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 3.1: Data Access Interface

FINAL VERSION

Due date of deliverable: 28/02/2018
Date accepted for publication/submission to EU: 11/03/2019
Actual submission date: 11/03/2019

Title	Data Access Interface
Creator	Aytor Naranjo
Editor	Aytor Naranjo
Brief Description	This document is a live document that will be updated through subsequent iterations but intends to provide a preliminary version about the data access interface at FANFAR.
Publisher	FANFAR Consortium
Contributors	Aytor Naranjo, Laia Romero, Guillermo Grau, Aytor Naranjo, Jafet Andersson, Bernard Minoungou, Mohamed Hamatan, Addie Shuaib Olorunoje
Type (Deliverable/Milestone)	Deliverable
Format	Report
Creation date	17/01/2017
Version number	v2
Version date	27/02/2017
Last modified by	Aytor Naranjo
Rights	Copyright “FANFAR Consortium”. During the drafting process, access is generally limited to the FANFAR Partners.
Audience	☐ internal X 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	11/03/2019

Version	**Date**	**Modified by**	**Comments**
v1	18/01/2019	Aytor Naranjo	
v2	27/02/2019	Aytor Naranjo	

Contents

List of Figures

List of Tables

Executive summary

This document documents and explains the data access interface for the operational forecasting and alert system to run systematically on the H-TEP environment. In this context a "data access interface" refers to an input data set to the hydrological forecasting model and the way this data is made available to the H-TEP environment.

The different input data streams to produce the hydrological forecasts and alerts on H-TEP consists of meteorological hindcast and forecast datasets, satellite altimetry-based water level, water bodies and in-situ observations (hydrometric observations of water level and streamflow).

At present, in-situ observations are provided to AGRHYMET by local meteorological agencies from a network of 20 hydrometric stations, currently from Guinea, Mali, Niger, Senegal and Nigeria (including NIHSA). Data from hydrometric stations contains river flows, water discharge and rating curves. Four stations collect data automatically on hourly basis. Readings from other stations are collected manually and sent once a month to AGRHYMET by email. AGRHYMET feeds all new in-situ data in their database. H-TEP repository can access to AGRHYMET's in-situ database via FTP, however, AGRHYMETs FTP is not yet updated operationally. So the data stream is not yet available to H-TEP & FANFAR processing. This document also includes how the data is controlled to ensure post-processing will yield accurate forecasts.

The water level from isardSAT is based on altimetry data from the satellites Sentinel 3, Jason-2 and SARAL-Altika. Input for the water level application, consists of a water mask of the water body or river of interest as well as altimetry data.

The meteorological data from SMHI consists of meteorological forecasts (ECMWF deterministic 1-10 day) and a mosaic product representing past meteorological conditions (HYDROGFD2.0). This is made available through FTP and is transferred to H-TEP every day. The AGRHYMET data consist of bias factors that will be accessible by SMHI via an FTP site, for the correction of HYDROGFD2.0 data.

The combination of these data streams into the FANFAR repository allows the daily production of hydrological hindcasts and 10-day forecasts to generate flood risk information.

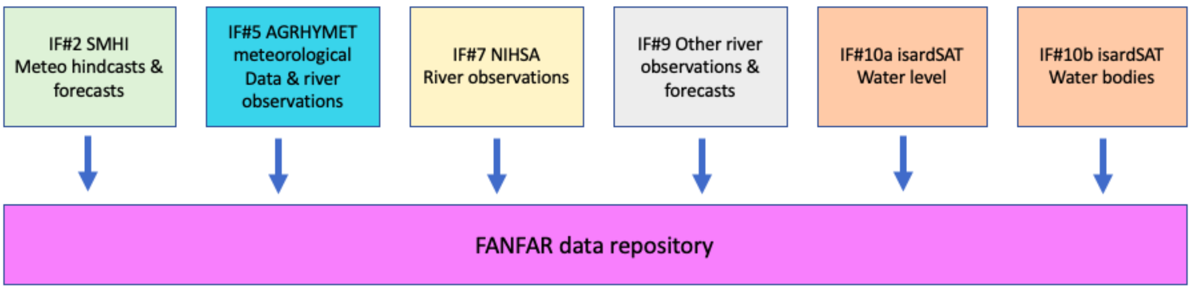

Figure 1. Illustration of the different data access Input interface sequence diagram

1. Introduction

This document is the Interface Control Document of the Operational Hydrological Forecasting and Alert system. It is aimed at documenting and explaining the data interface requirements for the FANFAR flood forecasting and alert system to run at an operational level.

Operational forecasting heavily relies on scheduled, automated execution and monitoring of the data flows and processing tasks on an ICT environment that is always up and running. In FANFAR, this is provided by the H-TEP environment.

This deliverable concentrates on the input data that is fed on daily basis to the H-TEP to produce a representation of future hydrological conditions in West Africa.

The core of the FANFAR system is a hydrological model, whose main function is to predict 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 use the HYPE model applied to the Niger River basin and to the entire West African domain. Two simulations are usually required to make a forecast with a hydrological model:
1. Hindcast: A simulation of a historic spin-up period up until the day before the forecast, producing a model state that represents the present hydrological conditions at the start of the forecast period (the initial state, t0). The core input for this simulation is meteorological ‘analysis’ data (i.e. fusion of meteorological observations and models representing historic conditions). Additional observations of the hydrological conditions (e.g. in-situ river flow observations or remotely sensed water extent) can also be used to improve the initial state through data assimilation.
2. Forecast: A simulation for the forecast period relying on the initial state and meteorological forecast data about future weather dynamics. The output of the forecast simulation is a representation of future hydrological conditions.

1.1 Interface definition

In this document, an interface is considered as the mean for data components to communicate with each other. In that sense, an interface is specified by the following elements:

- The information that is exchanged, expressed through a data format. The data format describes the structure of the data.
- The protocol, i.e. the set of rules that determine how this information is exchanged.

2. Overview of FANFAR system and data access interfaces

An overview of the FANFAR operational forecasting system is provided in Figure 2. The system consists of data sets, data flows, processing algorithms, distribution channels, scheduling, and underlying data stores and processing ICT infrastructure. This deliverable focusses on describing the inputs to the hydrological model HYPE in H-TEP, specifically number 2, 5, 7, 9 and 10.

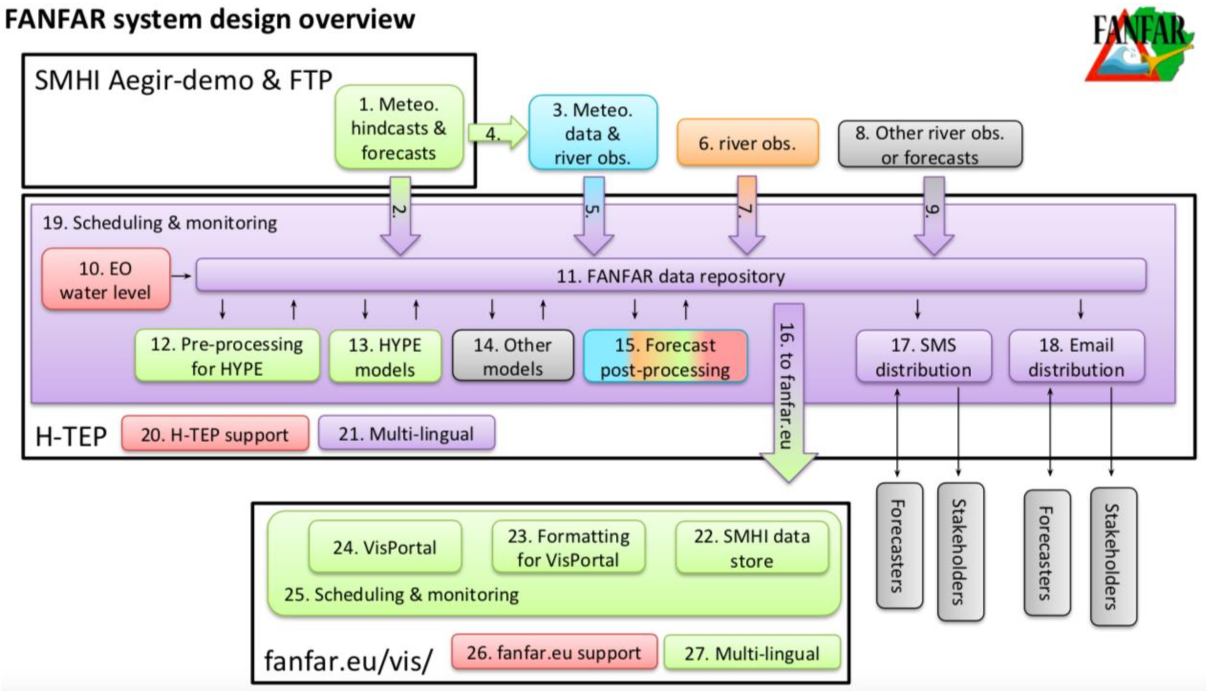

Figure 2. FANFAR system design

2.1 Overview of input data access interfaces

This deliverable covers the input data needed to derive useful forecast and alert information. The list of Internal System interfaces is provided in table 1 along with details on frequency, protocol and format.

Table 1.Internal system interface

ID	From	To	Name	**Frequency(** *Periodic/ event-based)*	**Protocol** ***(****FTP, HTTP, etc.)*	**Format** ***(****XML/JSON, etc.)*	Comments
IF2	SMHI	H-TEP Data repository	Meteorological hindcast & forecast data	Daily	FTP	netCDF	Pull-based
IF5	AGRHYMET	H-TEP Data repository	Water level/discharge (from manual stations)	Monthly	FTP	XLSX	Pull-based
			Rating curves	On demand	FTP	XLSX	Pull-based
			Water level/ Discharge (from automatic stations)	Hourly	FTP	TXT	Pull-based
			Meteorological hindcast data	Daily	FTP	netCDF	Pull-based

IF7	NIHSA	AGRHYMET	Water level/discharg e	Daily data collected and transmitted monthly	E-mail	XLSX	Email-based and Text msg from gauge readers
			Rating curves	On demand	E-mail	XLSX	Email-based
IF9	Other hydrological agencies	AGRHYMET	Water level/ Discharge	Monthly	E-mail	XLSX	Email-based
			Rating curves	On demand	E-mail	XLSX	Pull-based
IF10a	isardSAT	H-TEP Data repository	Water level	Daily	OGC WPS	csv	Request response
IF10b	isardSAT	H-TEP Data repository	Water bodies generation tool	On demand	FTP	csv in WKT	On demand

3. Interface descriptions

3.1 Interface#2 Meteorological hindcast and forecast data from SMHI

3.1.1 Introduction

A variety of meteorological datasets can be used to simulate the historical hydrological conditions up until the current time step (the 'hindcast'). To this end SMHI provides the HYDROGFD2.0 dataset (an updated version of the data and method presented by Berg et al 2018, https://www.hydrol-earth-syst-sci.net/22/989/2018/) for West Africa within the FANFAR project (Figure 3).

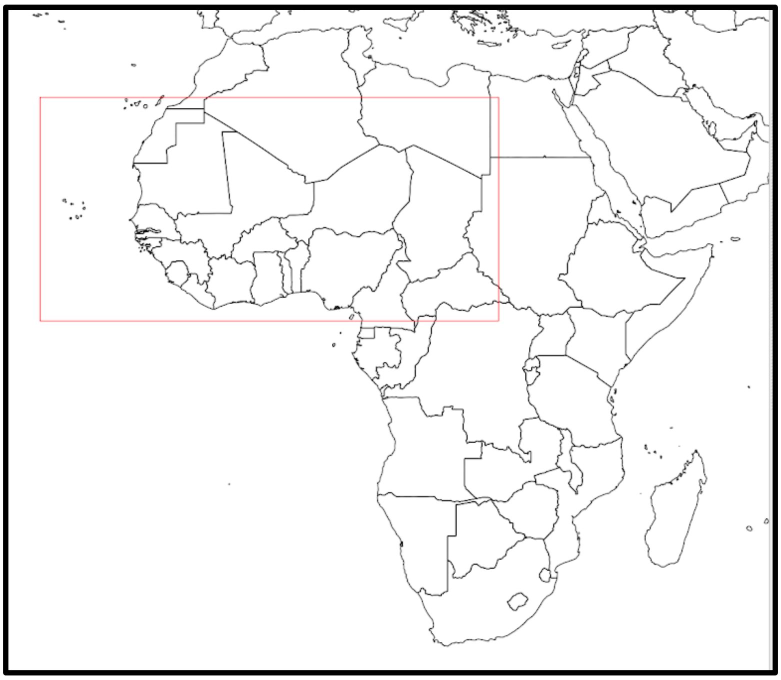

Figure 3. Domain of the meteorological data provided from SMHI (red square).

HYDROGFD2.0 consists of a set of component datasets (HYDROGFDEI, HYDROGFDOD and OD-DAILY) with different levels of adjustments toward observations. The meteorological forecasting dataset consists of deterministic 1-10day simulations by ECMWF. Table 2 provides more details on these data sets. In each dataset the following variables are provided on daily resolution: precipitation sum, mean air temperature, maximum air temperature and minimum air temperature.

Table 2.Overview of meteorological hindcast and forecast data from SMHI. t0 refers to the first day in the current month. tx is the current day, tf is the last day of the forecast period, and 3m referest to three months.

Time period	1979 – $t_{0\text{-}3m}$	$t_{0\text{-}3m}$ - t_0	t_0 - t_x	t_x - t_f
Original meteorological data	ERA-Interim reanalysis	Saved 1-day forecasts, ECMWF deterministic	Saved 1-day forecasts, ECMWF deterministic	Current 1-10 day forecasts, ECMWF deterministic
P observations used to adjust the original meteo. data	CPC-unified	CPC-unified	None	None
T observations used to adjust the original meteo. data	CPC-temp	CPC-temp	None	None
Frequency	Monthly	Monthly	Daily	Daily
Output dataset	HYDROGFDEI	HYDROGFDOD	OD-DAILY	ECOPER

3.1.1.1 Dependencies

Availability of inputs to HYDROGFDEI, HYDROGFDOD, OD-DAILY and ECOPER. Production of these datasets at SMHI’s servers. Storage on an FTP-site accessible by H-TEP.

3.1.1.2 Protocol

The transfer between SMHI and H-TEP is handled by the FTP protocol, in such a way that H-TEP pulls the data from SMHI’s FTP site every day. Subsequently H-TEP ingests the data in the H-TEP data repository with appropriate metadata to enable opensearch queries in order to find the data.

3.1.1.3 Data format

All files are provided in netCDF format, according to the Climate and Forecast metadata convention v 1.6 (http://cfconventions.org/). This allows descriptive metadata to accompany the data (Figure 4).

```
netcdf pr_hydrogfdei_201811_fanfar_SMHI {
dimensions:
        lon = 108 ;
        lat = 52 ;
        time = UNLIMITED ; // (30 currently)
        nb2 = 2 ;
variables:
        float lon(lon) ;
                lon:standard_name = "longitude" ;
                lon:long_name = "longitude" ;
                lon:units = "degrees_east" ;
                lon:axis = "X" ;
        float lat(lat) ;
                lat:standard_name = "latitude" ;
                lat:long_name = "latitude" ;
                lat:units = "degrees_north" ;
                lat:axis = "Y" ;
        double time(time) ;
                time:standard_name = "time" ;
                time:long_name = "time" ;
                time:bounds = "time_bnds" ;
                time:units = "hours since 1949-12-01 00:00:00" ;
                time:calendar = "standard" ;
        double time_bnds(time, nb2) ;
                time_bnds:units = "hours since 1949-12-01 00:00:00" ;
                time_bnds:calendar = "standard" ;
        float factor(lat, lon) ;
        float adjusted_points(lat, lon) ;
        float pr(time, lat, lon) ;
                pr:standard_name = "precipitation_flux" ;
                pr:long_name = "Precipitation" ;
                pr:units = "kg m-2 s-1" ;
                pr:code = 228 ;
                pr:table = 128 ;
                pr:cell_methods = "time: mean" ;
                pr:title = "pr" ;

// global attributes:
                :CDI = "Climate Data Interface version 1.6.0 (http://code.zmaw.de/projects/cdi)" ;
                :history = "Tue Feb 26 10:25:12 2019: cdo sellonlatbox,-28,26,3,29 /data/proj/Fouh/HydroGFD/hydrogfd2.0/data/hydrogfdei/pr/day/pr_GLB-0.5_GFD
day_201811.nc /data/proj/aegir/aegir/temp/fanfar/pr_hydrogfdei_201811_fanfar_SMHI.nc\nCreated = 20190210" ;
                :institution = "Swedish Meteorological and Hydrological Institute" ;
                :Conventions = "CF-1.6" ;
                :product = "observation" ;
                :frequency = "day" ;
                :references = "Berg et al. 2017, Near real-time adjusted reanalysis forcing data for hydrology; https://www.hydrol-earth-syst-sci-discuss.net
                :Version = "GFD2.0" ;
                :Contact = "Peter Berg (peter.berg@smhi.se)" ;
                :institute_id = "SMHI" ;
                :product_id = "obs" ;
                :CDO = "Climate Data Operators version 1.6.0 (http://code.zmaw.de/projects/cdo)" ;
}
```

Figure 4 Example of the header of a HYDROGFD2.0 netCDF file (precipitation from HYDROGFDEI for november 2018). The header provides the metadata, while the data itself is stored according to dimensions in each variable in arrays.

3.1.1.4 Exchange frequency

New files are searched for every day, however some datasets are only updated once per month (Table 2).

3.1.1.5 Data size

The data size varies but range from approximately 200Kb for daily files to 11Mb for monthly files.

3.1.1.6 Notification

The production of the datasets is monitored at SMHI and in case of production failures, a notification is provided to technical staff responsible for the production. At H-TEP, the data retrieval process is planned to be monitored in the FANFAR monitoring tool.

3.1.1.7 Data removal

Global data stored at SMHI. Data for the West African domain is stored on H-TEP. Data is regularly removed from the intermediate storage on SMHI's FTP system.

3.2 Interface #5 River flow and water level observations and meteorological data by AGRHYMET

AGRHYMET collates in-situ observations provided by local meteorological agencies from a network of hydrometric stations, currently from Guinea, Mali, Senegal, Niger and Nigeria. This input data consist of the following:

- Water level/discharge data collected manually from stations
- Water level data obtained from automatic stations
- Rating curves

The metadata details below (Figure 5) show all the current hydrometric stations that FANFAR has access to and which are stored in an Excel file accessible from an FTP. The map illustrating the locations of these stations in shown in Figure 6.

STATION	LONGITUDE	LATITUDE	PAYS	RIVER	BASIN	Id Station	Bassin_Area
FARANAH	-10.75	10.04	GUINEE	Niger	Niger	1171500115	3 180
KANKAN	-9.30	10.38	GUINEE	Milo	Niger	1171501705	9 900
TINKISSO	-10.58	11.23	GUINEE	Tinkisso	Niger	1171502510	
BARO	-9.69	10.61	GUINEE	Niandan	Niger	1171501805	
DIALAKORO	-8.9	11.45	GUINEE	Niger	Niger	1171500110	
MANDIANA	-8.68	10.62	GUINEE	Sankarani	Niger	1171502005	
KOUROUSSA	-9.88	10.65	GUINEE	Tinkisso	Niger	1171500120	
ANSONGO	0.52	15.53	MALI	Niger	Niger	1271500106	566 000
BANANKORO	-8.67	11.68	MALI	Bagoé	Niger		
BENENY KEGNY	-4.93	13.39	MALI	Bani	Niger	1271600105	116 000
BOUGOUNI	-7.45	11.40	MALI	Baoulé	Niger		15 700
DIOILA	-6.80	12.52	MALI	Baoulé	Niger		32 500
DIRE	-3.38	16.28	MALI	Niger	Niger	1271500118	366 500
KE-MACINA	-5.36	13.96	MALI	Niger	Niger		147 000
KIRANGO AVAL	-6.08	13.70	MALI	Niger	Niger		137 000
KORIOUME	-3.03	16.67	MALI	Niger	Niger		
DOUNA	-5.95	13.18	MALI	Bani	Niger	1271600108	101 600
KOULIKORO	-7.56	12.86	MALI	Niger	Niger	1271500142	120 000
SELINGUE	-8.17	11.58	MALI	Sankarani	Niger		
TAOUSSA (TOSSAYE)	-0.55	16.97	MALI	Niger	Niger		340 000
GARBEY-KOUROU	1.62	13.73	NIGER	Sirba	Niger	1271502403	38 750
KANDADJI	0.98	14.60	NIGER	Niger	Niger	1321500053	
NIAMEY	2.10	13.50	NIGER	Niger	Niger	1271500127	700 000
JIDERE BODE	4.08	11.23	NIGERIA	Niger	Niger	1331500019	563 500
KAINJI AVAL	4.60	10.03	NIGERIA	Niger	Niger	1331500020	
LOKOJA	6.75	7.80	NIGERIA	Niger	Niger	1331500023	1 089 000
JEBBA	4.5	9.1	NIGERIA	Niger	Niger	1331500114	6 348 000
UMAISHA	7.12	7.59	NIGERIA	Benue	Niger	1331700101	
KENDE	4.25	11.52	NIGERIA	Sokoto	Niger	1331501701	
BAFING MAKANA	-10.28	12.55	SENEGAL	Bafing	Sénégal	1272600105	21 700
BAKEL	-12.45	14.90	SENEGAL	Senegal	Sénégal		
KIDIRA	-12.22	14.45	SENEGAL	Falémé	Sénégal		
OUALIA	-10.38	13.60	SENEGAL	Bakoye	Sénégal	1272601412	84 400
DAKA-SAIDOU	-10.62	11.95	Sénégal	Bafing	Sénégal	1272600110	15 500
GOURBASSI	-11.63	13.40	Sénégal	Falémé	Sénégal	1272601606	17 100

Figure 5. Overview of the metadata file

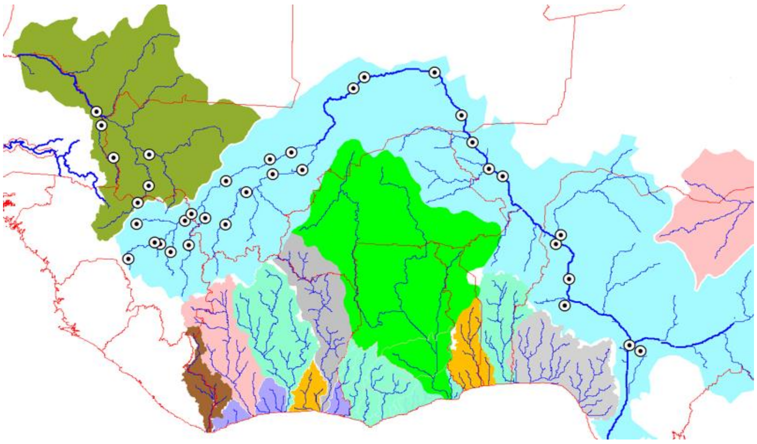

Figure 6. Location of stations

3.2.1 Water level/discharge from manual stations

This input data includes hydrometric measurements of water level surface elevation ("stage") and/or volumetric discharge (flow) collected manually.

3.2.1.1 Dependencies

National agencies send data to AGRHYMET and AGRHYMET prepares it for H-TEP (formatting and putting it on FTP)

3.2.1.2 Protocol

All data that will feed the system from AGRHYMET will be hosted by the ftp site: which is password protected.

3.2.1.3 Data format

The current data formats will be adapted to fit operational processing. It is also not yet operationally updated on a regular schedule.

The water level and the river discharge consists in excel files and in the near future the file name will be structured as follows: StationName_StationID_Variable_Resolution_DateStart_DateEnd.xls

a. Variable: water level or discharge
b. Resolution: hourly, or daily
c. DateStart: The first date in the file that there is data for.
d. DateEnd: the last date in the file that there is data for.

To treat the data automatically all xls files will have the same headers indicating the content of each column. Also, all stations within the file Auto_data and Manual_data will have to share the same ID as in the Metdata_stations.xlsx.

Figure 7 below is an example of the current data format which is in process to be modified and harmonised. The first two columns of figure 7 are identical to all files regardless of the service. The first indicates the date and the second the water levels (or discharge). However, the next two columns differ from one source to another. The third column provides information on the validity of the data. The fourth column indicates the possibility of modifying it.

	A	B	C	D	E
1	Date	Value	Validity	editable	
2	01/01/1981 00:00:00	368.2	N	O	
3	02/01/1981 00:00:00	354.9	N	O	
4	03/01/1981 00:00:00	325.6	N	O	
5	04/01/1981 00:00:00	311.7	N	O	
6	05/01/1981 00:00:00	299.1	N	O	
7	06/01/1981 00:00:00	286.7	N	O	
8	07/01/1981 00:00:00	276.5	N	O	
9	08/01/1981 00:00:00	272.1	N	O	
10	09/01/1981 00:00:00	269.1	N	O	
11	10/01/1981 00:00:00	267.2	N	O	
12	11/01/1981 00:00:00	259.8	N	O	
13	12/01/1981 00:00:00	253.8	N	O	
14	13/01/1981 00:00:00	248.5	N	O	
15	14/01/1981 00:00:00	229.5	N	O	
16	15/01/1981 00:00:00	218	N	O	
17	16/01/1981 00:00:00	217.2	N	O	
18	17/01/1981 00:00:00	215.6	N	O	
19	18/01/1981 00:00:00	210.5	N	O	
20	19/01/1981 00:00:00	207.1	N	O	
21	20/01/1981 00:00:00	193.3	N	O	
22	21/01/1981 00:00:00	175.5	N	O	
23	22/01/1981 00:00:00	162.5	N	O	
24	23/01/1981 00:00:00	156.3	N	O	
25	24/01/1981 00:00:00	151.9	N	O	
26	25/01/1981 00:00:00	152.9	N	O	
27	26/01/1981 00:00:00	157	N	O	
28	27/01/1981 00:00:00	155.4	N	O	
29	28/01/1981 00:00:00	146.9	N	O	
30	29/01/1981 00:00:00	138.2	N	O	

Figure 7. Overview of the current HYDROMET format of the data collected manually

3.2.1.4 Exchange frequency

The water level and water discharge data from manual stations is collected on monthly basis by field technicians who share the readings via email with AGRHYMET. The recipient connects periodically to the FTP server access updated in-situ datasets.

3.2.1.5 Data size

The size of the Excel files is generally small, under 20kB.

3.2.1.6 Notification

No notification is needed. The recipient connects periodically to the FTP server and downloads new data when available.

3.2.1.7 Data removal

AGRHYMET does not remove this datasets after creation from the FTP server.

3.2.2 Water level/discharge from automatic stations

This input data includes hydrometric measurements of water level surface elevation ("stage") and/or volumetric discharge (flow) collected automatically from the hydrometric stations. H-TEP pulls the water level/discharge from AGRHYMET's repository via FTP.

3.2.2.1 Dependencies

Readings from automatic stations are received every 1 to 3 hours, depending on the stations, except in case of connection problem of the GSM network.

3.2.2.2 Protocol

All data that will feed the system from AGRHYMET will be hosted by the ftp site.
The data of the automatic stations are accessible from the same ftp site.

3.2.2.3 Data format

The automatic stations measure the water level in centimetres and the supply voltage in volts (which informs about the state of the battery supplying the station. The battery itself is charged by a solar panel during the day). This information is saved in a TXT file as shown in the figure 8.

The contents of the automatic station files are described as follows.

T001: 05.06.2018,19: 00, PLL05288, CH01 = 166.88, CH32 = 13.18; corresponds to:

T001: the title of the transmission. There is only one push of data per hour.

05.06.2018: the date (June 5, 2018).

19:00: time of transmission of the data (local time).

PLL05288: identifier of the station (this one corresponds to the station of Kandadji).

CH01 = 166.88: channel 01 corresponds to that which measures the height of water (166.88 cm).

CH32 = 13.18: Channel 32 corresponds to the voltage of the battery.

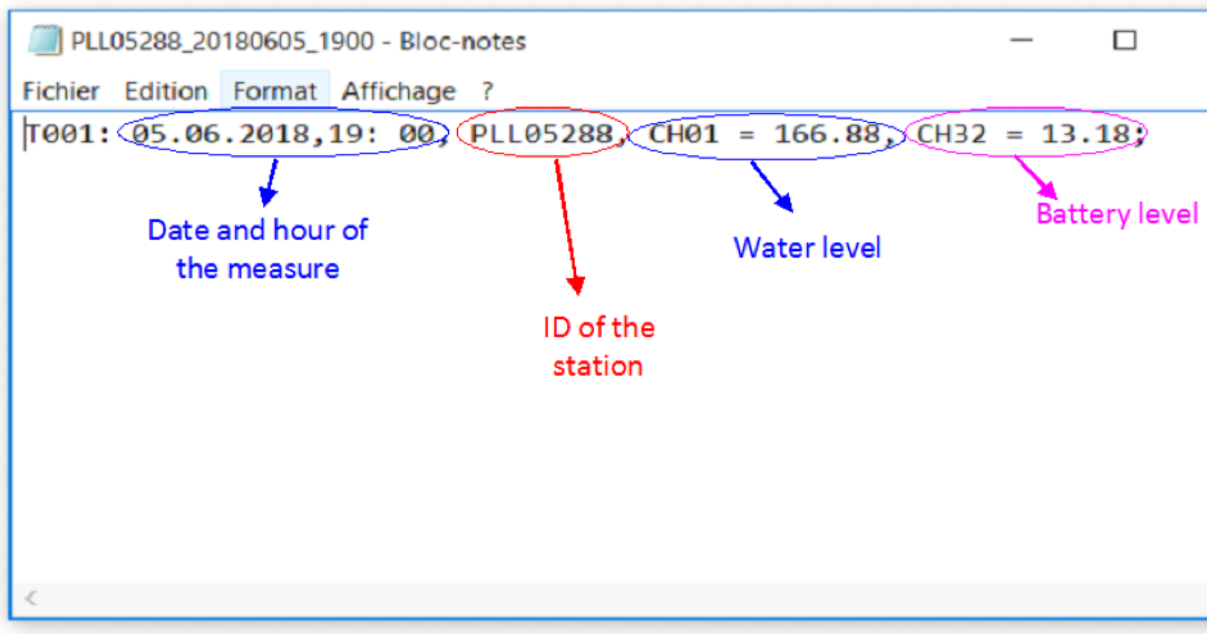

Figure 8.Overview of the data format of automatic stations

3.2.2.4 Exchange frequency

AGRHYMET receives the water level and river discharge from automatic stations every hour except in case of connectivity issues in the GSM network.

H-TEP connects to the FTP server every day to get access to the water level and water discharge from automatic stations.

3.2.2.5 Data size

The size of the Excel files is fairly small, around 1kB.

3.2.2.6 Notification

A failure message notification is needed in case of failure of the automated operational system. This is not set in the system as yet but should be considered to secure the operational production.

3.2.2.7 Data removal

AGRHYMET does not remove data files after creation from the FTP server.

3.2.3 Rating curves

AGRHYMET provides rating curves which consists in a graph of discharge versus water level for a given gauging stations, where the stream discharge is measured across the stream channel with a flow meter. H-TEP pulls the rating curves from AGRHYMET's repository via FTP.

NIHSA also have rating tables of some hydrometric (not graph) stations in the Niger Basin supplied to AGRHYMET.

3.2.3.1 Dependencies

The rating curves depend on the availability of the readings of water discharge and water level.

3.2.3.2 Protocol

Rating curves are provided by countries until an update is done. They are accessible from AGRHYMETs FTP site.

3.2.3.3 Data format

The rating curves consist in excel files as per figure 9 below.

H (cm)	Q (m3/s)	valididité
95	0,000	Oui
96	0,478	Oui
97	1,310	Oui
98	2,160	Oui
99	3,010	Oui
100	3,880	Oui
101	4,760	Oui
102	5,640	Oui
103	6,540	Oui
104	7,440	Oui
105	8,360	Oui
106	9,280	Oui
107	10,200	Oui
108	11,200	Oui
109	12,100	Oui
110	13,100	Oui
111	14,100	Oui

Figure 9. Data format of a rating curve

3.2.3.4 Exchange frequency

The rating curves are generated on demand by season. The same rating curve can be available for years.

3.2.3.5 Data size

The size of the Excel files is generally small, under 30kB.

3.2.3.6 Notification

A failure message notification is needed in case of failure of the automated operational system. This is not set in the system as yet but should be considered to secure the operational production.

3.2.3.7 Data removal

The latest updated rating curve remains the one to use, therefore, AGRHYMET will add a date stamp on the rating curves so the user knows which rating curve refers to what period.

4.2.4 Meteorological hindcast data by AGRHYMET

AGRHYMET data can be used to produce a corrected meteorological hindcast dataset in a similar way as HYDROGFD2.0. In principle the HYDROGFD2.0 algorithm will be applied (on the same input datasets as in Table 2) but instead using AGRHYMET's data for correction. These data are obtained by combining satellite data (TAMSAT daily Rainfall Estimates data for precipitation and Japanese 55-year Reanalysis data for extreme temperatures) and data from West African meteorological station observation networks. These data cover West Africa with a resolution of 0.5 ° X0.5 °. This consists of replicating the HYDROGFD method by correcting the HYDROGFD2.0 datasets with AGRHYMET datasets. These data are divided into two groups:

- Climatological data:
    - Monthly precipitation over the 1981-2010 climatological period;
    - Monthly minimum temperature over the 1981-2010 climatological period;
    - Monthly maximum temperature over the 1981-2010 climatological period;
    - Monthly mean temperature over the 1981-2010 climatological period;
    - Monthly number of wet days over the 1981-2010 climatological period.
- The anomalies last three months that will be used to calculate bias factors for HYDROGFD2.0 data correction:
    - Monthly anomalies of precipitation;
    - Anomalies of mean temperature;
    - Anomalies of number of wet days.

The figures 10 and 11 below show an overview of these data.

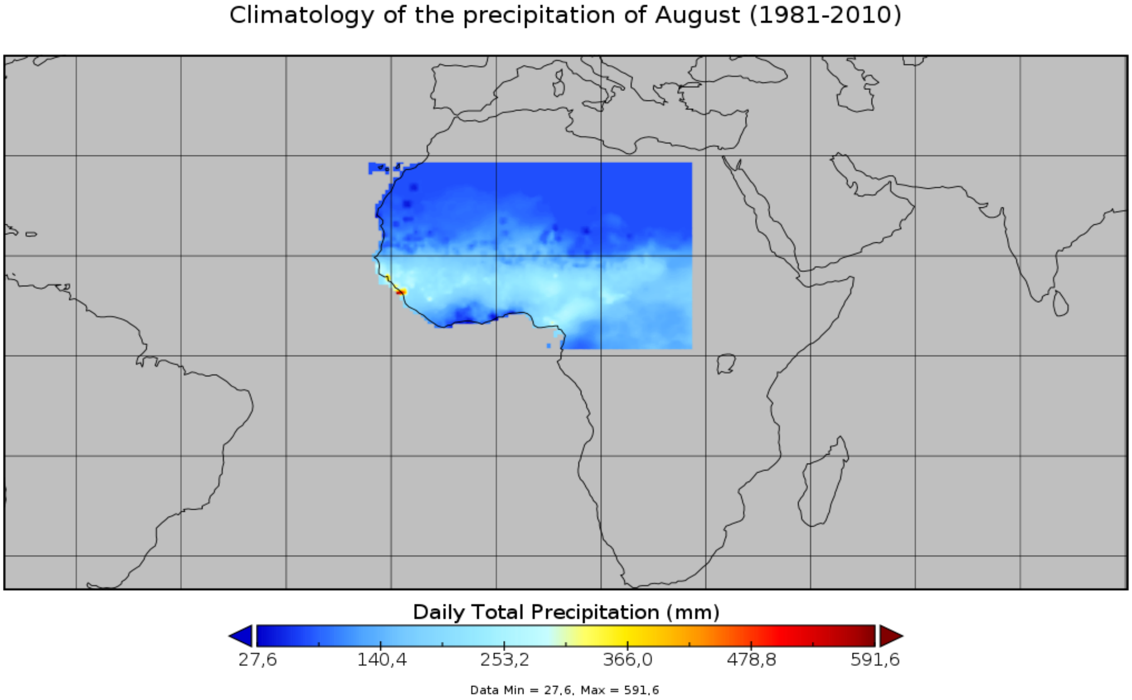

Figure 10. Climatology of the precipitation

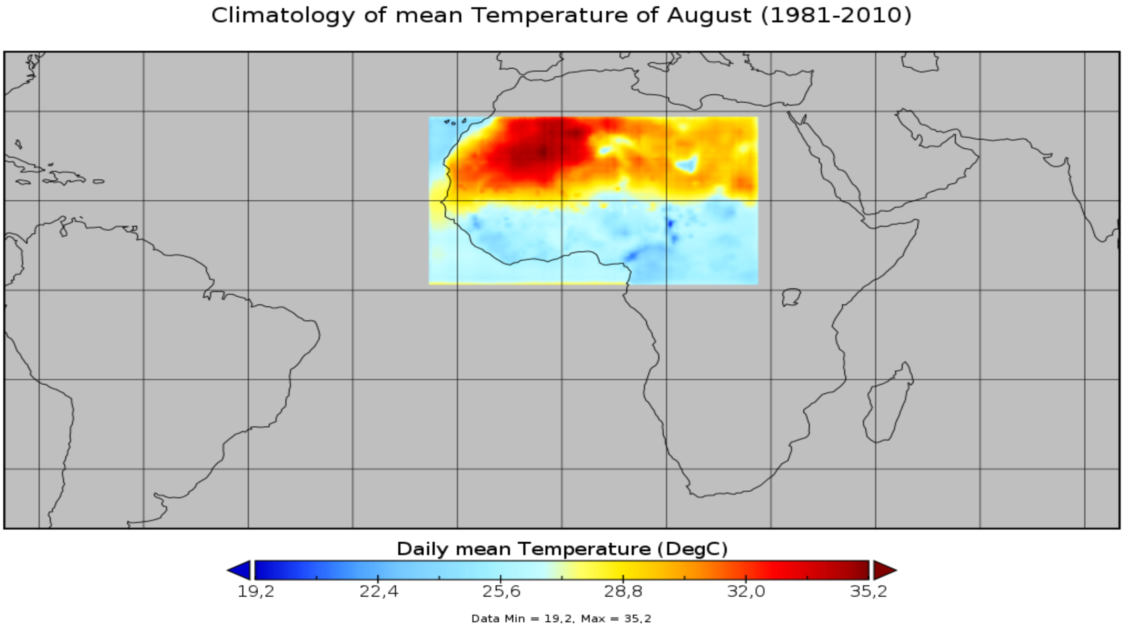

Figure 11. Climatology of mean temperature

4.2.4.1 Dependencies

These data will be accessible via the ftp site of AGRHYMET. Their availability depend on the availability of the inputs to the HydroGFD2.0 dataset (Table 2), which will be operationally transferred from SMHI to AGRHYMET.

4.2.4.2 Protocol

The transfer is not operational yet. This will implemented as a transfer from AGRHYMET to H-TEP by the FTP protocol, in such a way that H-TEP regularly pulls the data from AGRHYMET's FTP server.

4.2.4.3 Data format

All files are provided in netCDF format. The metadata are shown in the figure 12 below.

```
File D:/Merging/hydrogfd_format/climatology/precipitation/niger_pcpmerged_climatology_19812010_mon01.nc (NC_FORMAT_C

     2 variables (excluding dimension variables):
        float precip[Lon,Lat]
            long_name: Daily total of precipitation
            units: kg/m2
            missing_value: NaN
            standard_name: precipitation
            title: precip
        float nonNA[Lon,Lat]
            long_name: Fraction of the available data
            _FillValue: NaN
            missing_value: NaN

     2 dimensions:
        Lon  Size:95
            standard_name: longitude
            long_name: Lon
            units: degreeE
            axis: X
        Lat  Size:57
            standard_name: latitude
            long_name: Lat
            units: degreeN
            axis: Y

    3 global attributes:
        Version: Merged-4.0.0
        Created: 2019-02-15
        Contact: Bernard Minoungou (bernard.minoungou@cilss.int)
```

Figure 12. Meteorological hindcast format example.

4.2.4.4 Exchange frequency

New files are searched for every day, however some datasets are only updated once per month.

4.2.4.5 Data size

The data size is approximately 3170 Kb.

4.2.4.6 Notification

A failure message notification is needed in case of failure of the automated operational system. This is not set in the system as yet but should be considered to secure the operational production.

4.2.4.7 Data removal

Data will be store at AGRHYMET and H-TEP. Data will be regularly removed.

3.3 Interface#7 NIHSA's River observations

The data obtained from the hydrometric stations managed by NIHSA is sent by email to AGRHYMET who is then compiling it and sharing it with the SMHI following the same process as explained at 4.2 sub-section above (Interface #5). The format of the data produced from NISHA is the same as that shown in Figure 4. Daily data are collected from the gauge readers by text message or from the automatic stations every 3 hours and sent to AGRHYMET monthly.

3.3.1.1 Dependencies

N/A

3.3.1.2 Protocol

All data that will feed the system from AGRHYMET will be hosted by the ftp site: which is password protected.

3.3.1.3 Data format

The water level and the water discharge consist in TXT files.

3.3.1.4 Exchange frequency

The automatic stations transmit data every 3-hours. The manual station data are collected at 7.00am, 12.00noon and 6.00pm by the gauge readers and the average daily data are received by text message and sent to AGRHYMET monthly by Email from NIHSA.

3.3.1.5 Data size

Some stations with little historical produce data of about 330kb (i.e Baro Station) and those with up to-date data from 2.0 to 3.0 mb or more (Lokoja, kainji etc.)

3.3.1.6 Notification

A failure message notification is needed in case of failure of the automated operational system. This is not set in the system as yet but should be considered to secure the operational production.

3.3.1.7 Data removal

AGRHYMET does not remove datasets after creation from the FTP server.

3.4 Interface #9 River observations from other Met. Agencies

The format of the data received from other national hydrology services is in the format shown in figure 4. These data are also received by the AGRHYMET Regional Center by email on monthly basis.

3.5 Interface#10a Water level by isardSAT

3.5.1 Introduction

The Water Level Service developed by isardSAT provides the water level time series of a number of (small) lakes and rivers. The service is based on altimetry data from Sentinel 3, Jason-2 and SARAL-Altika. isardSAT's satellite-based water level products of the Niger River basin floodplain are sequentially assimilated into a hydrological model from SMHI to generate river discharge return periods in FANFAR and complements in-situ sensor networks. It can also be used as a proxy of streamflow, for model calibration and validation as well as for hydrologic data assimilation making possible for the FANFAR community to access to data where there was none before. An input from the water level service is included in the figure 13 below.

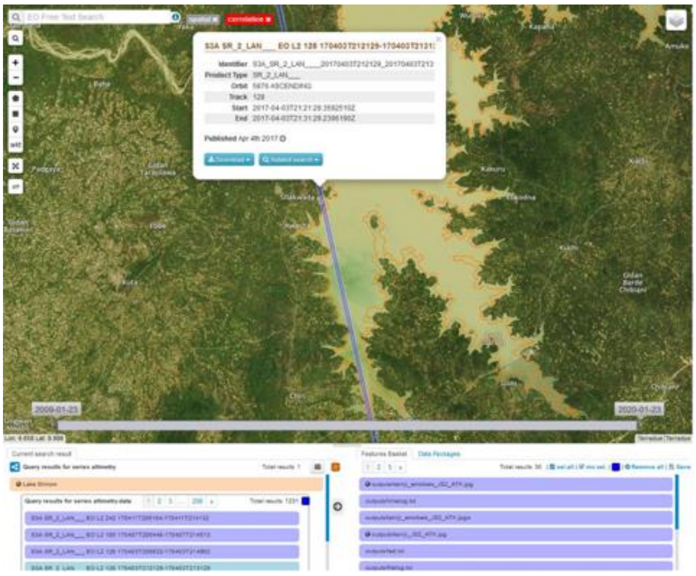

Figure 13. Track over Kainji lake, Western Nigeria, input to the water level service

3.5.1.1 Dependencies

Input for the water level application, consists of a water mask of the water body or river of interest as well as altimetry data. For altimetry data at the H-TEP the EO data available is obtained from multiple satellite sensors with varying spatial, temporal and spectral resolution.

3.5.1.2 Protocol

The protocol of the water level input is a OGC WPS

3.5.1.3 Data format

The water level files consist in a csv file.

3.5.1.4 Exchange frequency

The water level service on the H-TEP relies on data provided by different satellites, as mentioned earlier which have different orbital cycles. For example, Sentinel 3 (A+B) orbital cycle is 27 days (14+7/27 orbits per day, 385 orbits per cycle) which means that time taken for the satellite to pass over the same geographical point on the ground is every 27 days.

For example, for SEN-3 data H-TEP harvest all new available data from ESA's SciHub which is store in the H-TEP catalogue. Then, on daily basis a scheduler queries the catalogue on specific filters, such as a water body. When a new acquisition cross the water body the H-TEP launches a water level processing over that water body.

3.5.1.5 Data size

3.5.1.6 Notification

A failure message notification is provided by H-TEP is the water level processing job is not successful.

3.5.1.7 Data removal

The water level data results are stored persistently in the H-TEP catalogue and never removed.

3.6 Interface#10b Water bodies by isardSAT

3.6.1 Introduction

isardSAT has produced a set of 570 water bodies as a group of georeferenced polygons representing the intersection between the above mentioned satellite altimetry tracks and water masks. They are water bodies with potential for water level altimetry based estimations and are used as inputs to systematically execute water level time series. They are in a csv file format and WKT protocol format. The csv columns are the attributes for each polygon (i.e. area). Further attributes can be manually added on demand if needed.

The water bodies obtained from the satellite missions have been generated by isardSAT following the rule:

- Water coverage during 80% of the year and;
- Buffer zone extended to a 150-metre area alongside of the satellite tracks.

An example of a water body input is included in the figure 14 below.

Figure 14. Water body input generated at the Kpong dam, lower Volta river

3.6.1.1 Dependencies

There are no dependencies for the water bodies.

3.6.1.2 Protocol

The water bodies are provided to the H-TEP repository through FTP.

3.6.1.3 Data format

The process to obtain the water bodies is done externally and the resulting csv in WKT files are available on FTP on demand.

Feature	Value
WKT_INTERSECTIONS_WATER_BODIES	
WKT	POLYGON ((0.115097525947312 6.49450094988938,0.080919387719875 6.64749987578923,0.081025816962109
(Derived)	
(clicked coordinate X)	0.0928
(clicked coordinate Y)	6.5980
Area	10,331 km²
Closest vertex X	0.0823
Closest vertex Y	6.6475
Closest vertex number	6
Perimeter	36,789 km
Vertices	16
feature id	198
(Actions)	
	View feature form
WKT	POLYGON ((0.115097525947312 6.49450094988938,0.080919387719875 6.64749987578923,0.081025816962109
BUFF_DIST	300.00000000000
AREA	10330817.05

Figure 15. Water body data description

3.6.1.4 Exchange frequency

The process to obtain the water bodies is done externally and the resulting csv in WKT files are available on FTP on demand.

3.6.1.5 Data size

The size of a single water body is approx. 200kB.

3.6.1.6 Notification

A failure message notification is needed in case of failure of the automated operational system. This is not set in the system as yet but should be considered to secure the operational production.

3.6.1.7 Data removal

When a new water body is updated the former will be removed from the H-TEP repository.