The role of susceptibility, exposure and vulnerability as drivers of flood disaster risk at the parish level

Pedro Pinto Santos (pmpsantos@campus.ul.pt)
University of Lisbon Institute of Geography and Spatial Planning: Universidade de Lisboa Instituto de Geografia e Ordenamento do Territorio https://orcid.org/0000-0001-9785-0180

Susana Pereira
University of Lisbon Institute of Geography and Spatial Planning: Universidade de Lisboa Instituto de Geografia e Ordenamento do Territorio

Jorge Rocha
University of Lisbon Institute of Geography and Spatial Planning: Universidade de Lisboa Instituto de Geografia e Ordenamento do Territorio

Eusébio Reis
University of Lisbon Institute of Geography and Spatial Planning: Universidade de Lisboa Instituto de Geografia e Ordenamento do Territorio

Mónica Santos
Centre for the Research and Technology of Agro-Environmental and Biological Sciences, University of Trás-os-Montes and Alto Douro

Sérgio Cruz Oliveira
University of Lisbon Institute of Geography and Spatial Planning: Universidade de Lisboa Instituto de Geografia e Ordenamento do Territorio

Ricardo A. C. Garcia
University of Lisbon Institute of Geography and Spatial Planning: Universidade de Lisboa Instituto de Geografia e Ordenamento do Territorio

Raquel Melo
Universidade de Évora Centro de Geofísica de Évora: Instituto de Ciencias da Terra Polo da Universidade de Evora

José Luís Zêzere
University of Lisbon Institute of Geography and Spatial Planning: Universidade de Lisboa Instituto de Geografia e Ordenamento do Territorio

---

Research Article

Keywords: Floods, damages, regression trees, databases, risk drivers

Posted Date: June 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1649268/v1

Abstract

Fluvial flooding continues to be a process that has a major impact on society, the environment and the economy. Although its triggering factors are natural, the spatial configuration of exposure and vulnerability is expected to play a relevant role in explaining the damage records.

The starting point of this research is the use of existing flood susceptibility, exposure and social vulnerability mapping, produced at the parish level, as input data in a Classification and Regression Trees model. The model uses as a dependent variable, separately, two databases of flood damage: one with the most severe losses from the DISASTER database, and another that sums the DISASTER cases and the lower impact damages.

The results show a quite distinct classification of parishes, whether one database is used or the other. The DISASTER database reveals susceptibility has the most relevant flood risk driver in explaining the damage patterns, while the database with all the cases identifies exposure as the more relevant driver. In the end, the degree of damages has documented in databases is conditioned by the geographical distribution and overlay configuration of the three flood risk drivers.

Finally, the CART classification groups are analyzed at the light of the European Union's Floods Directive areas of significant potential flood risk. This analysis showed that the Directive's parishes are interpreted differently - in terms of their positioning in face of the risk drivers – which is explained by the use of distinct impacting-criteria in the construction of the flood damage databases.

1 Introduction

The recent report on human losses from disasters caused by natural hazards, based on the EM-DAT data from 2000 to 2019 (CRED & UNDRR, 2020) includes flooding within the hydrological hazards-type, along with landslides and wave action. Floods alone were responsible for 44% of all occurrences, 41% of affected people, 9% of casualties (representing 104,614 persons) and 22% of the economic losses (US$ 651 billion) recorded in the EM-DAT (CRED & UNDRR, 2020).

Natural and human environments, as in many other risks, intertwin significantly to generate flood risk (Zischg et al., 2018). The spatial relation of these environments is usually expressed by hazard, exposure and vulnerability with emphasis on social vulnerability (Koks et al., 2015; Santos et al., 2020). Several approaches are based on the traditional and current understanding of the risk concept (UNDRR, 2019) and define risk as a product of hazard, exposure and vulnerability. Other approaches perform a statistical analysis of the patterns of flood damages along with the characteristics of the hazard and exposure (Mazzoleni et al., 2020). Independently of the approach, databases of past floods and respective damages are essential in the calibration and validation process of flood hazard and risk models (e.g. Khosravi et al., 2019; Li et al., 2012; Termeh et al., 2018; Santos et al., 2019).

An increasingly number of global flood risk indexes has been produced by academic institutions and the business sector. The calculation of such indexes is taking advantage of the high availability of Earth Observation products and cloud computing capacity, producing static risk assessments (Assteerawatt et al., 2016; Phongsapan et al., 2019; Ward et al., 2013, 2015; Wing et al., 2018) or near-real time updated risk assessments (Dottori et al., 2017; Todini, 1999). Even so, resolution and diversity of input data representing exposure and vulnerability varies significantly among the available risk models according to the scale of the analysis.

Flood risk indexes have been designed worldwide to support risk characterization, analysis and management. In the Portuguese context, a static flood risk index was recently proposed at the municipality level, in support of strategic levels of decision in flood risk management (Santos et al., 2020). Subsequently, the research team involved in that work identified the usefulness of deepening the understanding of the risk drivers – susceptibility, exposure and vulnerability – in comparison with the historical records at a larger scale (parish). Such approach would be implemented using Geographic Information Systems (GIS) and machine learning.

The mix of GIS and other techniques like multivariate statistics, multicriteria analysis, physically-based and machine learning models is recognised to be appropriate to flood analysis and modelling (Arabameri et al., 2020; Bui et al., 2019). In this context, several methods have been used.

Regarding statistically-based methods, one can highlight the use of bivariate methods as the weights of evidence (Tehrany et al., 2014) and the frequency ratio (Samanta et al., 2018; Siahkamari et al., 2018). Other data driven methods include the entropy index (Hong et al., 2018), the statistical index (Khosravi et al., 2016), the evidential belief function (Bui et al., 2019), logistic regression (Ali et al., 2020) and the k-nearest neighbour (Costache et al., 2020; Costache et al., 2020; Costache et al., 2020). However, these procedures greatly rely on the relation between dependent and independent variables and are heavily influenced by the datasets size (McLay et al., 2001).

As for the multicriteria analysis to evaluate floods, several methods have also been used (Souissi et al., 2020), like the analytical network process (Cao et al., 2016), the analytical hierarchy process (Ghosh & Kar, 2018; Luu et al., 2018; Tang et al., 2018), the simple additive weighting (Khosravi et al., 2019) and the technique for order preference by similarity to ideal solution (Khosravi et al., 2019). These approaches function upon expert knowledge that can be twisted by confuse rulings and ambiguity (Miles & Snow, 1984).

More recent algorithms include machine learning techniques (Termeh et al., 2018; Zhao et al., 2018), Naive Bayes (Chen et al., 2020), decision tree models like random forest (Lee et al., 2017), artificial neural networks (Chapi et al., 2017), support vector machines (Choubin et al., 2019), support vector machine neuro-fuzzy inference system (Wang et al., 2019) and deep learning neural networks (Bui et al., 2020).

As far as we know, within the diverse decision tree models, like Random Forest (RF), Quick Unbiased and Efficient Statistic Tree, Classification and Regression Trees (CART) and Chi-squared Automatic Interaction

Detection (CHAID), only RF (Lee et al., 2017) and CHAID (Tehrany et al., 2013) have been applied until now in flood analysis and modelling.

In this work, the CART algorithm is used because this method has proved to work well on procedures with nonlinear behaviour and substantial inner heterogeneity (Ji et al., 2013). The CART has a number of proficiencies, like the insensitivity to outliers and data spatial distribution, the capability of integrating both categorical and continuous variables in the model and the ability of using several trees to characterize the modelling processes (Choubin et al., 2018).

The main objective of this paper is to understand the drivers of flood disaster risk at the parish level in the Northern region of Portugal. This study has the following specific objectives:

a) To identify the role of susceptibility, exposure and vulnerability, as the main drivers of flood risk, in justifying human losses and damages caused by floods in the XXI century (2000–2015), suggesting a classification of parishes based on that role;

b) To discuss the selection of the flood risk areas under the framework of the Floods Directive, according to the specific dominant disaster flood driving forces in each parish.

2 Study Area

The study area corresponds to NUTS II Northern region, located in Portugal, with an area of 21 287 km^2 and 3 689 682 inhabitants. The Northern region includes 86 municipalities and 1426 parishes since 2012 (2028 parishes at the 2011 Census date, before the administrative reorganization performed in 2012). The municipalities are organized in eight inter-municipal communities (NUTS III) (Fig. 1). The map also shows the six areas of significant potential flood risk (ASPFR) following the 1st cycle of the Floods Directive (EU Directive 2007/60/EC).

The study area is dominated by mountains in the West (with elevation above 1000 m), plateaus in the East, a narrow coastal platform, tectonic depressions and deeply incised valleys with steep slopes (Fig. 2). Low permeability rocks (granites and metamorphic rocks) dominate, which are often highly fractured and covered by weathered materials resulting from chemical weathering.

Regarding the basins of the main rivers that drain into the Northern region, three of them have their headwaters in Spain – Minho (9091 km^2, of which 8276 km^2 in Portugal), Lima (2522 km^2, of which 1199 km^2 in Portugal) and Douro (97 478 km^2, of which 18 588 km^2 in Portugal) - and the other main drainage basins are fully located in Portugal - Cávado (1699 km^2), Ave (1391 km^2) and Leça (185 km^2) (Fig. 2). The river regime is controlled by the pluvial regime. In these areas, despite snow accumulation is not dominant for the river regime, it may have some contribution to river flow on the transboundary basins (e.g. Diéz-Herrero et al., 2013).

According to the Köppen-Geiger-Pohl climate zones' classification, two subtypes of temperate climate can be found in the study area: temperate with dry or temperate summer (Csb) in the western and central areas, and temperate with dry or hot summer (Csa) in the inner eastern area. The climate in the western area is largely influenced by the proximity to the Atlantic Ocean, while in the east is more continental. The mean annual precipitation (MAP) is higher (2000–3500 mm) in the western mountains, mainly oriented parallel to the coastline, thus blocking the moist westerly winds blowing from the North Atlantic, while the eastern Douro river valley is one of the driest regions in the country (MAP = 300–500 mm). Precipitation concentrates during the autumn and winter seasons and the summer drought typically lasts for three months (June, July and August). Up to 50% of the total rainfall on winter (D-J-F months) is due to cyclonic and directional W and SW atmospheric rivers (Ramos et al., 2015). During the summer months rainfall is usually associated with N and E flows, although local factors such as relief and deep convective depressions are relevant in conditioning the spatial variability of intense rainfall events (Ramos et al., 2014). Spring and autumn months are transitional months and total rainfall can vary significantly (Gallego et al., 2011; Miranda et al., 2002; Trigo & DaCamara, 2000).

River floods have been occurred in large rivers with large drainage basins and were triggered by precipitation periods lasting for several days (Fig. 2, Pereira et al. 2017). Damaging river floods have been more frequent from November to February, although they also occurred on spring and autumn. On the other hand, flash floods are the result of short duration and intense precipitation events that affected small river basins with low concentration times, most frequently occurring during the autumn and winter, but may also occur in any month along the hydrological year (Zêzere et al., 2014).

3 Data And Methods

3.1 Flood damage databases

Data of damaging floods occurred in the Northern region of Portugal, regardless of the number of people affected or the economic value of the resulting damages, was obtained from national and regional newspapers.

Human damages (fatalities, missing, injured, displaced and evacuated people) caused by floods in mainland Portugal are gathered in the DISASTER database (Zêzere et al., 2014) for the period 1865–2010, that was further updated until 2020. The DISASTER database includes floods and landslides that caused human damages, independently of the number of people affected. The methodology used to collect flood DISASTER cases and human damages is summarized in Zêzere et al. (2014).

Minor consequences caused by floods (road and railroad circulation disruption, flooded buildings and corresponding damages) were collected by Santos et al. (2015, 2018) for the period 1865–2016 for the Northern region of Portugal, also using newspapers as data sources. Additional information on the methodology used to collect flood cases that caused material damages is summarized in Santos et al.

(2014, 2015). On both databases, a significant amount of work was developed to check and validate the flood cases, crosschecking different sources, from national to regional and local newspapers.

In this work, we explore the damaging floods (river floods and flash floods) that caused both human and minor (uniquely material) losses in the Northern region of Portugal for the period 2000–2015, for which the exact location is known (Fig. 3). Along this period, the DISASTER database counts 40 flood cases that caused human damages (3 fatalities, 1 injured, 267 evacuated and 197 displaced people). Adding to the DISASTER database, 213 flood cases that caused minor damages were recorded, of which 114 caused road disruptions (53.8%) and 94 caused damages in the built environment (44.3%) (Table 1).

Table 1
– Damaging floods between 2000 and 2015 the Northern region of Portugal

(Sources: DISASTER database; Santos et al., 2018)

DISASTER database			**Minor damages database**			**Total**	
Flash flood	**River floods**	**Total**	**Flash flood**	**River floods**	**Total**		
Nr. of cases	7	33	**40**	56	157	**213**	**253**
Nr. of fatalities	2	1	**3**	-	-	**-**	**3**
Nr. of injured people	0	1	**1**	-	-	**-**	**1**
Nr. of evacuated people	6	261	**267**	-	-	**-**	**267**
Nr. of displaced people	6	191	**197**	-	-	**-**	**197**
Nr. of cases with building damages	6	11	**17**	18	76	**94**	**111**
Nr. of cases with road network damages	5	6	**11**	32	82	**114**	**124**

Human damages caused by floods were recorded in 9 years (2000, 2001, 2002, 2004, 2006, 2010, 2011, 2013, 2014), while minor damages were recorded every year. The hydrological year of 2000/2001 recorded the highest number of damaging floods in the Northern region. River floods, in opposition to flash floods, are the fluvial process that causes most of flood damages (whether with human consequences or with only minor damages) in the study area (89%). The Douro river basin concentrates 48% of the damaging floods, followed by the Ave river basin with 26% (Fig. 3). The majority of the minor damages are located near the coastal urban areas.

3.2 Susceptibility

Flood susceptibility for the study area was calculated at the parish level from the results obtained in a national-scale assessment (Pedro Pinto Santos et al., 2019). The model evaluates the propensity of each stream to flooding using four types of raw input data: a hydrologically corrected DEM, land use, dominant parent material and the fine fraction of the topsoil. From the first input data, flow accumulation (F_{acc}) and average slope (S_{avg}) of each cell (with 3 arc-second resolution) was derived. The remaining data was used to derive the third stream flood susceptibility conditioning factor, relative permeability (P_{rel}). Weights assigned to the three factors are validated using historical evidences of past major floods (Zêzere et al., 2014) in four validation areas: mainland Portugal, the Northern region, and three sub-basins where a historical record of minor flood cases was also available. The cell-defined susceptibility score is based on the sum of those weighted values, transformed through the min-max method to a range from 0 to 5. The advantage of this method is the evaluation of flood susceptibility of each cell considering the contribution of the entire upstream drainage area. Additional information of the used method can be found in Santos et al. (2019). The flood susceptibility value calculated per each parish is the simple arithmetic average of the susceptibility scores (Fig. 4).

3.3 Exposure

In the exposure assessment to floods at the parish level, two variables were used: population density and the percentage of the artificialized territories (Fig. 5). Population density was computed for each parish using the resident population in the 2011 population Census, divided by the parish area in km^2. The artificialized territories include the urban areas, road and railroad networks, construction areas, quarries and urban green spaces. This data was extracted from the Land Cover map of 2010 (DGT, 2011) with a minimum cartographic unit of 1 ha, using the legend codes at level I, and computing the percentage of the artificialized areas in the total area of each parish. The final exposure value corresponds to the arithmetic average between population density and the percentage of the artificialized territories. Exposure in the Northern region of Portugal shows a very polarised spatial pattern, with a concentrated presence of population and artificial areas in the metropolitan area of Porto, around the mouth of the Douro river, and from here northwards to the major cities of Braga and Viana do Castelo (Fig. 5).

3.4 Vulnerability

Vulnerability is represented in this research by social vulnerability (SV). It is defined as the propensity of individuals, communities and systems to be harmed by hazardous processes, based on their social and demographic characteristics and territorial context (Chen et al., 2013; Cutter et al., 2003; Mendes, 2009; Mendes et al., 2019; Ogie & Pradhan, 2019; Tavares et al., 2018; Yoon, 2012). The statistical procedure to assess SV consists in the iterative application of principal component analysis until a robust and interpretable relationship of principal components is obtained from the explicative variables. In this research, an initial set of 24 variables collected at the parish level from the population Census of 2011 was used. After the iterative process, the final SV model ran with 15 variables, that were grouped around four principal components (PC) (Table 1).

PC1 represents the contexts of demographic ageing, low qualifications and low levels of women empowerment. PC2 represents the demographic and economic dynamism: positive loadings in variables related with qualifications, socially-valued employments, and internal and external attractiveness, whose PC scores needed to be inverted to be concordant with SV interpretation. PC3 represents the economic condition of residents: positive loadings in variables as overcrowded households and households without at least one basic infrastructure, opposing to the negative loading in variable car usage in daily commuting. Finally, PC4 represents the mobility, as a SV driver expressing the need of long commuting times and working or studying outside the municipality of residence.

A score of SV is obtained (Fig. 6) by applying weights to the components' scores based on the percentage of explained variance. The rural areas, with aged population and households without basic infrastructures are those with the highest social vulnerability. Some parishes in the old city centers and outskirts of the Porto city are also evidenced as highly vulnerable.

Table 2 – Principal components of social vulnerability in 2011 for the Northern region of Portugal. Darker backgrounds mean stronger loadings.

Explicative variables	Loading within each principal component			
	PC1	PC2	PC3	PC4
Mean age (years)	0.952	0.007	-0.089	-0.026
Households with a person aged 65 years old and over (%)	0.912	0.037	0.030	-0.020
Women activity rate (%)	-0.862	0.246	-0.105	-0.042
Illiteracy rate (%)	0.800	-0.247	0.173	0.007
Variation of population between 2001-2011 (%)	0.691	-0.354	0.135	-0.058
Population with higher education completed (%)	-0.392	0.761	-0.186	-0.078
Professionals socially more valued (%)	-0.160	0.755	-0.215	-0.165
Population 5 years before living in other municipality (%)	0.030	0.748	0.029	0.288
resident population of foreign nationality (%)	-0.096	0.689	0.336	0.068
Overcrowded households (%)	-0.477	-0.089	0.731	0.109
Car usage on daily journeys (%)	-0.339	0.010	-0.657	-0.088
Households without at least one basic infrastructure (%)	0.396	-0.299	0.445	0.143
School leavers rate (%)	0.041	0.056	0.296	-0.204
Population working or studying in another municipality (%)	-0.191	0.079	-0.188	0.837
Average time spent on commuting (min)	0.178	0.032	0.261	0.740
Cardinality	+	-	+	+
% of variance explained	33.505	13.349	11.634	7.965

3.5 Geographical and statistical analysis

The geographical and statistical analysis of the flood damages, susceptibility, exposure and vulnerability scores is performed with the transformed values to the range [0, 1], using the min-max method.

The reported damaging floods were split in two databases, to which the geographical and statistical analysis was conducted:

- Database of DISASTER-damaging flood cases, which only assembles the DISASTER-type cases;
- Database of All-damaging flood cases, which assembles damaging flood cases with minor consequences and DISASTER-type cases.

The purpose of this procedure is to test if the risk components – susceptibility, exposure and vulnerability – play the same role independently of the type of damaging flood that is reported.

Classification using regression trees (CART) was developed by Breiman et al. (2017). It is a recurrent distribution non-parametric algorithm used to analyse and predict data relations (Choubin et al., 2018), applied both to predict categorical, i.e., classification, and continuous variables, i.e., regression (Felicísimo et al., 2013; Youssef et al., 2016). The dissimilarity amongst classification and regression approaches is on the methods used to split and assemble the data (Prasad et al., 2006).

CART was applied on two groups of cases (a group with All-damaging flood cases and another with DISASTER-damaging flood cases) to divide the flood cases into a sequence of classes upon their internal homogeneity and built a model for each group. The objective of this procedure is to create a tree and evaluate the set of logical ifthenelse split conditions used. In this case, starting from the root, the method splits the data subsets, generating two child nodes. If the separated data belongs to the same class, then it is merged and forms a leaf. If not, the division process continues. The sample is divided and catalogued according to the squared residuals minimization algorithm (Timofeev, 2004). This method creates data subsets the most likely similar to the dependent variable (Mahjoobi & Etemad-Shahidi, 2008). More information concerning the CART algorithm can be found in (Choubin et al., 2018).

3.6 Validation and sensitivity

In order to have an enhanced knowledge of our analysis accuracy a k-fold cross validation procedure was applied. One randomly split the training data into 20 subsamples, i.e., folds, and trained each one of them with 80% of the data. The remaining

4 Results

4.1 Classification, regression trees

Using the All-damaging flood cases database and using the DISASTER-damaging flood cases database (i.e., that with only the most severe human consequences), results in a quite distinct classification of parishes according to the regression trees algorithm.

Both trees are simple structures with three nodes but the similarities stop there. The first two leaves of the tree ran with All-damaging flood cases (Fig. 7) come from splitting exposure values using somewhat low thresholds i.e., low exposure (0.299 and 0.143 in a 0 to 1 scale), due to the high asymmetric distribution of urban areas in the study area (cf. Figure 5). These two nodes classify 20% of the parishes giving us groups '2' and '4'. Finally, the last splitting is supported by vulnerability and separate the parishes with low values (= < 0.5, 25%) from those with high values (> 0.5, 55%). This gives us groups '5' and '6', respectively.

For the database that uses only the DISASTER-damaging flood cases (Fig. 8) the situation is diverse. The three splitting nodes use sequentially: susceptibility, exposure and vulnerability. The first retrieves a leaf (4.9%) of high susceptibility values (> 0.6) and corresponds to group '2'. The second, splits between the 0.297 exposure threshold, defining group '4' (4.9%) as the one with values above it. Finally, the third one, separates two leaves, groups '5' and '6', the first with vulnerability values below or equal to 0.479 and the second with values above it.

In summary, using the database with All-damaging flood cases, the 4 groups are interpreted as:

- Group '2': parishes with high exposure;
- Group '4': parishes with intermediate exposure;
- Group '5': parishes with low exposure and low vulnerability:
- Group '6': parishes with low exposure and high vulnerability.

It is worth mentioning that in this regression tree, flood susceptibility does not play a key role in classifying the parishes using historical flood cases.

The groups resulting from using only the DISASTER-damaging flood cases can be interpreted as:

- Group '2': parishes with high susceptibility:
- Group '4': parishes with low susceptibility and high exposure:
- Group '5': parishes with low susceptibility, low exposure and low vulnerability:
- Group '6': parishes with low susceptibility, low exposure and high vulnerability.

Spatially, the groups based on the analysis with the All-damaging flood cases (Fig. 9) tend to isolate important urban (highly populated) parishes and their expansion crowns, whereas in the DISASTER-damaging flood cases classification (Fig. 10) the mark of these transition/influence areas is more fuzzy and the influence of the proximity to two important rivers, i.e., Minho and Douro, is highlighted.

4.2 Validation and sensitivity

From Table 3, more confidence is put on the model created upon the DISASTER-damaging flood cases than in the model with the All-damaging cases. The former has a much lower – and acceptable – misclassification estimation and a residual standard error.

Table 3
– Models misclassification analysis

	Misclassification	
	Estimate	**Std. Error**
All-damaging flood cases	0.568	0.223
DISASTER-damaging flood cases	0.026	0.006

One used a measure of how much the model predicts value changes according to the dependent variable in response to changes on the independent variables, in order to analyse the variables importance. The normalized value is just the importance of the variable regarding the highest importance value, displayed as percentage.

Analysing the results in Table 4, it can be drawn that in both databases exposure has a high importance when comes to justify the cases: normalized importance of 100% using the All-damaging and 98.5% using the DISASTER-damaging flood cases. The less important variable is the vulnerability, whereas with four times more importance when analysing the DISASTER-damaging flood cases in comparison with the All-damaging cases database: 25.1% comparing to 6.5% of the normalized importance, respectively (Table 4). The extreme behaviour is given by susceptibility, that has no importance in discriminating the All-damaging flood cases but arises as the most important when analysing the DISASTER-damaging cases.

Table 4
– Risk components importance

Independent Variable	Variable Importance	Normalized Importance (%)
All-damaging flood cases		
Exposure	0.036	100
Vulnerability	0.002	6.5
DISASTER-damaging flood cases		
Susceptibility	0.000429	100
Exposure	0.000422	98.5
Vulnerability	0.000107	25.1

5 Discussion

From the 2028 parishes, 57 parishes are included in the six Floods Directive areas of significant potential flood risk (ASPFR). Considering the three drivers of flood damages – susceptibility, exposure and vulnerability – the ASPFR parishes present low exposure, with the exception of the Porto ASPFR (Fig. 11). Porto (5) and Peso da Régua (6) ASPFR are the ones located along the major river, the Douro river, which explains the high susceptibility index value. Comparing the parishes covered by the Directive with the parishes outside the ASPFR (Fig. 11), it is verified that exposure and susceptibility are, on average, higher in the first group, while vulnerability is higher in the parishes outside the ASPFR. This confirms the results presented in Table 4, which assign less importance to vulnerability as a splitting criterion in the CART classification method. Looking at the number of flood cases per parish, the Directive's parishes are capturing 23% of the DISASTER cases (9 in 40 cases) and 31% of the All-damaging flood cases database (79 in 253 cases). This means that around ¾ of the DISASTER-damaging flood cases have occurred in streams not covered by the Floods Directive. Although not available yet, it is expected that several of these locations will be covered in the 2nd cycle, as the number of ASPFR in the study area has been expanded to 17.

A more detailed analysis of CART results combined with the Flood Directive areas demonstrates that the Directive's parishes differ significantly within each ASPFR as well, and that using criteria-differentiated flood impact databases provides quite distinctive interpretations for a same parish (Fig. 12). Using the database with All-damaging flood cases, 18 parishes are individualized from the set of 57 simply considering the exposure driver (group 2), with relevance in the Porto ASPFR. The DISASTER-damaging flood cases complement the analysis by introducing the role of susceptibility: i.e., 8 of the 10 parishes of Peso da Régua ASPFR are classified in the All-damaging cases database as having low exposure and high vulnerability, while considering the DISASTER cases allows to characterize the hazard dimension, in this case with high susceptibility. Chaves ASPFR is another interesting example of improved interpretation capacity using the two CART analysis: in the all flood cases CART, 2 parishes show high exposure and 8 show low exposure; in the DISASTER-damaging flood cases CART, the same 2 and 8 parishes maintain the exposure classifications, but this analysis adds that they present low susceptibility.

Although the DISASTER-damaging flood cases database produces a classification with less standard error (Table 3), both databases render an understanding about flood risk drivers, linked to small and frequent flood episodes, that is simultaneously relevant to flood risk decision-makers. Regarding the use of historical flood damage databases in risk characterization, this study concludes that using databases differentiated, or filtered, by the type of impact provides an equally differentiated classification of the role of risk drivers. This means that the role of susceptibility, exposure and vulnerability varies according to the damage characteristics of each record. When filtered by the degree of damage, the two distinct databases related differently with the flood risk drivers.

Narrow inclusion criteria of flood damage databases – requiring the existence of 10 or more fatalities, 100 or more affected, declaration of a state of emergency and a call for international assistance, as in the case of the EM-DAT – is optimal for a global comparison of contexts by reporting only the most important events. In local studies, like the one here presented for the Portuguese Northern Region, the

narrower set of data (the DISASTER-damaging flood cases), eventually and surprisingly, provided a more holistic representation of flood drivers when compared to the All-damaging flood cases database. This conclusion is partly correct: in fact, minor and small disasters are more frequent and are explained by societal and human-related drivers – not so dependent on the hazard characteristics – than the most impacting flood events. In resume, and considering three types of impact databases – global ones like the EM-DAT, national-level like the DISASTER and the all-cases type like those used in (Santos et al., 2018; Santos & Reis, 2018) –, their different inclusion criteria provide complementary illustrations of the role of flood risk drivers. This redirect us to the discussion on the standardization methods for monitoring damages from extreme and small disasters worldwide, and the need to consider the scale and purposes for which the data will be used for (De Groeve et al., 2013; Gall et al., 2009; Kron et al., 2012).

6 Conclusions

Two types of flood damage databases were used as dependent variables to train a machine learning classification method, based on regression trees (CART), that uses flood susceptibility, exposure and vulnerability as independent variables. All the data is expressed at the civil parish, a local-level administrative unit, summing 2028 territorial units of analysis in the Northern region of Portugal.

The results show that the risk drivers assume distinct roles in relation to the dependent variable, according to the inclusion criteria adopted in the definition of the two databases. Ultimately, the degree of damages (serious and minor), has documented in databases, is conditioned by the geographical distribution of the flood risk drivers. The inclusion of all types of damages highlights the role of exposure first and vulnerability later, but when the cases are filtered to include only human consequences – casualties, injuries, missing people, evacuated or displaced people – flood susceptibility emerges as the most relevant driver, followed by exposure and vulnerability.

Next steps of research in the field should explore in greater detail the combined effect of i) the scale at which damage data is aggregated, ii) and the distinct inclusion criteria on databases, over the explicative and predictive capacity of the three flood risk drivers.

References

1. Ali SA, Parvin F, Pham QB, Vojtek M, Vojteková J, Costache R, Thi Thuy Linh N, Nguyen HQ, Ahmad A, Ghorbani MA (2020) GIS-based comparative assessment of flood susceptibility mapping using hybrid multi-criteria decision-making approach, naïve Bayes tree, bivariate statistics and logistic regression: A case of Topľa basin, Slovakia. Ecol Ind 117:106620–2020v117. https://doi.org/10.1016/j.ecolind.2020.106620
2. Arabameri A, Saha S, Chen W, Roy J, Pradhan B, Bui DT (2020) Flash flood susceptibility modelling using functional tree and hybrid ensemble techniques. J Hydrol 587:125007. https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.125007

3. Assteerawatt A, Tsaknias D, Azemar F, Ghosh S, Hilberts A (2016) Large-scale and High-resolution Flood Risk Model for Japan. *E3S Web of Conferences*, *7*, 1–5. https://doi.org/10.1051/e3sconf/20160711009
4. Breiman L, Friedman JH, Olshen RA, Stone CJ (2017) Classification and regression trees. Routledge
5. Bui DT, Tsangaratos P, Ngo P-TT, Pham TD, Pham BT (2019) Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods. Sci Total Environ 668:1038–1054. https://doi.org/10.1016/j.scitotenv.2019.02.422
6. Cao C, Xu P, Wang Y, Chen J, Zheng L, Niu C (2016) Flash Flood Hazard Susceptibility Mapping Using Frequency Ratio and Statistical Index Methods in Coalmine Subsidence Areas. In *Sustainability* (Vol. 8, Issue 9). https://doi.org/10.3390/su8090948
7. Chapi K, Singh VP, Shirzadi A, Shahabi H, Bui DT, Pham BT, Khosravi K (2017) A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environ Model Softw 95:229–245
8. Chen W, Li Y, Xue W, Shahabi H, Li S, Hong H, Wang X, Bian H, Zhang S, Pradhan B, Ahmad B, Bin (2020) Modeling flood susceptibility using data-driven approaches of naïve Bayes tree, alternating decision tree, and random forest methods. Sci Total Environ 701:134979. https://doi.org/10.1016/j.scitotenv.2019.134979
9. Chen W, Cutter SL, Emrich CT, Shi P (2013) Measuring social vulnerability to natural hazards in the Yangtze River Delta region, China. Int J Disaster Risk Sci 4(4):169–181. https://doi.org/10.1007/s13753-013-0018-6
10. Choubin B, Moradi E, Golshan M, Adamowski J, Sajedi-Hosseini F, Mosavi A (2019) An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. Sci Total Environ 651(Pt 2):2087–2096. https://doi.org/10.1016/j.scitotenv.2018.10.064
11. Choubin B, Zehtabian G, Azareh A, Rafiei-Sardooi E, Sajedi-Hosseini F, Kişi Ö (2018) Precipitation forecasting using classification and regression trees (CART) model: a comparative study of different approaches. Environ Earth Sci 77(8):1–13
12. Costache R, Hong H, Pham QB (2020) Comparative assessment of the flash-flood potential within small mountain catchments using bivariate statistics and their novel hybrid integration with machine learning models. Sci Total Environ 711:134514. https://doi.org/10.1016/j.scitotenv.2019.134514
13. Costache R, Pham QB, Sharifi E, Linh NT, Abba SI, Vojtek M, Vojteková J, Nhi PT, Khoi DN (2020) Flash-Flood Susceptibility Assessment Using Multi-Criteria Decision Making and Machine Learning Supported by Remote Sensing and GIS Techniques. In *Remote Sensing* (Vol. 12, Issue 1). https://doi.org/10.3390/rs12010106
14. Costache R, Popa MC, Bui T, Diaconu D, Ciubotaru DC, Minea N, Pham QB (2020) Spatial predicting of flood potential areas using novel hybridizations of fuzzy decision-making, bivariate statistics, and machine learning. J Hydrol 585:124808. https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.124808

15. CRED, & UNDRR (2020) *Human cost of disasters. An overview of the last 20 years (2000–2019)*. https://reliefweb.int/report/world/human-cost-disasters-overview-last-20-years-2000-2019
16. Cutter SL, Boruff BJ, Shirley WL (2003) Social vulnerability to environmental hazards. Soc Sci Q 84(2):242–261. https://doi.org/10.1111/1540-6237.8402002
17. De Groeve T, Poljansek K, Ehrlich D(2013) *Recording disaster losses recommendations for a European approach*. https://doi.org/10.2788/98653
18. Dottori F, Kalas M, Salamon P, Bianchi A, Alfieri L, Feyen L (2017) An operational procedure for rapid flood risk assessment in Europe. Nat Hazards Earth Syst Sci 17(7):1111–1126. https://doi.org/10.5194/nhess-17-1111-2017
19. Felicísimo ÁM, Cuartero A, Remondo J, Quirós E (2013) Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslides 10(2):175–189. https://doi.org/10.1007/s10346-012-0320-1
20. Gall M, Borden KA, Cutter SL (2009) When do losses count? Bull Am Meteorol Soc 90(6):799–809. https://doi.org/10.1175/2008BAMS2721.1
21. Gallego MC, Trigo RM, Vaquero JM, Brunet M, García JA, Sigró J, Valente MA (2011) Trends in frequency indices of daily precipitation over the Iberian Peninsula during the last century. J Geophys Res Atmos 116(2):1–18. https://doi.org/10.1029/2010JD014255
22. Ghosh A, Kar SK (2018) Application of analytical hierarchy process (AHP) for flood risk assessment: a case study in Malda district of West Bengal, India. Nat Hazards: J Int Soc Prev Mitigation Nat Hazards 94(1):349–368. https://doi.org/10.1007/s11069-018-3392-y
23. Hong H, Tsangaratos P, Ilia I, Liu J, Zhu A-X, Chen W (2018) Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China. Sci Total Environ 625:575–588. https://doi.org/10.1016/j.scitotenv.2017.12.256
24. Ji Z, Li N, Xie W, Wu J, Zhou Y (2013) Comprehensive assessment of flood risk using the classification and regression tree method. Stoch Env Res Risk Assess 27(8):1815–1828. https://doi.org/10.1007/s00477-013-0716-z
25. Khosravi K, Pourghasemi HR, Chapi K, Bahri M (2016) Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: a comparison between Shannon(')s entropy, statistical index, and weighting factor models. Environ Monit Assess 188(12):656. https://doi.org/10.1007/s10661-016-5665-9
26. Khosravi K, Shahabi H, Pham BT, Adamowski J, Shirzadi A, Pradhan B, Dou J, Ly H-B, Gróf G, Ho HL, Hong H, Chapi K, Prakash I (2019) A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. J Hydrol 573:311–323. https://doi.org/https://doi.org/10.1016/j.jhydrol.2019.03.073
27. Koks EE, Jongman B, Husby TG, Botzen WJW (2015) Combining hazard, exposure and social vulnerability to provide lessons for flood risk management. Environ Sci Policy 47:42–52. https://doi.org/10.1016/j.envsci.2014.10.013

28. Kron W, Steuer M, Löw P, Wirtz A (2012) How to deal properly with a natural catastrophe database - Analysis of flood losses. Nat Hazards Earth Syst Sci 12(3):535–550. https://doi.org/10.5194/nhess-12-535-2012
29. Lee S, Kim J-C, Jung H-S, Lee MJ, Lee S (2017) Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomatics Nat Hazards Risk 8(2):1185–1203. https://doi.org/10.1080/19475705.2017.1308971
30. Li K, Wu S, Dai E, Xu Z (2012) Flood loss analysis and quantitative risk assessment in China. Nat Hazards 63(2):737–760. https://doi.org/10.1007/s11069-012-0180-y
31. Luu C, Meding J, Von, Kanjanabootra S(2018) Assessing flood hazard using flood marks and analytic hierarchy process approach: a case study for the 2013 flood event in Quang Nam, Vietnam. *Natural Hazards*, *90*(3), 1031–1050. https://doi.org/DOI:101007/s11069-017-3083-0
32. Mahjoobi J, Etemad-Shahidi A (2008) An alternative approach for the prediction of significant wave heights based on classification and regression trees. Appl Ocean Res 30(3):172–177. https://doi.org/https://doi.org/10.1016/j.apor.2008.11.001
33. Mazzoleni M, Mård J, Rusca M, Odongo V, Lindersson S, Di Baldassarre G (2020) Floodplains in the Anthropocene: A global analysis of the interplay between human population, built environment and flood severity. Water Resour Res 57(2). https://doi.org/10.1029/2020WR027744. e2020WR027744
34. McLay CD, Dragten R, Sparling G, Selvarajah N (2001) Predicting groundwater nitrate concentrations in a region of mixed agricultural land use: a comparison of three approaches. Environ Pollution (Barking Essex: 1987) 115(2):191–204. https://doi.org/10.1016/s0269-7491(01)00111-7
35. Mendes JM (2009) Social vulnerability indexes as planning tools: Beyond the preparedness paradigm. J Risk Res 12(1):43–58. https://doi.org/10.1080/13669870802447962
36. Mendes JM, Tavares AO, Santos PP (2019) Social vulnerability and local level assessments: a new approach for planning. Int J Disaster Resil Built Environ 11(1):15–43. https://doi.org/10.1108/IJDRBE-10-2019-0069
37. Miles RE, Snow CC (1984) Designing strategic human resources systems. Organ Dyn 13(1):36–52. https://doi.org/https://doi.org/10.1016/0090-2616(84)90030-5
38. Miranda PMA, Espírito F, Coelho S, Rodrigues Tomé A, Valente MA, Pires HO, Pires VC, Miranda PMA, Coelho FES, Tomé AR, Valente MA, Carvalho A, Pires C, Pires HO, Pires VC, Ramalho C(2002) 20th century Portuguese Climate and Climate Scenarios. In F. D. Santos, K. Forbes, & R. Moita (Eds.), *Climate Change in Portugal: Scenarios, Impacts and Adaptation Measures (SIAM Project)* (pp. 23–83). Gradiva
39. Ogie RI, Pradhan B (2019) Natural Hazards and Social Vulnerability of Place: The Strength-Based Approach Applied to Wollongong, Australia. Int J Disaster Risk Sci 10(3):404–420. https://doi.org/10.1007/s13753-019-0224-y
40. Phongsapan K, Chishtie F, Poortinga A, Bhandari B, Meechaiya C, Kunlamai T, Aung KS, Saah D, Anderson E, Markert K, Markert A, Towashiraporn P (2019) Operational Flood Risk Index Mapping for

Disaster Risk Reduction Using Earth Observations and Cloud Computing Technologies: A Case Study on Myanmar. Front Environ Sci 7:191. https://doi.org/10.3389/FENVS.2019.00191/BIBTEX

41. Prasad AM, Iverson LR, Liaw A (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9(2):181–199
42. Ramos AM, Trigo RM, Liberato MLR, Tomé R (2015) Daily Precipitation Extreme Events in the Iberian Peninsula and Its Association with Atmospheric Rivers. J Hydrometeorol 16(2):579–597. https://doi.org/10.1175/JHM-D-14-0103.1
43. Razavi Termeh SV, Kornejady A, Pourghasemi HR, Keesstra S (2018) Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. Sci Total Environ 615:438–451. https://doi.org/10.1016/j.scitotenv.2017.09.262
44. Samanta RK, Bhunia GS, Shit PK, Pourghasemi HR (2018) Flood susceptibility mapping using geospatial frequency ratio technique: a case study of Subarnarekha River Basin, India. Model Earth Syst Environ 4(1):395–408. https://doi.org/10.1007/s40808-018-0427-z
45. Santos M, Fragoso M, Santos JA (2018) Damaging flood severity assessment in Northern Portugal over more than 150 years (1865–2016). Nat Hazards 91(3):983–1002. https://doi.org/10.1007/s11069-017-3166-y
46. Santos PP, Reis E(2018) Assessment of stream flood susceptibility: a cross-analysis between model results and flood losses. *Journal of Flood Risk Management*, *11*. https://doi.org/10.1111/jfr3.12290
47. Santos PP, Reis E, Pereira S, Santos M(2019) A flood susceptibility model at the national scale based on multicriteria analysis. *Science of the Total Environment*, *667*. https://doi.org/10.1016/j.scitotenv.2019.02.328
48. Santos P, Pinto, Pereira S, Zêzere JL, Tavares AO, Reis E, Garcia RAC, Oliveira SC (2020) A comprehensive approach to understanding flood risk drivers at the municipal level. J Environ Manage 260(January). https://doi.org/10.1016/j.jenvman.2020.110127
49. Santos P, Pinto, Reis E, Pereira S, Santos M (2019) A flood susceptibility model at the national scale based on multicriteria analysis. Sci Total Environ 667:325–337. https://doi.org/10.1016/j.scitotenv.2019.02.328
50. Siahkamari S, Haghizadeh A, Zeinivand H, Tahmasebipour N, Rahmati O (2018) Spatial prediction of flood-susceptible areas using frequency ratio and maximum entropy models. Geocarto Int 33(9):927–941. https://doi.org/10.1080/10106049.2017.1316780
51. Souissi D, Zouhri L, Hammami S, Msaddek MH, Zghibi A, Dlala M (2020) GIS-based MCDM – AHP modeling for flood susceptibility mapping of arid areas, southeastern Tunisia. Geocarto Int 35(9):991–1017. https://doi.org/10.1080/10106049.2019.1566405
52. Tang Z, Zhang H, Yi S, Xiao Y (2018) Assessment of flood susceptible areas using spatially explicit, probabilistic multi-criteria decision analysis. J Hydrol 558:144–158. https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.01.033
53. Tavares AO, Barros JL, Mendes JM, Santos PP, Pereira S (2018) Decennial comparison of changes in social vulnerability: A municipal analysis in support of risk management. Int J Disaster Risk Reduct

31:679–690. https://doi.org/10.1016/j.ijdrr.2018.07.009

54. Tehrany MS, Pradhan B, Jebur MN (2013) Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. J Hydrol 504:69–79. https://doi.org/https://doi.org/10.1016/j.jhydrol.2013.09.034
55. Tehrany MS, Pradhan B, Jebur MN (2014) Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. J Hydrol 512:332–343. https://doi.org/10.1016/j.jhydrol.2014.03.008
56. Termeh SVR, Kornejady A, Pourghasemi HR, Keesstra S (2018) Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. Sci Total Environ 615:438–451
57. Tien Bui, Dieu, Hoang N-D, Martínez-Álvarez F, Ngo P-TT, Hoa PV, Pham TD, Samui P, Costache R (2020) A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area. Sci Total Environ 701:134413. https://doi.org/10.1016/j.scitotenv.2019.134413
58. Tien Bui, Duie, Khosravi K, Shahabi H, Daggupati P, Adamowski JF, Melesse AM, Pham T, Pourghasemi B, Mahmoudi HR, Bahrami M, Pradhan S, Shirzadi B, Chapi A, K., & Lee S(2019) Flood Spatial Modeling in Northern Iran Using Remote Sensing and GIS: A Comparison between Evidential Belief Functions and Its Ensemble with a Multivariate Logistic Regression Model. In *Remote Sensing* (Vol. 11, Issue 13). https://doi.org/10.3390/rs11131589
59. Timofeev R (2004) Classification and regression trees (CART) theory and applications. Humboldt University, Berlin, pp 1–40
60. Todini E (1999) An operational decision support system for flood risk mapping, forecasting and management. Urban Water 1(2):131–143. https://doi.org/10.1016/S1462-0758(00)00010-8
61. UNDRR (2019) Global Assessment Report on Disaster Risk Reduction (GAR2019). United Nations Office for Disaster Risk Reduction. https://gar.undrr.org/report-2019
62. Wang Y, Hong H, Chen W, Li S, Panahi M, Khosravi K, Shirzadi A, Shahabi H, Panahi S, Costache R (2019) Flood susceptibility mapping in Dingnan County (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm. J Environ Manage 247:712–729. https://doi.org/10.1016/j.jenvman.2019.06.102
63. Ward PJ, Jongman B, Salamon P, Simpson A, Bates P, De Groeve T, Muis S, De Perez EC, Rudari R, Trigg MA, Winsemius HC (2015) Usefulness and limitations of global flood risk models. Nat Clim Change 5(8):712–715. https://doi.org/10.1038/nclimate2742
64. Ward PJ, Jongman B, Weiland FS, Bouwman A, Van Beek R, Bierkens MFP, Ligtvoet W, Winsemius HC (2013) Assessing flood risk at the global scale: model setup, results, and sensitivity. Environ Res Lett 8(4):044019. https://doi.org/10.1088/1748-9326/8/4/044019
65. Wing OEJ, Bates PD, Smith AM, Sampson CC, Johnson KA, Fargione J, Morefield P (2018) Estimates of present and future flood risk in the conterminous United States. Environ Res Lett 13(3). https://doi.org/10.1088/1748-9326/aaac65

66. Yoon DK (2012) Assessment of social vulnerability to natural disasters: A comparative study. Nat Hazards 63(2):823–843. https://doi.org/10.1007/s11069-012-0189-2
67. Youssef AM, Sefry SA, Pradhan B, Alfadail EA (2016) Analysis on causes of flash flood in Jeddah city (Kingdom of Saudi Arabia) of 2009 and 2011 using multi-sensor remote sensing data and GIS. Geomatics Nat Hazards Risk 7(3):1018–1042. https://doi.org/10.1080/19475705.2015.1012750
68. Zêzere JL, Pereira S, Tavares AO, Bateira C, Trigo RM, Quaresma I, Santos PP, Santos M, Verde J (2014) DISASTER: A GIS database on hydro-geomorphologic disasters in Portugal. Nat Hazards 72(2):503–532. https://doi.org/10.1007/s11069-013-1018-y
69. Zhao G, Pang B, Xu Z, Yue J, Tu T (2018) Mapping flood susceptibility in mountainous areas on a national scale in China. Sci Total Environ 615:1133–1142. https://doi.org/https://doi.org/10.1016/j.scitotenv.2017.10.037
70. Zischg AP, Hofer P, Mosimann M, Röthlisberger V, Ramirez JA, Keiler M, Weingartner R (2018) Flood risk (d)evolution: Disentangling key drivers of flood risk change with a retro-model experiment. Sci Total Environ 639:195–207. https://doi.org/10.1016/j.scitotenv.2018.05.056
71. Zischg, A. P., Hofer, P., Mosimann, M., Röthlisberger, V., Ramirez, J. A., Keiler, M., & Weingartner, R. (2018). Flood risk (d)evolution: Disentangling key drivers of flood risk change with a retro-model experiment. *Science of the Total Environment*, *639*, 195–207. https://doi.org/10.1016/j.scitotenv.2018.05.056

Figures

Figure 1

Population density (no. inhabitants/km^2) in the parishes of the Northern region (NUTS II) of Portugal (Population Census, 2011). Note: ASPFR study areas: (1) Chaves; (2) Esposende; (3) Ponte da Barca; (4) Ponte de Lima; (5) Porto; (6) Peso da Régua.

Figure 2

Elevation, main drainage basins and damaging floods from 2000 to 2015 in the Northern region of Portugal (Sources: DISASTER database; Santos et al., 2018).

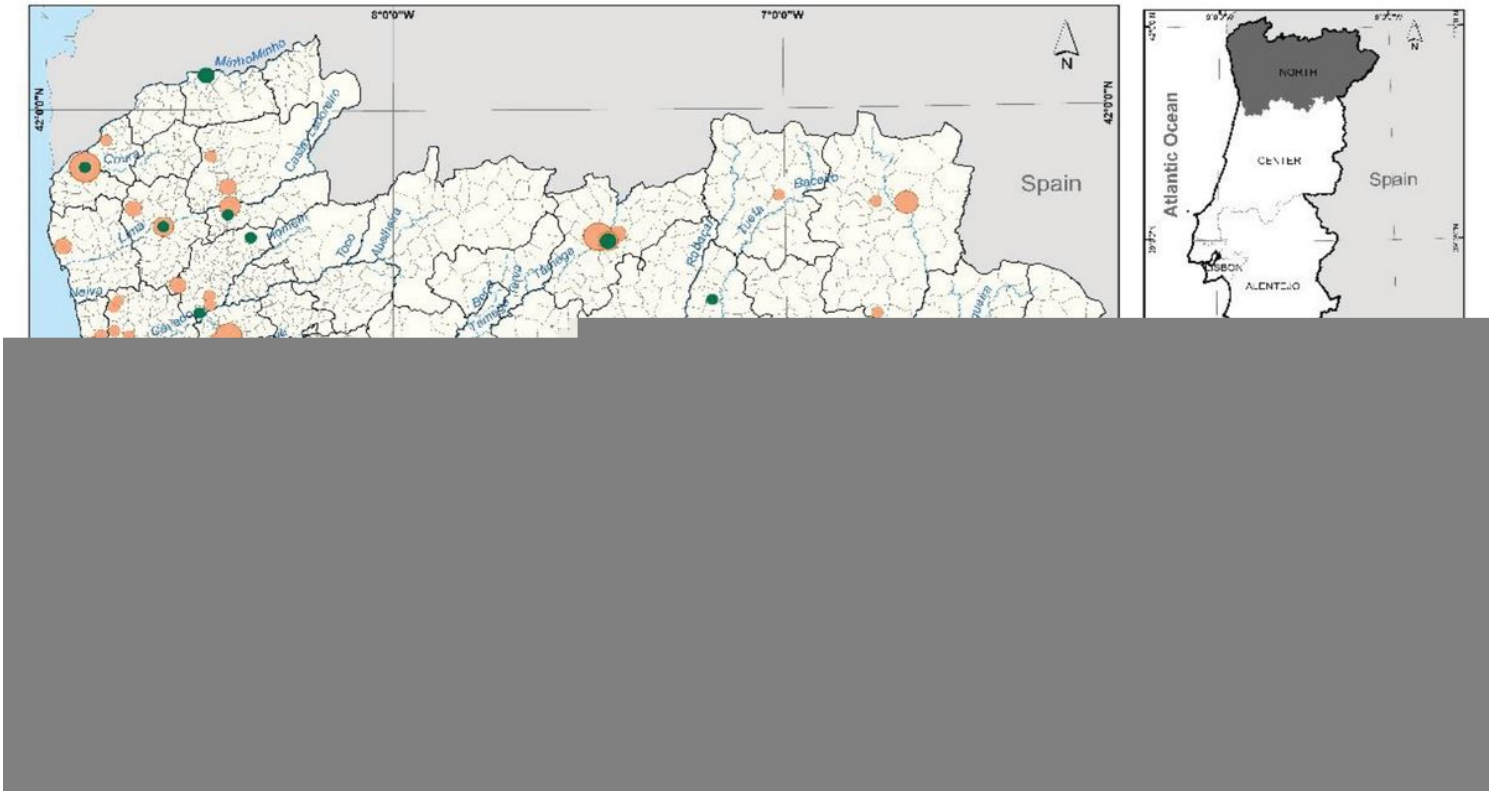

Figure 3

Flood susceptibility index per parish in the Northern region of Portugal (Source: Santos et al., 2019).

Figure 4

Percentage of artificialized territories in 2010 per parish in the Northern region of Portugal (Source: Land Cover Map of the Directorate-General of the Territory, 2010).

Figure 5

Social Vulnerability per parish in the Northern region of Portugal.

Figure 6

Regression tree for All-damaging flood cases, per parish, in the Northern region of Portugal.

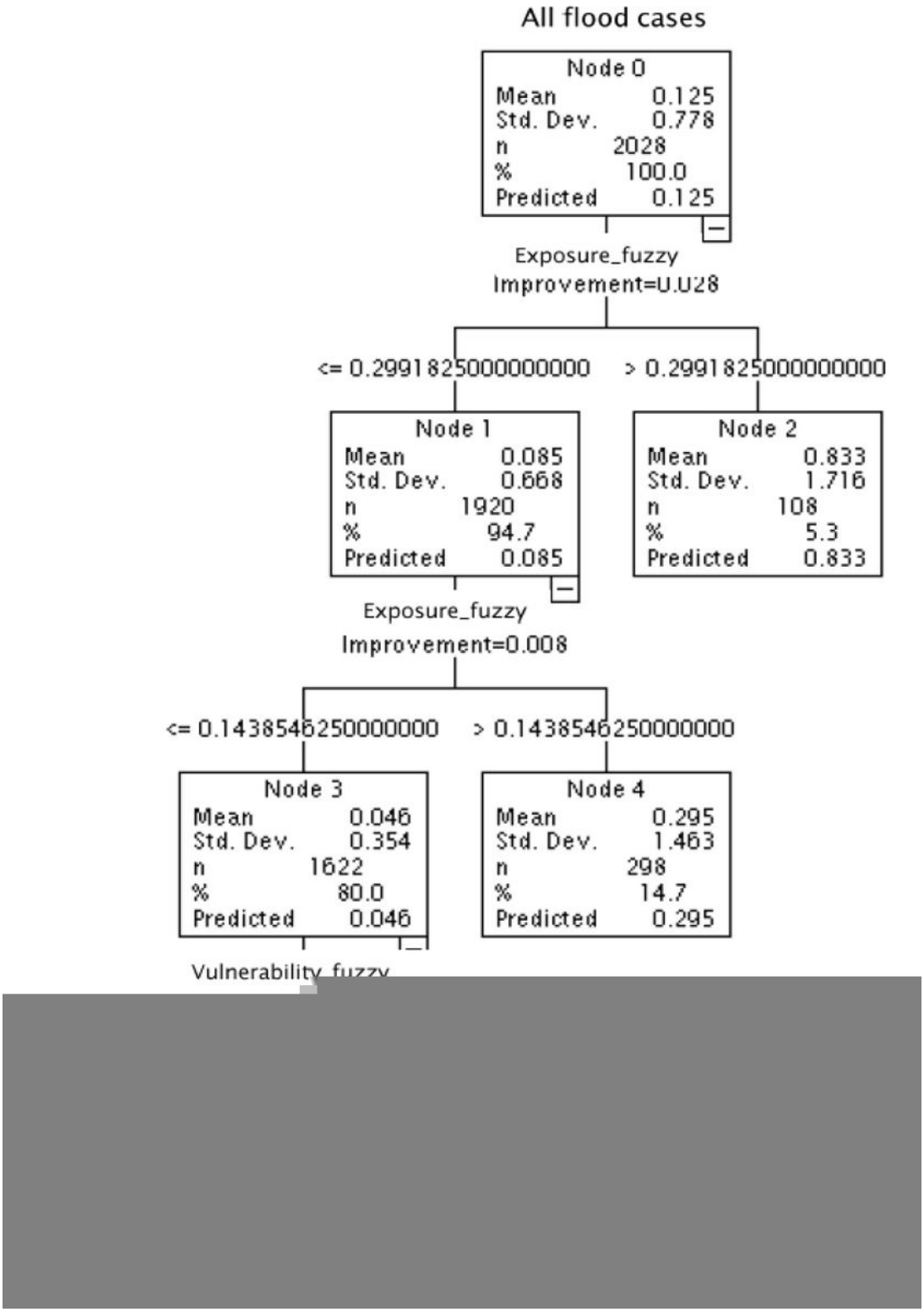

Figure 7

Regression tree for DISASTER-damaging flood cases, per parish, in the Northern region of Portugal.

Figure 8

Classification model using the All-damaging flood cases database, per parish in the Northern region of Portugal. Note: ASPFR study areas: (1) Chaves; (2) Esposende; (3) Ponte da Barca; (4) Ponte de Lima; (5) Porto; (6) Peso da Régua.

Figure 9

Classification model using the DISASTER-damaging flood cases database, per parish in the Northern region. Note: ASPFR study areas: (1) Chaves; (2) Esposende; (3) Ponte da Barca; (4) Ponte de Lima; (5) Porto; (6) Peso da Régua.

Figure 10

Susceptibility, exposure and vulnerability index and flood impacts in parishes covered and not covered by areas of potential significant flood risk (Floods Directive, 1st cycle).

Figure 11

Distribution of the Directive's parishes among the CART groups generated using the all-flood cases and the DISASTER flood cases databases.