Assessing the forecast skill of GLOFAS and validating historical impact data for their use in the trigger methodology of the EAP for river floods in Zambia

Table of Contents

[Summary of Early action plan 3](#_Toc531863810)

[Alignment with humanitarian principles 3](#_Toc531863811)

[Stakeholders 3](#_Toc531863812)

[Flood risks in Zambia 4](#_Toc531863813)

[Current status: forecast skill of WARMA 5](#_Toc531863814)

[Materials and methods: assessing the forecast skill of GLOFAS 6](#_Toc531863815)

[Materials and methods: validating historical impact data 7](#_Toc531863816)

[Results: Trigger model (impact based forecasting model) 7](#_Toc531863817)

[Table 2Trigger Methodology : preliminary analysis 7](#_Toc531863818)

[Riverine Discharge data 7](#_Toc531863819)

[Historical Floods Data 8](#_Toc531863820)

[Validation of historical flood events using satellite imaging 9](#_Toc531863821)

[Analysis of hindcast discharge with GLOFAS and estimate of trigger level 12](#_Toc531863822)

[Comparison with observations 13](#_Toc531863823)

[Skill of the Model 19](#_Toc531863824)

[Possible workflow in the FbF system based on preliminary analysis 19](#_Toc531863825)

[Menu and selection of forecast 19](#_Toc531863826)

[Definition of the level of impact 19](#_Toc531863827)

[Selection of actions 20](#_Toc531863828)

[Monitoring evaluation, responsibility and learning 20](#_Toc531863829)

[National society capacity for early action 20](#_Toc531863830)

[Finances and logistics 20](#_Toc531863831)

[EAP approval and validation with other organization 20](#_Toc531863832)

[Conclusion 20](#_Toc531863833)

[Annexes 20](#_Toc531863834)

# Summary of Early action plan

## Alignment with humanitarian principles

In developing and drafting the EAP, the FbF team took into account the humanitarian principles (e.g. humanitarian impartiality, neutrality, impartiality, independence, 'do no harm' and local participation) related to the Humanitarian Code of the Red Cross and NGOs, the principles and rules of humanitarian assistance of the Red Cross and Red Crescent, the principles of 'good humanitarian donorship', the principles of SPHERE protection (e.g. principle 1 to 3) and the SPHERE minimum standards.

Basically, the `FbF mechanism' aims to directly bolster the (pro)-active capacity for action of the population at the moment of being in situations of possible affectation and stress of survival (principle 7 "good humanitarian donorship").

In addition, the early actions in this EAP have been defined based on consultation and participation of a national technical committee, which includes all lead institutions in sectoral disaster management and chaired by the Disaster management and Mitigation unit (DMMU) of Zambia, with the coordination of Zambian Red Cross.

The strategies, plans and guidelines of the government as well as capacity of Zambian red cross capacity in coordinating response to take ‘early action’ have been considered.

During the development of the EAP, the team was guided by the SPHERE standards for action identification, relations with DRM (Disaster Risk Management) authorities, analysis and monitoring.

Humanitarian supplies and sensitizations will strengthen local family solutions and are suitable for use during the extreme event. For purchases, priority has always been given to local assets in the same region.

# Stakeholders

The EAP for flooding has been designed to be implemented by ZRC in conjunction with other organizations such as DMMU, WARMA, ZAM , community leaders, and other regional stakeholders.

Each of the actors involved in this EAP has a key role to play during the coordination and activation of the Forecast Based Preparation mechanism. Thus, this EAP should be part of the contingency plans of the national, regional and provincial government entities as well as the National Society of the ZARC.According to their institutional mandate or the coordination carried out within the framework of this protocol, each institution has some roles and responsibilities . Basically, these roles are divided into three categories: coordination, communication and implementation of field actions

Zambian Red cross

DMMU

WARMA

ZAM …

A table listing all the organizations

# Flood risks in Zambia

Flood risks in Zambia consist of:

* Fluvial/river flood: mostly large space and time scale riverine floods, characterized by usually high direct damage and displacement of many people due to their wide flood risk extent, and generally densely populated flood plain areas.
* Flash flood, that occur at smaller time and spatial scales. A flood that rises and falls quite rapidly with little or no advance warning, usually the result of intense rainfall over a relatively small area (Glossary of the American Meteorological Society, 2000 edition). Key aspect of the definition is the time scale: sudden hydrological response to the causative event.
* Pluvial flooding: direct runoff over land causing local flooding in areas not previously associated with natural or manmade water courses. Key aspect of the definition is the lack of proper drainage network in the area impacted by the flood. (see http://ec.europa.eu/environment/water/flood\_risk/index.htm)

And what is the impact of the dams in e.g. the Zambezi delta plain (ref two MSc studies UNESCO-IHE students) in causing floods?

What triggers flash floods is more complex to understand than for riverine floods, as these are often caused by a combination of high local precipitation rate due to convective storm-related events, associated to specific hydrological situation and geomorphological characteristics of the catchment. In general, flash and fluvial floods have shorter lead times and are harder to predict than river floods.

AB : One exercise could be to identify areas and flood plains that are prone to Flash floods. Indeed, Flash floods triggers and impacts are different than riverine floods and should be studied independently.

Conclusions: the difficulty to forecast pluvial and flash floods combined with the fact that in Zambia, river floods cause the most humanitarian impact (TO BE CHECKED!), we focus in the development of the triggers for the EAP on river floods.

Table 1 Overview of flood prone districts (REF analysis Aki?)

|  |  |  |
| --- | --- | --- |
|  |  | **Floods** |
|  | Province | District |
| 1 | Western | Kalabo |
|  |  | Mongu |
|  |  | Sioma |
|  |  | Shangumbo |
|  |  | Sesheke |
|  |  |  |
| 2 | Southern | Mazabuka |
|  |  | Namwala |
|  |  | Namwala |
|  |  | Kazungula |
|  |  |  |
| 3 | Central | Kabwe |
|  |  | Itezhi Tezhi |
|  |  | Mumbwa |
|  |  |  |
| 4 | Lusaka | Lusaka |
|  |  | Kafue |
|  |  |  |
|  |  |  |
| 5 | Eastern | Luangwa |
|  |  | Chipata |
|  |  | Mambwe |
|  |  |  |
| 6 | Copperbelt | Kitwe |
|  |  | Lufwanyama |
|  |  | Masaiti |
|  |  | Chambeshi |
|  |  |  |
| 7 | Northern | Mpulungu |
|  |  | Kasama |
|  |  | Mpika |
|  |  |  |
|  |  |  |
| 8 | Luapula | Mansa |
|  |  | Chembe |
|  |  | Kawambwa |
|  |  |  |
| 9 | Muchinga | Chinsali |
|  |  | Nakonde |
|  |  | Chama |
|  |  |  |
| 10 | North Western | Chavuma |
|  |  | Zambezi |
|  |  | Kabompo |
|  |  | Manyinga |

# Current status: forecast skill of WARMA

First, describe the hydrological modelling NHMSs use: models used, spatial and temporal resolution usually achieved. Temporal resolution: lead time of warning probabilistic and/or deterministic, meteorological and hydrological observational network.

*Table 1 Generic overview of NHMS hydrological models*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hydrological analysis** | **Models used** | **Forecast products** | **Spatial resolution** | **Temporal resolution** |
| Rainfall– runoff |  |  |  |  |
| Hydrodynamic |  |  |  |  |
| Data assimilation (DA) |  |  |  |  |

So GLOFAS could add value xx

# Materials and methods: assessing the forecast skill of GLOFAS

Flood risk analysis includes analysis of historical impact data, vulnerability indicators and exposure indicators. To complete this analysis, a review of primary data from the different sources of information at the national level was conducted and secondary sources were also used

The continuous evaluation of GloFAS performance over different domains and temporal scales is necessary to both gain the trust of the end-users and to guide its further improvement (ref Bischiniotis). This research presents an evaluation framework that assesses GloFAS forecasting skills in Zambia, which is a country that has experienced many devastating flood events over the past decade. The predictive skill is assessed on a so-called hindcast mode, using daily forecasts for the years xx over lead times from 1 to 30 days.

The skill is examined from two different perspectives:

a) by calculating several verification scores at every river point of the Zambian river network

b) through an event-based analysis, by comparing the flood signals against collected information from multiple disaster databases such as xx.

Finally, the quantile mapping technique was applied in order to evaluate whether a simple post-processing may be of use to humanitarian organizations and decision-makers for better preventive flood risk management planning.

The basin shape files and coordinates for the selected locations for Zambia were used to extract the GloFAS data for hindcast period. The ensembles of 30 days flow outlook was provided in NetCDF format. The ensemble mean was extracted for all individual stations. Performance of ensemble mean was analyzed again the observations. The performance of forecast with lead time of 5-days, 10-days, 15-days, 20-days, 25-days and 30-days is presented in this document. Various statistical parameters including mean error (ME), root mean square error (RMSE), nash-sutcliffe efficiency (NSE), percentage bias (PBIAS) and coefficient of determination (R2) were considered. The hydrographs for the analysis period for different lead time of forecast and histograms representing the sampling of quantiles is also provided for better comparison of flow patterns (ref RIMES document).

SHOULD WE ALSO EVALUATE 5, 10, 15, 20 AND 15 DAYS FORECASTS?

Table 2 Comparison of flow statistics of the observed and GLOFAS forecasted flow series for xx area, unit m^3/s

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Statistics | Observation | 5-day forecast | x-day forecast | 30-day forecast |
| Minimum |  |  |  |  |
| Maximum |  |  |  |  |
| 1st Quartile |  |  |  |  |
| Median |  |  |  |  |
| 3rd Quartile |  |  |  |  |
| Mean |  |  |  |  |
| Standard deviation (n-1) |  |  |  |  |
| Standard error of the mean |  |  |  |  |

Table 3 Statistical criteria for examining the accuracy of the forecast time series for xx area

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | MAE | ME | SS | PBIAS | NSE | R^2 |
| 5-day forecast |  |  |  |  |  |  |
| x-day forecast |  |  |  |  |  |  |
| 30-day forecast |  |  |  |  |  |  |

# Materials and methods: validating historical impact data

Get here all information on the methodology separated from results

Historical analysis of satellite imagery to validate impact data.

Can we bring the two analyses together: GLOFAS and satellite imagery analysis? Can we also use the historical imagery to get a flood extent map for a specific GLOFAS point forecast? Or should we use for this global models that produce flood extent maps like GLOFRIS etc (see Trigg paper).

# Results: Trigger model (impact based forecasting model)

## Table 2Trigger Methodology : preliminary analysis

### Riverine Discharge data

In order to determine the danger levels for a flood in Zambia we use historical simulated data from GLOFAS. The Global Flood Awareness System (GloFAS), jointly developed by the European Commission and the European Centre for Medium-Range Weather Forecasts, is a global hydrological forecast and monitoring system that delivers global ensemble discharge forecast for major rivers.

We acquired historical simulations from GLOFAS for 24 stations in Zambia. The simulations span over a time range of 22 years (from 1997 to 2018) and represent the hindcasted discharge at a given point in time for the next 30 days calculated with 10 ensemble members. Figure 1 shows the position of the GLOFAS data extraction points in Zambia.

Figure 1 Map showing the position of the monitoring station in Zambia represented within the GLOFAS model. The borders represent the administrative borders of the districts (source OCHA). [SOME OF THE NAMES NOT SHOWING: CHECK ALSO NOT ALL THE STATIONS ARE SHOWN]

iuj

AB : It would be interesting to sort the 24 stations by Catchments . Indeed, looking at the river order in which the stations are located, and the relief of the catchment, we could identify if the location is prone to riverine or Flash flood. In addition, we expect that river gauges in the same river or catchment will respond in the same way to a flood. Triggers could then be identified per group of stations.

### Historical Floods Data

We collected data on historical flood events from the following sources:

* Reports from DMMU
* DREF reports from IFRC
* Newspapers

What kind of data was in these reports? Use table a bit like this?

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Description** | | | **Data Quality** | | | | **Costs of data** | | |
| **Data set** | **Source of data** | **Date of source/disasters covered** | | **Source reliability** | **Content accuracy** | **Granularity** | | **Type** | **Structure level** |
| Flood report | IFRC relief workers | Historical floods | | 3 | 4 | District | | Report | unstructured |
| Disaster Database | EM-DAT | 1967-2017 | | 4 | 4 | National | | Database | structured |
| Internally displaced persons | OCHA | 2010-2016 | | 4 | 3 | District | | Dataset | structured |
| Damage and loss assessment | Government | 2015 | | 3 | 4 | District | | Report | Semi-structured |

We compiled a list of reported flood events for each district. We use these data to validate the forecast and calculate the skill of the model.

Figure 2 histogram showing the number of flood reports in Zambia for each year that have been analyzed for this study. [SHOWS ALSO DATA AFTER 2014: ARE WE USING THOSE?]

# Validation of historical flood events using satellite imaging

The reported events give us a sample of historical floods but not enough information on the severity of the flood. The reports mention impact data inconsistently, the number of displaced or the crop loss but it is challenging to compare the different events based on these data. We can validate the dataset and extract more information on the flood severity analyzing satellite images of water extent collected during the same period as that of the reporting. We ca use this information to test the following hypothesis:

* The reports are representative of the largest floods in Zambia
* There is a correlation between number of reports and extent of the flood event
* There is a correlation between the extent of the flood event and the impact of the event

AB : Not sure if this would be feasible, but Focus Group Discussion in villages at location close to every station would really help to estimate the intensity of each flood events.

We validate the list of reported events using satellite images from the Global Surface Water by EC JRC. This dataset is based on three million images from the Landsat satellites and provide estimates of the changes in the global water extent between 1984 and 2015 at 30m resolution.

Using this dataset on a google engine we extract for each district in Zambia the maximum flood extent recorded on each year and its timestamp. For each event we calculate the deviation from the mean area flooded in a given district and the value of the standard deviation from the mean. We analyze the

The hypothesis we want to test is that extreme events that have consistent impact on the population are those that deviate maximally from the average flooding. In this way if we are considering a flood plane or a generally swampy region the absolute flooding may be large but not very impactful on the population because considered “normal”.

Visual validation with LANDSAT imaging (Roberto’s method)

Zambia is home of 5 major flood planes that can be identified visually from satellite images. For this purpose we use LANDSAT 9 images which are provided free of charge by NASA[[1]](#footnote-2) and specifically images from the Operatinal Land Imager (OLI). OLI collects images in the visual band (XXX-XXX nm) [INSERT IMAGE EXAMPLE].

Floodplains are the areas that are most prone to seasonal flooding and where the most extreme flood events are expected to happen. We use LANDSAT data for identification of flood-planes combining it with topographical data on the whole country to identify flood prone areas around rivers and lakes. [NEED PRECISE METHODOLOGY DESCRIPTION]. Figure 3 shows the identified flood planes in Zambia.

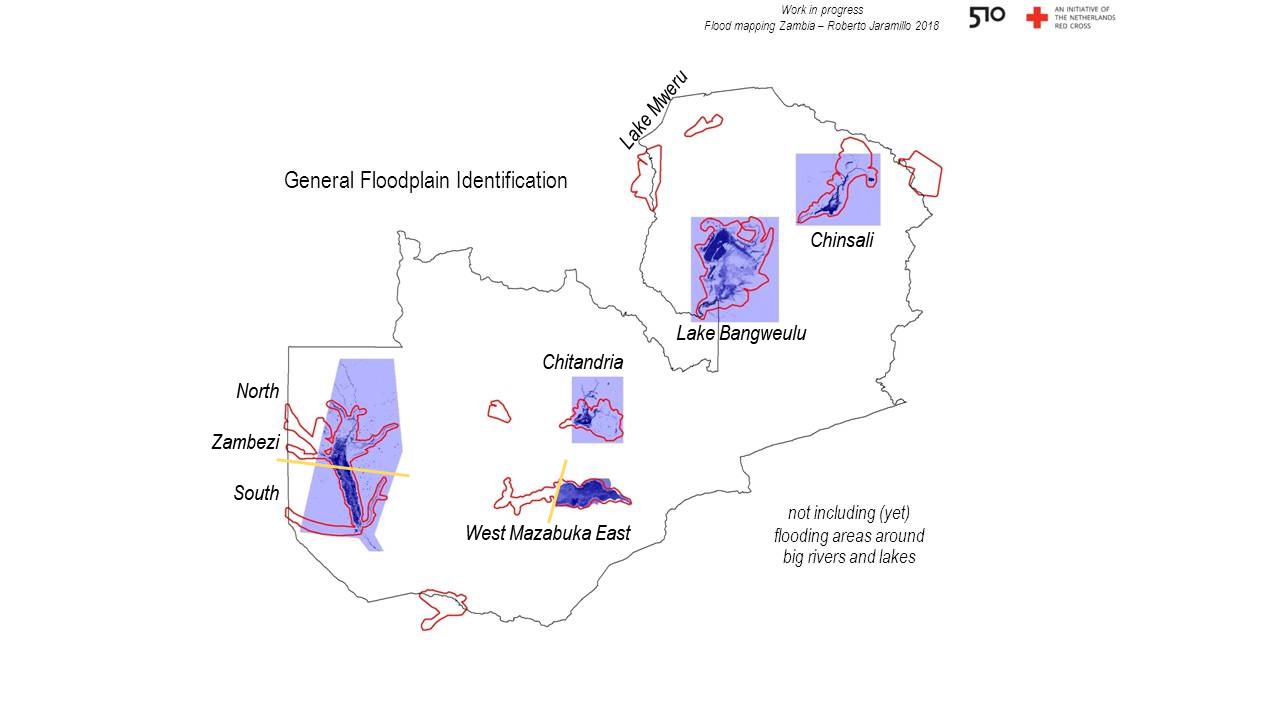


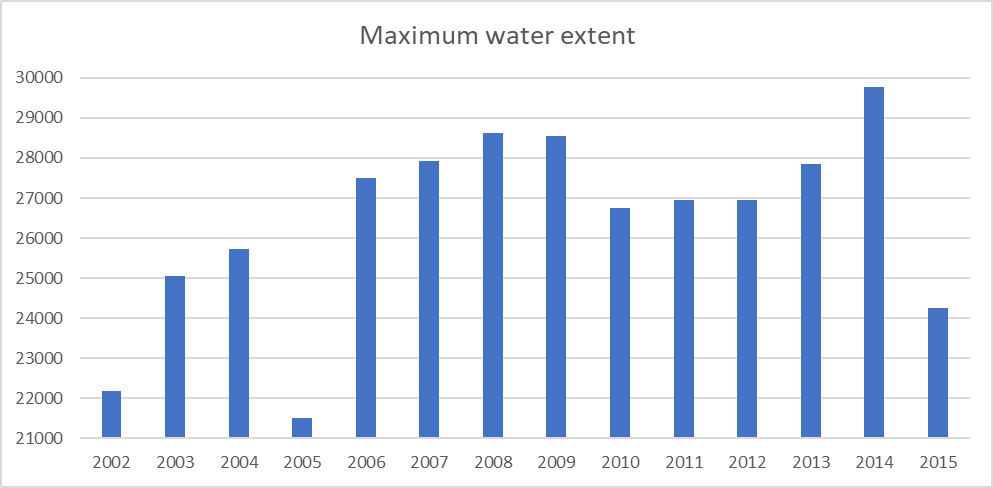
Figure 3

For each flood plain images are extracted for the period 1984 – 2018. For each year the images are visually inspected (HOW?) by a flood expert[[2]](#footnote-3) for flood events. A qualitative scale of magnitude (1-8) is used to classify the events with respect to the size of the flood extent in the image. 1 is low flood extent and 8 is very high level of flooding in comparison to the water extent in normal condition. Everything above 4 is considered an above average event. There is a relation between the return period and the magnitude of the flood.

[GRAPH]

AB : Is it the total flooded area of all flood plains or only Zambesi river? With the data we have, could we correlate these maximum flood extents per flood plains with maximum discharge from GLOFAS? I guess the idea would be to illustrate the relation between Magnitude (flood extent), and GloFAS discharges data and return period?

|  |  |  |  |
| --- | --- | --- | --- |
| year |  | area | rank |
| 2002 |  | 22174.38 | 13 |
| 2003 |  | 25057.08 | 11 |
| 2004 |  | 25728.3 | 10 |
| 2005 |  | 21498.39 | 14 |
| 2006 |  | 27500.94 | 6 |
| 2007 |  | 27932.04 | 4 |
| 2008 |  | 28630.44 | 2 |
| 2009 |  | 28557.09 | 3 |
| 2010 |  | 26756.55 | 9 |
| 2011 |  | 26940.15 | 7.5 |
| 2012 |  | 26940.15 | 7.5 |
| 2013 |  | 27838.35 | 5 |
| 2014 |  | 29757.24 | 1 |
| 2015 |  | 24266.7 | 12 |



We use this data to validate the historical events identified by the reports.

Table 3 Magnitude of flood extent for the Zambesi River flood plane

|  |  |
| --- | --- |
| YEARS | MAGNITUDE OF FLOOD EXTENT |
| 2003 | 4 |
| 2004 | 7 |
| 2005 | 2 |
| 2006 | 4 |
| 2007 | 6 |
| 2008 | 4 |
| 2009 | 8 |
| 2010 | 7.5 |
| 2011 | 0 |
| 2012 | 0 |
| 2013 | 0 |
| 2014 | 7 |
| 2015 | 4 |
| 2016 | 4 |
| 2017 | 4 |
| 2018 | 5 |

### Analysis of hindcast discharge with GLOFAS and estimate of trigger level

Figure 2 shows the hindcasted riverine discharge at Katima Mulilo station (extracted from GLOFAS) for multiple years. Since these are simulated historical data and a forecast refers to the future, we refer to these simulations as “hindcast data”. Figure 3 shows the hindcast timeseries data (blue line) for a monitoring station in Katima Mulilo. is located in the Southern district at the border with Namibia and often experience floods during the rainy season (December - March). If we look at the historical data we can easily identify the peaks of discharge within a given year. Within the scope of this study we want to test the hindcast skill of this model for extreme events only.

Caughlan et al. (2016) describes a methodology to identify extreme events in a riverine discharge time series, by qualitatively selecting the 95% percentile as a proxy for extreme event. From analyzing the performance of this proxy in our data we find that using 90% percentile gives a better detection of the flood events when compared to impact data and we will use this threshold from now on. IS 90% A SENSIBLE PERCENTILE TO USE?

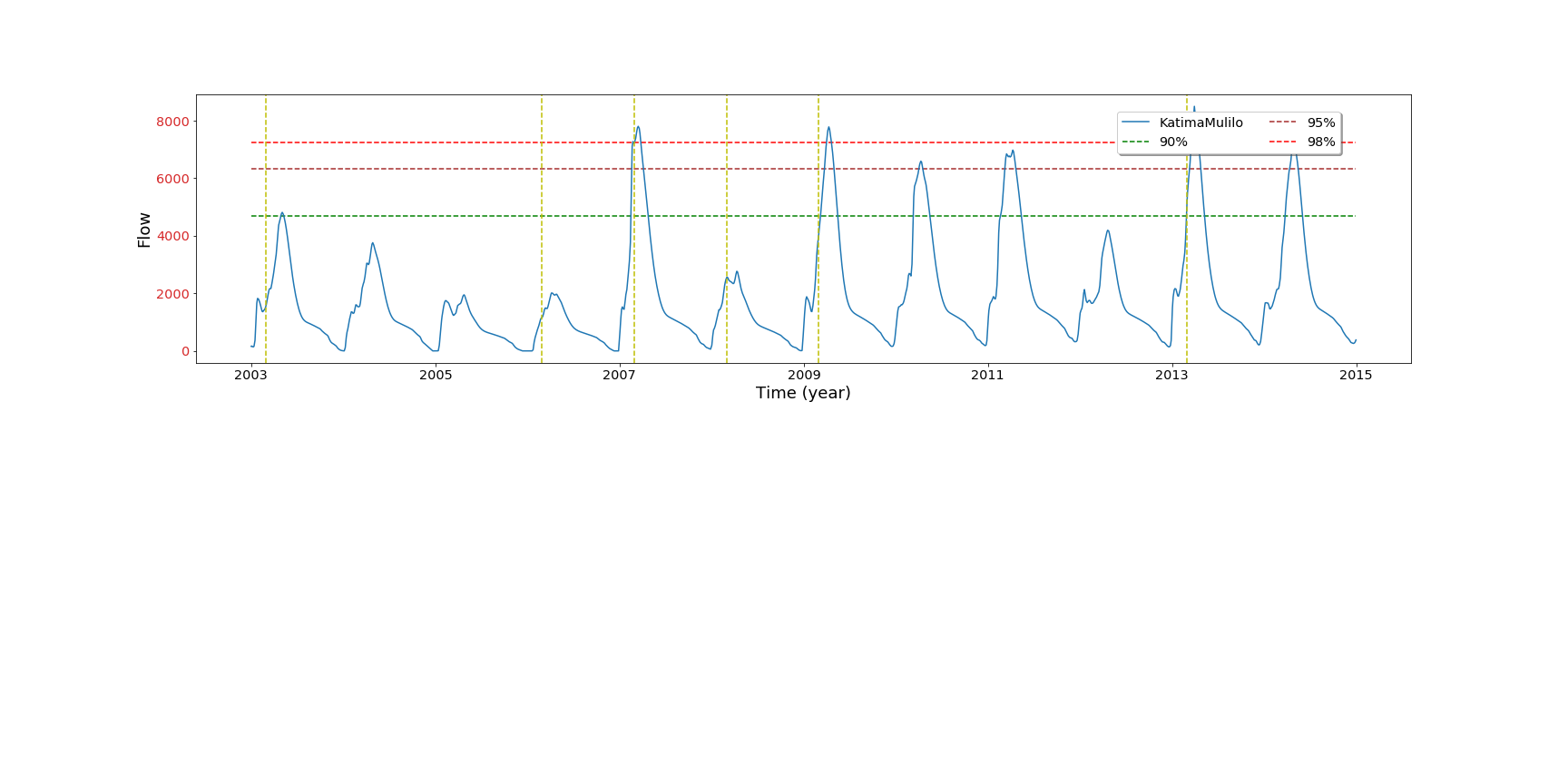


Figure 4 Hindcast data from GLOFAS for the station of Katima Mulilo in the Lukulu district. The green, purple and red dashed line mark the 90%,95% and 98% percentiles respectively. The yellow dotted vertical lines mark the years when significant flood has been reported. Blue line indicate the median of the ensembles for time step 1

In our case this danger level (90%) would have been exceeded in 2003, 2006, 2007, 2009 and 2010, 2011 and 2013. Out of these 7 peaks above threshold, 4 correspond to a report for flood in this district while 3 do not appear in the reports. This could be either false positive or not reported events.

### Comparison with observations

We use observed discharge data from the WARMA monitoring station to validate GLOFAS hindcast data. Out of the 24 GLOFAS stations we have overlapping data for 9 stations. The monitoring data is unfortunately not continuous in time however we compare overlapping time ranges.

We find that the quality of the comparison of hindcast data and the observed discharge varies for each station (see Figure 3 ). For some station even if the general pattern of higher discharge peaks is the same in observation and simulations, there is large difference in the absolute number. SHOULD WE CORRECT FOR BIAS USING OBSERVED DATA?HOW? Figure below show a test possible implementation of bias correction based on observed data.

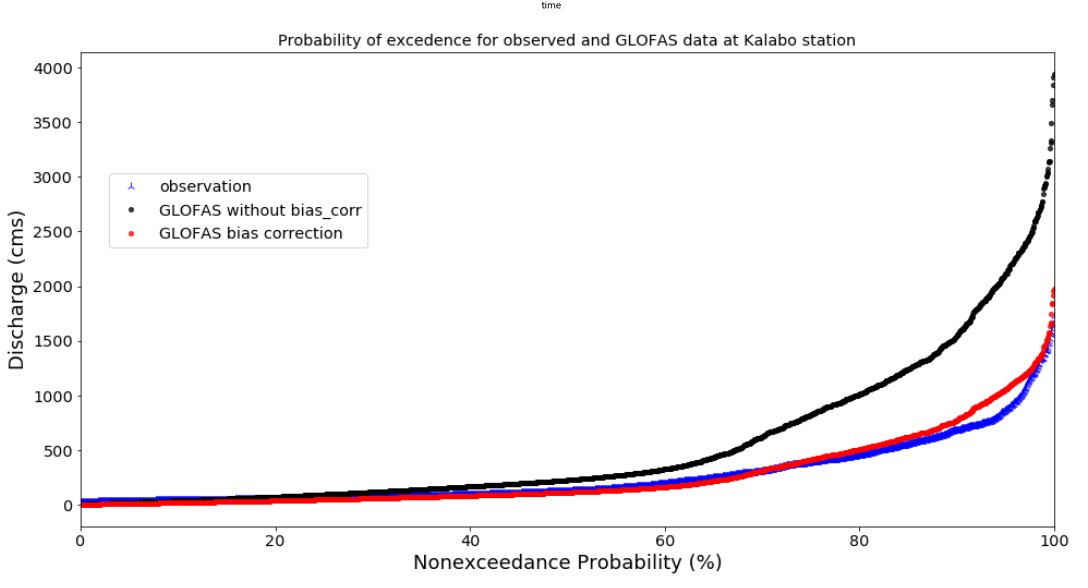


Figure 5 Plot of the discharge from GLOFAS (black markers) and observations (blue) for Kalabo station against the probability of not exceeding the flood treshold. The red line shows the GLOFAS data after a correction for bias has been applied. The correction in this case is a 0.5 factor qualitatively chosen. Figure 6 Comparison of observed discharge (orange line) with hindcast data (blue line). The Horizontal line are the 90,95 and 98% percentile calculated from the hindcast data. The yellow vertical lines mark the years where a flood has been reported in the area.

## Which trigger Level and What forecast probability should trigger action?

To guide the discussion of selecting a “danger level” we selected a list of possible threshold values which are defined based on percentile and flow return period, for each GLOFAS station in ZAMBIA. This thresholds are defined based on the ensemble median values of glofas simulation with a time step=0 and used as a proxy for the amount of discharge that causes a flood. The list of the selected threshold are:

* 98 percentile,
* 95 percentile,
* 90 percentile,
* 5 Year return Period,
* 10 Year return Period,
* 15 Year return Period and
* 20 Year return period.

AB : is percentile not related to Return period ?

In total 7 x 24(for each GLOFAS station) we defined threshold, this will help provide sufficient information for the discussion of the technical committee.

A close up of a map

Description generated with high confidence A picture containing text

Description generated with high confidence

A screenshot of a cell phone

Description generated with high confidence

For FBF activation our interest is to test the forecasting skill of GLOFAS with a lead time that gives as a window of opportunity to take actions, thus we did skill assessment of GLOFAS simulation with time step of 4 and 7. For each Threshold we calculate probability (based on frequency) for false alarm and hit, this was done by

* first comparing the maximum simulated flows against threshold (glofas signal to detect flood) for each year
* then we compared glofas signal with impact data. If impact data showed a flood was reported for that specific period a positive glofas signal is translated as hit. If glofas missed the flood event it will be a miss. On the other hand if Glofas indicated a positive signal (flooding) whereas no impact was reported for that specific year glofas signal is interpreted as false alarm
* For each GLOFAs Signal probability was calculated by comparing the % of ensemble members indicating the identified signal.

For humanitarian actors the most interesting indicator for forecast verification is the false alarm ratio (FAR) which is defined as the number of forecast-based actions that were not followed by a flooding, divided by the total number of actions that were triggered by the system( False alarm + Hit Rate). False alarm represents the proportion of actions that are taken “in vain”. Following the methodology in Weeink (2010), we can compute a number of scores to assess the skill of a forecast method. the False Alarm Ratio (FAR), the Probability of detection (POD) and the critical success index (CSI) (WMO, 2007). Which are defined as follows:

Where hits are events identified as above the threshold in the model and corresponding to a flood report for the same time and district. Misses are reported floods event that are not showing as above threshold in the model and false alarms are events above threshold in the model but not corresponding to a reported flood event for that time or place.

SHOULD WE ONLY USE FORECAST WITH PROBABILITY HIGHER THAN X? WHAT IS A SENSIBLE THRESHOLD? HOW DO WE COMBINE THE INFORMATION ON THE PROBABILITY OF THE FORECAST WITH THE PROBABILITY OF DETECTION?

Discussion also in terms of: upstream versus downstream monitoring stations? Possibly better GLOFAS forecast skill e.g. for upstream stations? Influence of dams/irrigation infrastructure etc?

## Possible workflow in the FbF system based on preliminary analysis

## Menu and selection of forecast

## Definition of the level of impact

# Selection of actions

Process of preparation and activation of the EAP

# Monitoring evaluation, responsibility and learning

# National society capacity for early action

# Finances and logistics

# EAP approval and validation with other organization

# Conclusion

# Annexes

References

[1] K. Bischiniotis, B. van den Hurk, E. Coughlan de Perez, E. Zsoter, M. Grillakis, J. Aerts, Evaluation of ensemble streamflow predictions and flood warnings in Peru, in: Geophys. Res. Abstr., 2018. https://meetingorganizer.copernicus.org/EGU2018/EGU2018-15318.pdf (accessed May 2, 2018).

[1] M.A. Trigg, C.E. Birch, J.C. Neal, P.D. Bates, A. Smith, C.C. Sampson, D. Yamazaki, Y. Hirabayashi, F. Pappenberger, E. Dutra, P.J. Ward, The credibility challenge for global fluvial flood risk analysis, Environ. Res. Lett. 11 (2016). doi:10.1088/1748-9326/11/9/094014.

[1] GloFAS, Performance analysis of long lead flood forecast information from GloFAS, (2016) 1–29. RIMES

1. Data can be found at [USGS website] [↑](#footnote-ref-2)
2. Roberto Jaramillo in this case [↑](#footnote-ref-3)