

# MANUAL

## Early Warning System for Heavy Rainfall

COVID-19 GLOBAL RESPONSE

Corina Markodimitraki, Melissa El Hamouch and Georgios Roullis

2021.05.31 – Version 1

With support of



# Contents

- Introduction ..... 1
  - Product in a nutshell..... 1
  - What is the Early Warning System for Extreme Rainfall and what can you do with it? ..... 1
  - Who is the Early Warning System for Extreme Rainfall for?..... 1
  - Why use the Early Warning System for Extreme Rainfall?..... 1
- Case study ..... 3
- Data responsibility..... 4
  - Datasets ..... 4
  - Data processing ..... 4
  - Non-discrimination ..... 4
  - Human oversight ..... 4
  - Risks ..... 4
- Requirements ..... 5
  - User knowledge ..... 5
  - Hardware ..... 5
  - Software..... 6
  - Time ..... 7
- The product..... 8
  - What does the Early Warning System for Heavy Rainfall consist of? ..... 8
  - Locations to download the product ..... 8
  - Advantages & Limitations ..... 8
- Setting up the Early Warning System for Heavy Rainfall: a step-by-step guide ..... 10
- Abbreviations ..... 15
- Resources ..... 15

## **Introduction**

Since March 2020 and within the COVID-19 Global Response project, the 510 Data Team of the Netherlands Red Cross (NLRC) has aided the Egyptian Red Crescent by providing a weather forecast-based system that issues early warning prior to heavy rainfall. This manual provides the information that one needs to know when planning on setting up and using this framework.

### **Product in a nutshell**

The framework presented in this manual consists of programming scripts and an online dashboard that support the collection of weather forecast data and analysis thereof for the early detection and warning of heavy rainfall.

### **What is the Early Warning System for Extreme Rainfall and what can you do with it?**

The 510 Data team has set up Early Warning System for Extreme Rainfall to support the Egyptian Red Crescent. The system assesses the weather forecast and issues warnings in geographical locations that are at risk from heavy rainfall. This is done in the following way:

1. By collecting information on past rainfall occurrences and setting a threshold.
2. By collecting information on upcoming rainfall occurrences and comparing them to the set threshold. Any predicted rainfall that measures above the set threshold is marked as dangerous and the early warning system is activated.

The predictions can be visualized in QGIS or on an interactive dashboard that does not require specific skills other than basic computer and internet skills and is easily interpretable by the end user.

### **Who is the Early Warning System for Extreme Rainfall for?**

The warning system can be used by the disaster management team of any National Society (NS) or other non-governmental organisation looking to quickly evaluate the possibility of extreme weather in the form of heavy rainfall, and organize their early response. Until now the 510 Data Team has aided the Egyptian Red Crescent to assess heavy rainfall.

### **Why use the Early Warning System for Extreme Rainfall?**

Heavy rainfall can lead to soil erosion, dam failure, water pollution, destruction of roads, housing and farms and cause human, livestock and mineral wealth losses. The warning system presented in this manual offers an open source, fast and automated way of early

prediction of heavy rainfall. The system can help disaster management teams, with fast decision making and response to help populations in the severely endangered areas. The warning system repeats the forecasting multiple times within 24 hours, to stay up to date.

## **Case study**

The Early Warning System for Extreme Rainfall is a tool that is designed to notify a NS that a certain geographic location is at risk of receiving heavy rainfall. The tool helps NS predict potential disasters, such as flooding, and allows for proper emergency response actions to be put in place. With the help of the tool, the NS can assure that its branches in high flood risk areas are equipped with the necessary gear and that evacuation and rescue plans are communicated.

To date, the Early Warning System for Extreme Rainfall tool has been supporting the Egyptian Red Crescent but has also become an interest to other NS, such as the Jordan Red Crescent Society. The tool uses data of past and upcoming rainfall occurrences, making it easily adaptable to fit various contexts.

## **Data responsibility**

### **Datasets**

The Early Warning System for Extreme Rainfall tool uses data of past and upcoming rainfall occurrences of a specific geographic location. The predicted meteorological data could be found online or through local weather forecasting channels. The tool could also harbor data on soil composition, roads, location of dams, topography and others.

### **Data processing**

The tool processes the data using programming scripts and an online dashboard that supports the analysis of the collected information. The tool is fully automated and repeats the forecasting multiple times in order to stay updated. Since various browsers could be used to collect, analyze and visualize the data, it is important to note that local knowledge from the NS should be taken into account before any geographic information is made public.

### **Non-discrimination**

The tool is non-discriminatory.

### **Human oversight**

In order to set up the programming script, an individual will need to have expertise in Python. Though the end user, does not need any programming experience, is still required to understand the output of the early warning system in order to make future decisions.

### **Risks**

The foreseeable risks of such a tool are dependent on the availability of predictions of meteorological data as well as the accuracy of the predictions. The information will need to be verified with local knowledge that can account for any missing information that is relevant to the specific geographical context.

## Requirements

### User knowledge

In this section we describe the skills that someone would need to set up the early warning system for heavy rainfall presented in this manual. In addition to the skills mentioned below, it is crucial the user has the following skills:

- Analytical thought
- Problem-solving attitude
- Data responsibility (see also section “Data responsibility” of this manual)
- GIS skills

To collect and analyze rainfall data the user will need the following skills:

- **Computer skills:** It is essential the user has advanced computer skills including internet, database and spreadsheet experience.
- **Python:** it is necessary that the user has knowledge and experience with Python to execute and adjust Python scripts.
- **GIS skills:** in case the user wants to visualize the collected data, it is necessary that they have knowledge and experience with QGIS desktop (or any other similar GIS program), in order to load the final output of the script.

### Hardware

Below you can find the hardware requirements for the Early Warning System for Heavy Rainfall explained in this manual.

#### To collect and analyze the data:

- **Python** (versions: 2.7.X, 3.6.X):
  - Processor: Intel Atom® processor or Intel® Core™ i3 processor
  - Disk space: 1 GB
  - Operating systems: Windows\* 7 or later, macOS, or Linux

**Note:** for a specific command in one of the Python scripts, Linux (Ubuntu) operating system is preferred (see step 4 in “Setting up the Early Warning System for Heavy Rainfall: a step-by-step guide” of this manual).

- **Excel Microsoft Office** (2019):
  - Processor: Windows OS: 1.6 GHz, 2-core / macOS: Intel processor
  - RAM: Windows OS: 4 GB RAM; 2 GB RAM (32-bit) / macOS: 4 GB RAM
  - Storage: Windows OS: 4 GB available disk space / macOS: 10 GB of available disk space
  - Screen resolution: Windows OS: 1280 x 768 (32-bit requires hardware acceleration for 4K and higher) / macOS: 1280 x 800
  - Graphics: Windows OS: DirectX 9 or later, with WDDM 2.0 or higher for Windows 10 (or WDDM 1.3 or higher for Windows 10 Fall Creators Update) / macOS: No graphics requirements.

### **To visualize the data:**

- For **QGIS Desktop**:
  - Processor: minimum Core i3 2.7 Ghz; recommended Core i7 3.5 Ghz
  - RAM: minimum 2 GB; recommended 8GB or more
  - Disk space: minimum 500 GB SATA; recommended SSD de 128 GB or 500 GB SATA
  - Graphics: minimum 1GB RAM; recommended 2 GB RAM (NVIDIA Geforce)
  - Operating systems: Windows 7-10, Mac OSX, Linux, Unix, or Android

## **Software**

Below we list the software required for the Early Warning System for Heavy Rainfall explained in this manual.

### **To collect, analyze and visualize the data**

- Internet browser (for computers) such as



- Mozilla Firefox  
[www.mozilla.org/en-US/firefox/new/](http://www.mozilla.org/en-US/firefox/new/)
- Google Chrome  
[www.google.com/chrome/](http://www.google.com/chrome/)
- Spreadsheet-based program such as
  - Excel (Microsoft Office)  
[www.microsoft.com/en-us/microsoft-365/excel](http://www.microsoft.com/en-us/microsoft-365/excel)
- Installed Python  
<https://www.python.org/>
- Installed QGIS  
<https://www.qgis.org/en/site/>

## Time

Here we indicate which parameters influence the duration of setup and usage of the Early Warning System for Heavy Rainfall, and state a real-life example of setup by a NS.

The duration of setting up and using the Early Warning System for Heavy Rainfall is the total duration of the following steps: collecting the geographical area shapefiles, collecting the historical rainfall data, running the script that calculates the threshold, collecting the forecasted data, running the script that compares the predicted values to the set threshold, and visualizing the output. All this depends on lead time (how many hours ahead are the predictions for) and user Python programming skills.

For reference, the 510 Data Team of the NLRC department spent approximately 1 month to set up the model for the Early Warning System for Heavy Rainfall. As the script is run automatically once it is setup, it takes the end user 5 minutes to open the end results in QGIS.

## The product

### What does the Early Warning System for Heavy Rainfall consist of?

The prediction of heavy rainfall can be achieved by using all of the following:

- Python scripts
- (Optionally) QGIS or any other GIS program

### Locations to download the product

Here we provide the links and locations to find the tools needed to set up and use the Early Warning System for Heavy Rainfall.

- Github page containing the needed scripts: <https://github.com/rodekruis/Rainfall-only-trigger>. Note: only the following scripts are used in the pipeline described in this manual:
  - `calculate_thresholds_returnperiods.py`
  - `rainfall_only_pygrib.py`
- (Optional) GIS visualization tool such QGIS: [www.qgis.org](http://www.qgis.org)

### Advantages & Limitations

There are several advantages and limitations linked to using the Early Warning System for Heavy Rainfall. Here we briefly discuss these.

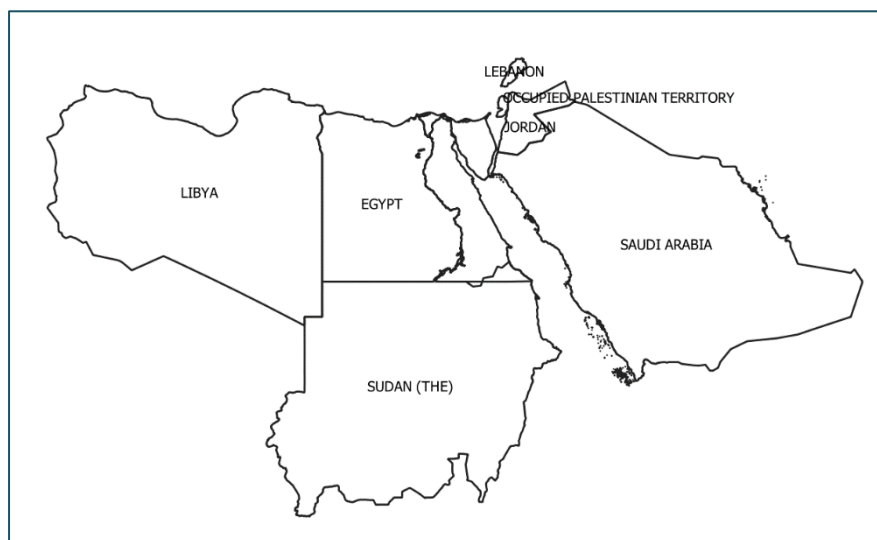
- The user is depended on the historical meteorological data being available. This could translate to a bottleneck in the pipeline if there are no data available.
- The user is depended on the forecasted meteorological data being available on <https://www.ncei.noaa.gov/data/global-forecast-system/access>. This could translate to a bottleneck in the pipeline if there are no data available.
- The pipeline described in this manual requires expertise in Python to setup, which could be a limitation if there is no expert present in the organization aiming to use the Early Warning System.
- The end user of the Early Warning System is not required to have experience using Python or GIS skills.
- The pipeline is fully automated, which provides the latest meteorological information to the end user.

- The pipeline is adjustable, meaning that the user can choose the lead time (time forecasted) and the geographical location they wish to include.

# Setting up the Early Warning System for Heavy Rainfall: a step-by-step guide

## 1. Collect the shapefiles for the region of interest

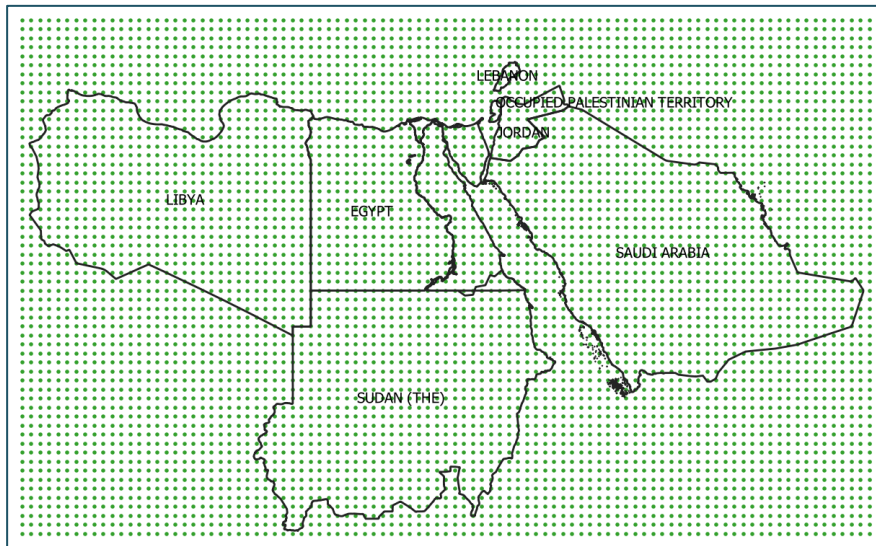
The shapefiles are necessary to visualize the areas in danger of heavy rainfall in QGIS or another GIS-based program. Shapefiles for most country boundaries can be found on <https://data.humdata.org/dataset>. The files can be loaded onto a GIS-based program such as QGIS, to visualize the area of interest, as seen in **figure 1**.



**Figure 1** • Visualization of a shape file by using the QGIS program. The area of interest is depicted.

## 2. Collect historical forecasted meteorological data

Manually collect historical meteorological data from a valid website such as <https://registry.opendata.aws/noaa-gefs-reforecast/>. Define the time period of interest and download the data for that period. As an example, the early warning system for heavy rainfall in Egypt uses the data of a 30-year period: 1991-2020. The downloaded file will contain meteorological data such as sun radiation, wind, precipitation for the whole globe. Save the files in the following path, as advised on the Github page: `Rainfall-only-trigger-master/calculate_thresholds/data/` The data can be visualized by using QGIS as seen in **figure 2**.



**Figure 2** • Visualization of the historical forecast data file by using the QGIS program. The data are projected onto the area of interest. Each green pixel represents data covering an area of 50 km<sup>2</sup>.

### 3. Calculate the threshold of the most extreme historical forecasted rainfall data

For this step, the Python script `calculate_thresholds_returnperiods.py` is needed and can be found on the Github page, see section “The product” of this manual. In this step, a rainfall value is defined as the threshold to which the present forecast data will be compared. If the present forecast data values in millimeters (mm) are equal or higher than the set threshold, they are marked as extreme. To calculate the threshold, the downloaded historical data of step 2 are filtered for area, type of meteorological measurement and lead time of interest, and these can be adjusted. Please note that the calculations are for each data “pixel” representing an area of 50 km<sup>2</sup> as shown in **figure 2**. Briefly, the script `calculate_thresholds_returnperiods.py`:

- a. Filters the data for the rainfall data by “total precipitation”.
- b. Filters the data for the geographical area of interest by latitude and longitude
- c. Keeps forecasted data for 3-, 5- and 7-day lead time.
- d. Applies the Generalised Extreme Value (GVE) distribution model to the block maximas of the data to calculate the thresholds in terms of return periods. In the example of Egypt, the block maximas are annual maximas of the 30-year period. From this, the thresholds of 2, 5, 10, 20, 50 and 100 year-return periods are defined. These threshold values are then saved as a csv file (**Figure 3**).

	lat	lon	threshold_99perc	2yr_threshold	5yr_threshold	10yr_threshold	20yr_threshold	50yr_threshold	100yr_threshold	forecast_time
0	8	9	7.970.800.028	1.866.207.366	2.384.868.944	270.429.348	2.993.903.207	3.345.811.777	3.593.526.098	3
1	8	9	6.029.453.125	1.238.384.053	1.691.345.586	2.043.637.938	2.427.057.745	2.999.289.486	3.492.343.481	5
2	8	9	5.585.471.992	1.062.939.928	1.305.545.752	1.448.465.776	1.573.655.634	1.720.021.232	1.819.175.051	7
3	8	10	8.060.703.125	1.701.070.239	2.382.954.044	2.924.244.516	3.522.939.825	4.432.914.557	5.231.035.618	3
4	8	10	6.185.774.961	1.339.098.014	177.520.218	20.662.102	2.347.063.095	271.310.284	2.989.262.407	5
5	8	10	5.625	1.133.861.662	1.312.135.989	1.397.276.186	146.072.655	1.522.701.232	1.557.750.053	7
6	8	11	66.875	1.516.109.681	2.012.393.632	2.314.122.709	2.584.964.042	2.910.399.534	3.136.943.928	3
7	8	11	5.125	1.115.927.577	1.434.631.412	1.627.795.315	1.800.767.849	2.008.047.094	215.195.357	5
8	8	11	4.709.140.625	1.041.193.767	1.271.942.594	1.414.380.882	154.375.607	1.701.277.817	1.812.381.391	7
9	8	12	709.375	1.485.336.454	1.926.600.803	2.205.171.698	2.462.650.433	2.782.338.024	3.012.236.383	3
10	8	12	5.680.898.437	1.181.030.511	1.477.694.266	1.702.204.899	1.941.348.283	2.289.674.434	258.276.261	5
11	8	12	5.560.703.125	1.070.149.293	1.290.752.386	1.435.927.417	1.574.525.134	1.752.974.327	1.885.996.864	7
12	8	13	7.614.492.187	1.580.010.664	2.051.564.917	2.312.183.784	2.529.477.437	2.769.908.876	2.924.069.164	3
13	8	13	6.199.999.809	1.196.834.457	1.465.036.481	1.626.924.353	1.771.428.087	1.943.974.667	2.063.342.933	5
14	8	13	5.900.000.095	1.136.149.603	1.320.214.158	1.408.651.745	1.474.864.549	1.539.869.185	1.576.822.411	7
15	9	9	9.091.375.294	1.925.534.256	2.470.659.675	2.766.716.601	3.010.249.239	3.275.693.051	3.443.347.233	3
16	9	9	7.378.850.088	1.475.155.842	1.805.911.622	1.955.875.388	2.063.285.637	2.163.688.927	2.217.995.998	5
17	9	9	6.639.726.562	1.224.376.854	1.545.366.114	1.732.125.808	1.894.149.049	2.081.551.047	2.207.150.611	7
18	9	10	6.557.695.312	1.409.656.808	1.761.398.771	1.956.646.829	2.119.984.213	2.301.391.536	2.418.143.007	3
19	9	10	5.099.999.905	1.059.883.316	1.338.574.854	1.490.777.248	1.616.515.887	1.754.224.828	1.841.619.758	5

**Figure 3** • The output of the `calculate_thresholds_returnperiods.py` script are threshold values per return period. The csv file is opened in excel.

#### 4. Collect the most recent forecasted meteorological data

For this step, the Python script `rainfall_only_pygrib.py` is needed and can be found on the Github page, see section “The product” of this manual. Briefly, the script `API_download_rainfall_modified2.py`:

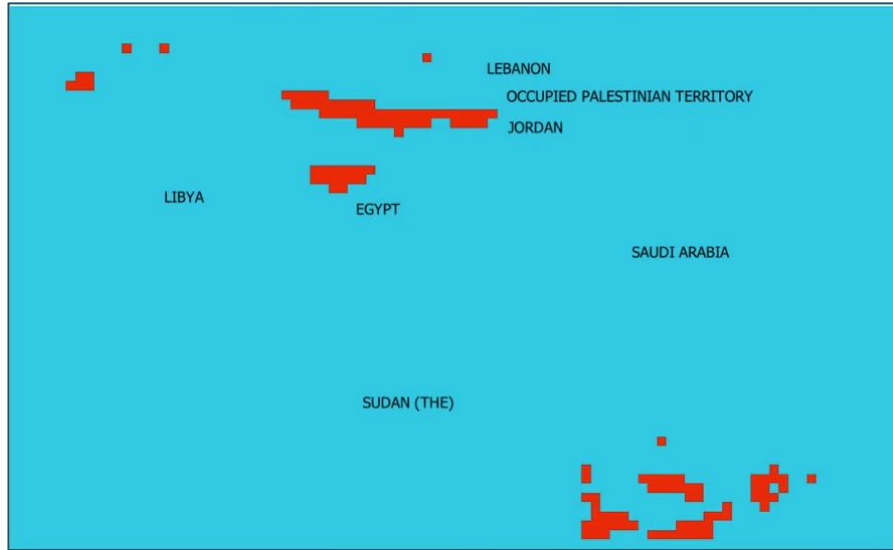
- a. Defines the geographical area of interest from the shapefiles mentioned above, by creating a bounding box. This will be used later on to compare the forecasted rainfall data with the thresholds set earlier.
- b. Collects the latest forecast data from <https://www.ncei.noaa.gov/data/global-forecast-system/access/grid-004-0.5-degree/forecast/> which are generated based on the model Global Forecast System by the National Oceanic and Atmospheric Administration (NOAA). The forecast model is run four times in 24 hours: at 00:00, 06:00, 12:00, 18:00 (four run cycles) and predicts up to 16 days ahead. The downloaded file will contain meteorological data such as sun radiation, wind, precipitation for the whole globe.
- c. Extracts the precipitation data from the downloaded grib2 file. The script extracts the data for lead time=3, 5 and 7 days and combines them into a single file. The lead time parameter can be adjusted.

Important note: In line 143 of the Python script, the command `wgrib2` is used. This extracts the precipitation data ACP from the downloaded grib2 file containing the forecasts. The command works best when executed in a Linux (Ubuntu) environment.

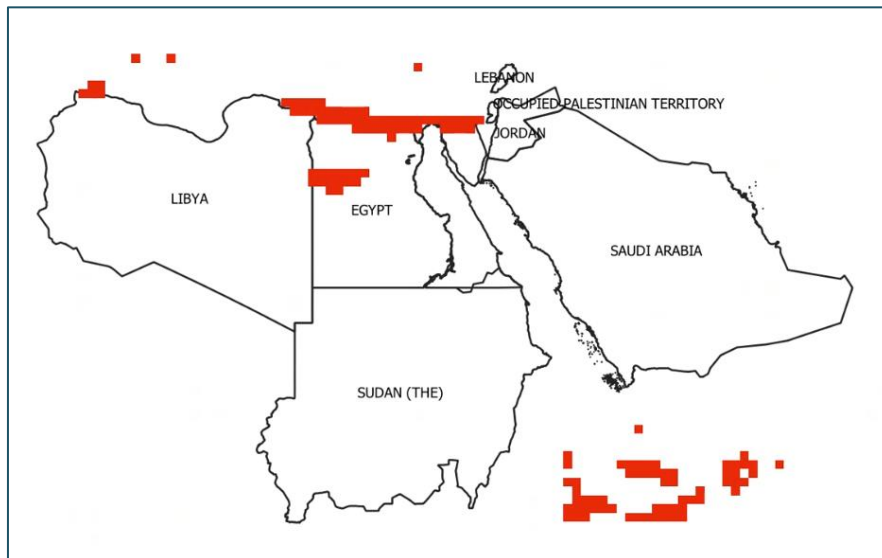
## 5. Compare the forecasted rainfall data to the threshold set by historical rainfall data

For this step, continue using the Python script `rainfall_only_pygrib.py`. The chosen threshold for the comparison depends on the National Society and its capacity for early action. Generally, the threshold defined by a 5-year return period is lower than the threshold defined by a 100-year return period. This means that the chance of a forecasted rainfall occurrence being equal or higher to the 5-year return period threshold is relatively high. This would result in the system giving an early and calling for early action more often than if the threshold used was defined based on a 100-year return period. The return period on which the threshold was calculated is defined in the script and can be adjusted. Briefly, the script:

- a. Extracts the data from the forecast file that fall into the bounding box (defined in step **4a**) so that only these are compared to the threshold. The new data frame will contain the data for the area of interest.
- b. Compares the set threshold (step **3**) to the forecasted rainfall data of the area of interest for each lead time. The rainfall values used are the average values of all run cycles of the day.
- c. Outputs a raster file (tiff) containing binary values representing the result of the comparison (**Figure 4**). A value of 0 means that the compared forecasted data was lower than the threshold value and there is a low possibility of an extreme rainfall occurrence. A value of 1 means that the compared forecasted data was equal or higher than the threshold value which indicates there is a high possibility that the forecasted rainfall will be extreme. The generated raster file can be projected on the area in QGIS or other GIS-based program, indicating the regions where there is a high chance of extreme rainfall (**Figure 5**).



**Figure 4** • The output of the `rainfall_only_pygrib.py` script is a raster file containing the binary values indicating extreme rainfall data. Non-extreme values are marked with a binary value of 0 and are depicted as cyan in this figure. Extreme values are marked with a binary value of 1 and are depicted as red. The raster file is opened in QGIS.



**Figure 5** • Extreme values of the raster file are marked with a binary value of 1 and are depicted as red on top of the region of interest. The cyan values are hidden in this figure. The raster file is opened in QGIS.



## Abbreviations

NLRC	Netherlands Red Cross
SARS-CoV-2	Severe Acute Respiratory Syndrome coronavirus 2
COVID-19	The disease caused by the SARS-CoV-2 virus
NS	National Society/Societies
hrs	hours
API	Application Programming Interface
GIS	Geographic Information System
GVE	Generalised Extreme Value

## Resources

- The model applied to the 99<sup>th</sup> quantile of historical rainfall data to calculate the threshold value  
<https://kikocorreoso.github.io/scikit-extremes/Module%20models.classic.html>
- The wgrub2 package used in the `rainfall_only_pygrib.py` Python script  
<https://www.cpc.ncep.noaa.gov/products/wesley/wgrib2/>
- The link to the Egypt early warning dashboard is available upon request