MONTHLY FLOOD FREQUENCY REGIONALIZATION FOR COMPREHENSIVE FLOOD DAMAGE ASSESSMENT TO CROPS

Charlie Dayane Paz Idarraga [1], Anna Rita Scorzini [2], Daniela Molinari [1]

(1) Department of Civil and Environmental Engineering, Politecnico di Milano (2) Department of Civil, Environmental and Architectural Engineering, University of L'Aquila

KEY POINTS

- *Plant vulnerability varies across phenological phases, establishing an intrinsic link between crop damage and month of flood occurrence*
- *Flood risk assessment in the agricultural sector requires transitioning from annual probability-based inundation scenarios to monthly hazard estimations*
- *The results of the approach enable the assignment of distinct monthly flood probabilities to all relevant catchments in a river district, thereby contributing to a comprehensive flood damage assessment to crops*

1 INTRODUCTION

Quantitative flood risk assessments rely on damage models, which relate information on flood hazard and vulnerability features of exposed assets to estimate expected losses. In contrast to other sectors, crop damage depends not only on typical hazard variables (including water depth, flow velocity, inundation duration, water salinity and yield of sediments and/or contaminants), but also on the month of flood occurrence (Molinari et al., 2019), since plant vulnerability changes over the different phenological phases that are strictly connected with the seasonality of crop production. This implies the necessity of shifting from the traditional representation of inundation scenarios based on annual probability to monthly-based hazard estimations. Such knowledge is indeed essential for assessing the expected annual damage to crops, which is calculated as the sum of incremental probability of occurrence over different flooding scenarios times the corresponding weighted average of possible losses occurring in the different months of the year, with weights given by the probability of flooding in each specific month. A detailed understanding of those seasonal flood patterns in the different sub-catchments of the basin, including un-gauged ones, can be then considered crucial when risk assessment is carried out at the river basin level.

In this study, we present a clustering approach to flood frequency regionalization applied to the Po River District in Northern Italy, within the risk mapping process required by the European Floods Directive. Previous studies have utilized clustering analysis to determine flood regimes, as is the case of *Beurton & Thieken* (2009). However, the nationwide scale of their model and its exclusive focus on seasonal flood distribution, prevent obtaining a more detailed characterization at the subbasin scale and a spatial transfer to other contexts. In our study, we aim to account for the complex climatic and topographic conditions of the region of interest by adding physical criteria to the statistical analysis of floods. This allows the applicability of the approach to other geographical contexts given the representativeness of the case study.

The observed monthly flow data from over 100 gauging stations are used to identify the seasonal probability classes. The catchment areas of these stations are characterized according to morphometric and land use features, in addition to the presence of upstream dams. These factors are then considered in a clustering procedure which form the basis for a monthly flood frequency regionalization, allowing to assign a hydrological pattern to ungauged basins based on similarity criteria.

2 STUDY AREA AND AVAILABLE DATA

The area of interest covers the Po River District in Northern Italy, which is characterized by significant climate variability. The atmospheric fluxes from the Mediterranean Sea and the intricate topography, which encompasses the southern range of the western and central Alps, as well as the northern Apennines range, are the main factors influencing the climate in the area (*De Michele & Rosso*, 2002).

Over 100 hydrometric stations with minimum 20 years of observed monthly data are available in the regions

of Lombardy, Piedmont, and Emilia-Romagna (Figure 1). The corresponding time series and Digital Elevation Models are obtained from the regional geo-portals, whereas land use is extracted from the CORINE land cover provided by the Copernicus Land Monitoring Service. In addition, information regarding the location and characteristics of upstream dams is obtained from the dam register provided by the Italian General Directorate for Dams and Water Infrastructures.

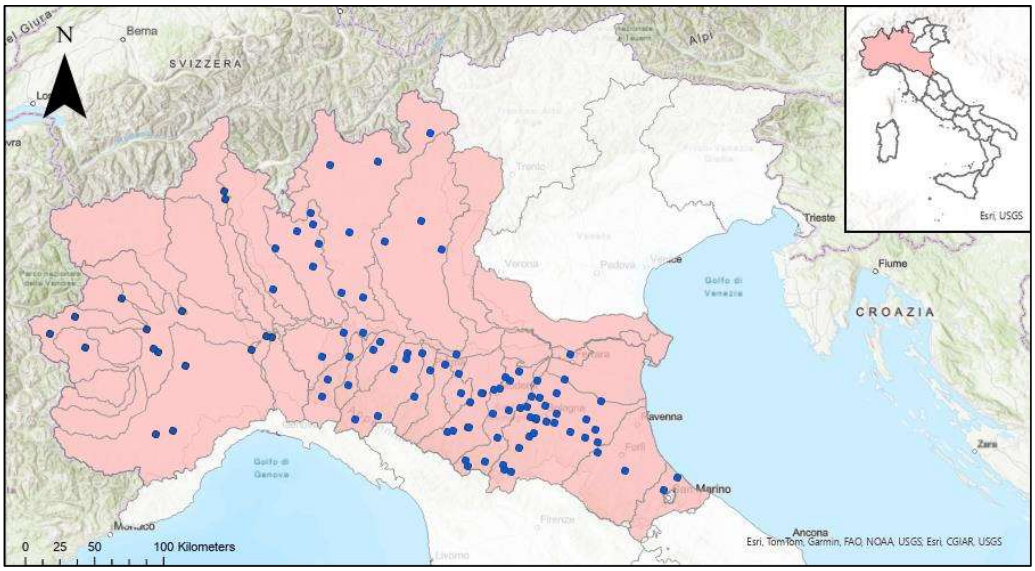

Figure 1. Study area: the Po River District

3 METHODOLOGY

The main cores of the methodology are: 1) selection of gauging stations and identification of their seasonal flood probability classes based on available time series, 2) delineation and characterization of the subbasins, 3) clustering and sensitivity analysis, and 4) assignment of hydrological patterns to ungauged basins (regionalization of monthly flood frequency). The detailed framework of the methodology is presented in **Figure 2**.

3.1 Seasonal flood probability classes

The selected gauges cover a minimum of 20 years of monthly mean discharge and/or hydrometric data. At each station and for each year, the maximum value and month of occurrence are obtained to calculate the probability of having a peak flow for each month. Thus, each station/subbasin is characterized by 12 probability values (January to December). The months are then grouped according to their season (winter (January to March), spring (April to June), summer (July to September) and autumn (October to December)), to identify which season(s) correspond to the highest probability of peak flow.

3.2 Delineation and characterization of the subbasins

Each station is considered as an outlet section to delineate its corresponding subbasin, which has been then characterized by a series of morphometric (slope, mean height, drainage density, etc.) and land use (permeable and impervious areas) features, as well as by a binary indicator for the presence of upstream dams (water reservoirs). Continuous descriptive features are discretized into a suitable number of classes to work exclusively with categorical variables in the clustering process.

3.3 Clustering

Unsupervised hierarchical clustering, a standard method in multivariate statistics (e.g., *Murtagh & Contreras*, 2012), is employed in this study to group the subbasins based on their similarities in flood generation mechanisms and seasonality, involving both physical (morphometric and dam conditions) and

statistical (seasonal flood probability) criteria. The method constructs a hierarchy of clusters by iteratively merging them based on their similarity, creating a tree-like structure known as a dendrogram. The Jaccard's index is selected as the distance metric, as proper for categorical data. Furthermore, a sensitivity analysis is performed to evaluate the stability and reliability of the identified clusters. This involves systematically varying clusterization parameters and assumptions, such as the definition of the seasons and the number of features considered in the process.

3.4 Monthly flood frequency regionalization

Identifying homogeneous clusters allows hydrological patterns to be assigned to ungauged basins using few key and simple variables. Thus, the basins in the Po River district are characterized by a monthly flood frequency pattern.

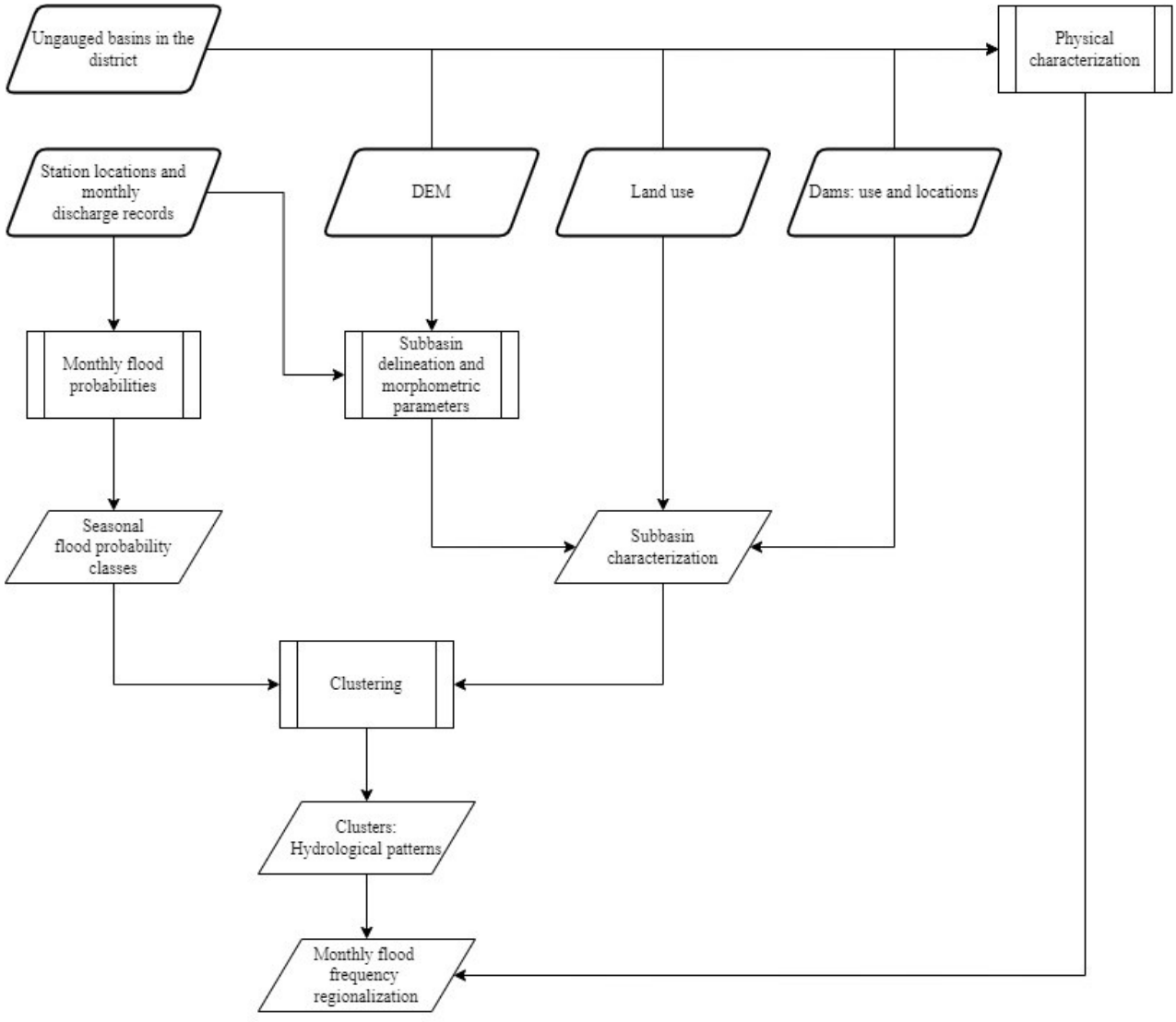

Figure 2. Methodological framework.

4 RESULTS AND DISCUSSION

Monthly probability analysis shows that each station is associated with one or two peak flood probability maxima. Most of the gauging stations in the area presents probability maxima in autumn-winter and spring periods. These results confirm the heterogeneity of meteorological processes within the district. Moreover, it is possible to recognize the heterogeneity of the region in terms of morphometric and land use features. As an

illustrative example, Figure 3 presents the outcomes of a clustering into five classes, delineating distinct hydrological patterns for both the Po River and the Apennine and Alpine basins. The computation of Cramer's V, utilized as a measure of association between the identified clusters and the explanatory variables, reveals that the primary determinants influencing the clustering results are the classification of flood seasonality and the presence of dams upstream from the gauging station. Among the morphometric features, which have a comparatively lesser impact, the most significant ones are the mean elevation of the basin and the fraction of impervious area.

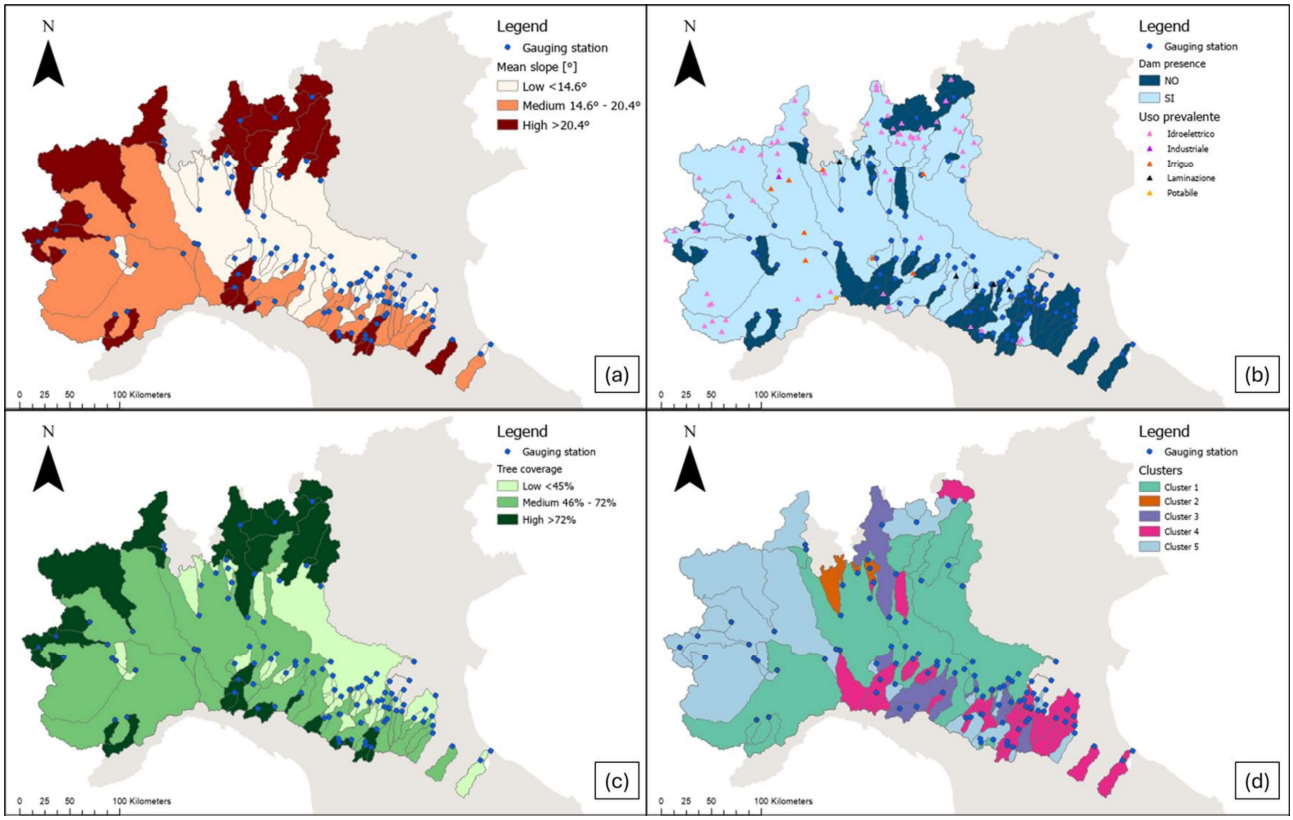

Figure 3. (a) Example of morphometric parameters (b) Presence of dams (c) Land use parameters (d) Clustering results

5 CONCLUSIONS

In this study, we presented a framework for the identification of homogeneous regions in terms of flood generation mechanisms and seasonality based on the use of observed monthly flow data in combination with physical and statistical criteria. The clustering approach can provide a flood frequency regionalization that contributes to the necessary shift from the traditional representation of inundation scenarios based on annual probability to monthly based hazard estimations. This enables the assignment of distinct monthly flood probabilities to all catchments within a district, thereby supporting a comprehensive flood risk assessment for the agricultural sector.

REFERENCES

Beurton, S. & Thieken, A.H. Seasonality of floods in Germany, Hydrological Sciences–Journal–des Sciences Hydrologiques, 2009, 54(1), 62-76.

De Michele, C. & Rosso, R. A multi-level approach to flood frequency regionalisation, Hydrology and Earth System Sciences, 2002, 6(2), 185-194.

Molinari, D., Scorzini, A.R., Gallazzi, A. & Ballio, F. AGRIDE-c, a conceptual model for the estimation of flood damage to crops: development and implementation. Natural Hazards and Earth System Sciences, 2019, 19(11), 2565-2582.

Murtagh, F. & Contreras, P. Algorithms for hierarchical clustering: an overview. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2012, 2(1), 86-97.