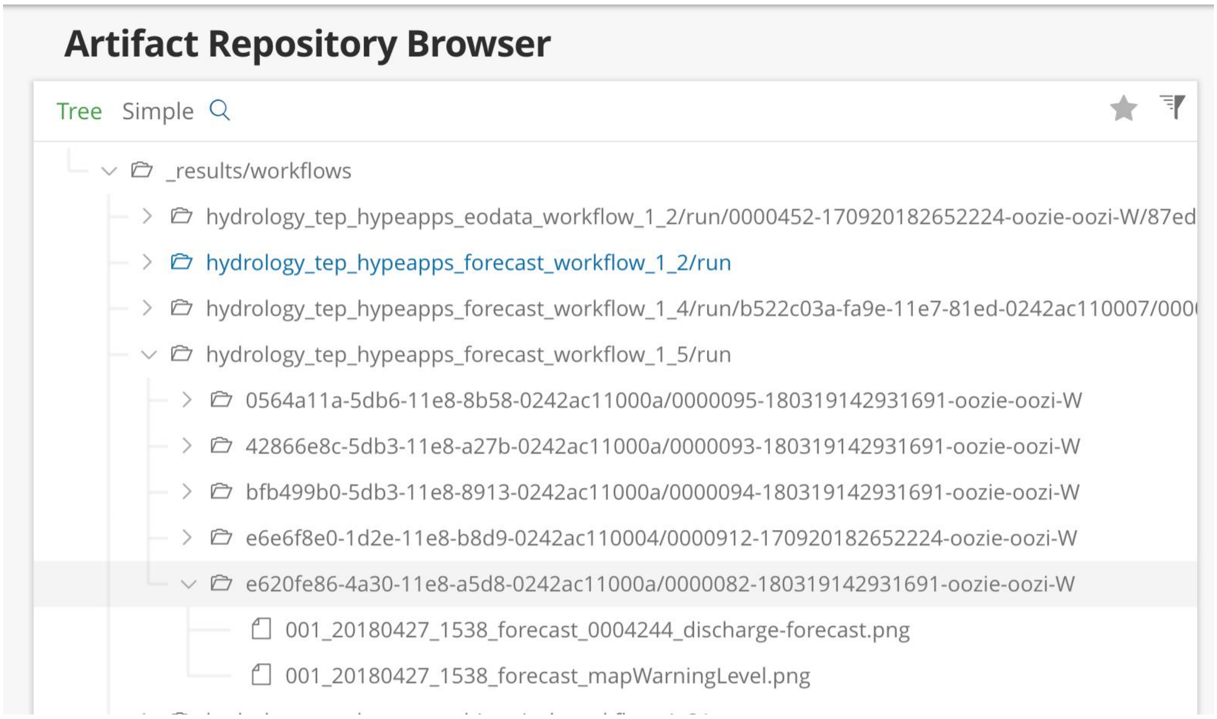


Figure 24 H-TEP storage containing job results, grouped by user, service and runs

Each of these folders can be associated to a distribution channel into which data is published periodically.

In the case of the systematic processing, a synchronization agent is configured at the root of the result folders by settings predefined properties such as the target remote system(s), the transfer protocol(s), the rules for the naming of the remote directories etc. as in the following example:

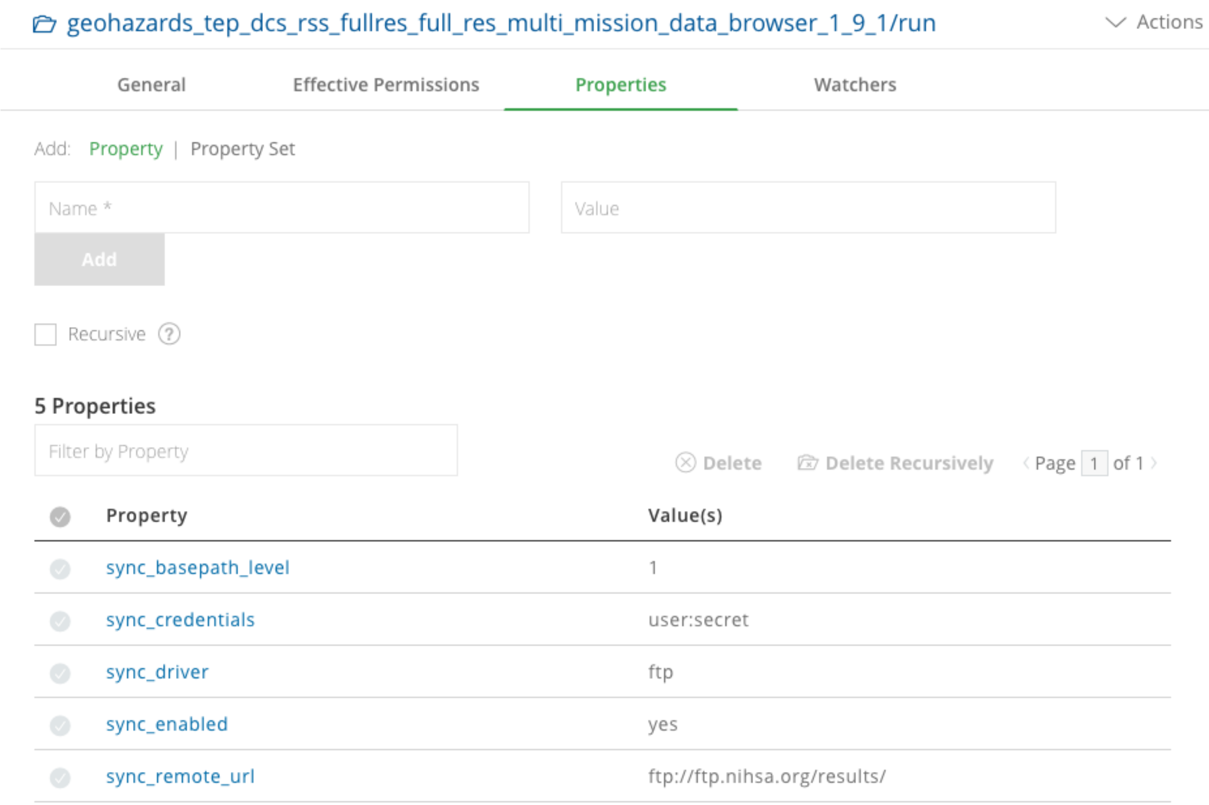


Figure 25 Distribution channel configuration example

That folder is therefore exported to the remote system(s) following the indicated protocols as soon as new produced data is added in that folder.