**Regional Frequency Analysis (RFA) for rainfall extremes with probabilistic neural networks: enhancement projects**

**Brief resume of the context and scope of the study**

Rainfall (or precipitation) events are extremely variable in time and space. In fact, precipitation with different durations (e.g., rainfall that lasts 1 hour and rainfall that lasts 24 hours) are due to very different climatic forcing (e.g., advection or convection). Also, according to local climate, precipitation events with the same duration can be very different (e.g., in the Alps the events are different than in the Po Plain).

We are interested in extreme precipitation: measures of Annual Maximum rainfall for some durations (1, 3, 6, 12, 24h) are present in some measurement stations (AMS, Annual Maximum Series). In general, Annual maxima follow a different probability distribution at each station and each duration.

Where AMSs with good length are available, local frequency analysis is performed: a suitable probability distribution is fitted on the data (e.g., Gumbel or Generalized Extreme Value distribution). Where they are not available, some relation is needed to obtain the probability distribution using the data from the measure stations. The neural network models can be used to obtain this relation. During training, neural networks should learn how to derive probability distributions for the AMS with different durations using morpho-climatic descriptors.

The descriptors are the following:

|  |  |  |
| --- | --- | --- |
| **Descriptor** | **Description** | **Information origin** |
| **1** | mean altitude within a circle with … radius whose center is the gauged location | MERIT DEM |
| **2** | standard deviation of the altitude within a circle with … radius whose center is the gauged location | MERIT DEM |
| **3** | mean slope (I.e., ration between vertical and horizontal distance) within a circle with … radius whose center is the gauged location | MERIT DEM |
| **4** | standard deviation of the slope within a circle with … radius whose center is the gauged location | MERIT DEM |
| **5** | mean aspect (I.e., direction of maximum slope) within a circle with … radius whose center is the gauged location | MERIT DEM |
| **6** | the standard deviation of the aspect within a circle with … radius whose center is the gauged location | MERIT DEM |
| **7** | minimum distance from the Adriatic coast | MERIT DEM |
| **8** | mean elevation within the distance between the gauged location and the Adriatic coast | MERIT DEM |
| **9** | standard deviation of elevation within the distance line between the gauged location and the Adriatic coast | MERIT DEM |
| **10** | Maximum elevation within the distance line between the gauged location and the Adriatic coast | MERIT DEM |
| **11** | minimum distance from the Tyrrhenian coast | MERIT DEM |
| **12** | mean elevation within the distance between the gauged location and the Tyrrhenian coast | MERIT DEM |
| **13** | standard deviation of elevation within the distance line between the gauged location and the Tyrrhenian coast | MERIT DEM |
| **14** | Maximum elevation within the distance line between the gauged location and the Tyrrhenian coast | MERIT DEM |
| **15** | Mean annual precipitation | BIGBANG dataset |
| **16** | Mean annual snow precipitation | BIGBANG dataset |
| **17** | Standard deviation of annual precipitation within the 1919-2019 record period | BIGBANG dataset |
| **18** | Standard deviation of annual snow precipitation within the 1919-2019 record period | BIGBANG dataset |
| **19** | Longitude | Mazzoglio et. al |
| **20** | Latitude | ... |

**DATASET 1**

1. **AMS\_descritt\_noSM\_meltD\_adim.csv**: stations for training, validation, test.
   1. AMS   
      Annual Maximum Series of rainfall values: one value for each duration (1, 3, 6, 12, 24 hours) for each year of measurement. Target variable, whose probability distribution (for a given duration) is searched.   
      It is dimensionless (I.e.: the rainfall depth, in *mm*, is divided for the mean for each station and each duration, which is column *MeanIdD*).  
      To obtain the original AMS column (I.e., the one in millimeters), just multiply *AMS* column by *MeanIdD* column.  
      In this CSV, every row is a record, meaning that every row is a measure of maximum rainfall depth for a specific year and a specific duration
   2. Columns to use as morphoclimatic descriptors:   
      'HMerit\_mean(m)', 'HMerit\_stdev(m)', 'Slope\_mean(deg)', 'Slope\_stdev(deg)', 'Aspect\_mean(deg)', 'Aspect\_stdev(deg)', 'MinDistAdriaticC\_(km)', 'OrogrAdri\_mean (m)', 'OrogrAdri\_stdev(m)', 'OrogrAdri\_max(m)', 'MinDistTirrenicC\_(km)', 'OrograTirr\_mean(m)', 'OrograTirr\_stdev(m)', 'OrograTirr\_max(m)', 'MAP(mm)', 'MASnowP(mm)', 'VarClimPT(mm)', 'VarClimSowP(mm)', 'X', 'Y'
   3. Column *'duration[h]':*  
      we could decide to model separately every duration
   4. Column *‘ID’*:  
      code to identify each station
2. **gumMap\_statbench\_Gumfit\_NEW.csv**:   
   the first 100 occurences of the *ID* column are the identification numbers of the test stations. To obtain test set: select rows where column *ID* of CSV (1) is equal to the first 100 occurences of *ID* column in CSV (2).  
   The resting rows from CSV (1) are the training set (I.e.: training + validation set, if a validation set is needed when training the neural networks)

**DATASET 2 (*Files\_CNN\_Michele*)**

The following are raster files that just represent some geomorphic descriptors for the Panaro river catchment (all with 1km x 1km cells):

1. The altitude: file .tif named *Merit\_Panaro.tif*, resolution ~90m x 90m
2. The slope: file .tif named *Slope\_merit\_Panaro.tif*, resolution ~90m x 90m
3. The aspect: file .tif named *Aspect\_merit\_Panaro.tif*, resolution ~90m x 90m
4. The MAP: file .tif named *MAP\_mean1951\_2019\_Panaro.tif*, resolution 1km x 1km
5. The MASnow: file .tif named *MASnow\_mean1951\_2019\_Panaro.tif*, resolution 1km x 1km
6. MAP standard deviation file .tif named *MAP\_std1951\_2019\_Panaro.tif,* resolution 1km x 1km
7. MASnow Standard deviation file .tif named *MAP\_std1951\_2019\_Panaro.tif*, resolution 1km x 1km
8. The rainfall measured stations in the Panaro river catchment area: .csv file named *Panaro\_rainfall\_stations.csv.* To be used to select the stations from the CSV file (1) of project (1).   
   Graphical file is *Panaro\_rainfall\_stations.shp*