International Journal of Disaster Risk Reduction 105 (2024) 104398

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: www.elsevier.com/locate/ijdrr

Predictive modeling of severe weather impact on individuals and populations using Machine Learning

Jordi Iglesias [a,*], Ildefonso Cuesta [a], Clara Salueña [a], Jordi Solé [b], David O. Prevatt [c], Alexandre Fabregat [a]

[a] *Universitat Rovira i Virgili, Av. Països Catalans, 26, Tarragona, 43007, Spain*
[b] *Universitat de Barcelona, Department of Ocean and Earth Dynamics, Barcelona, Spain*
[c] *University of Florida, Engineering School of Sustainable Infrastructure & Environment, Gainesville, FL, USA*

ARTICLE INFO

Keywords:
Machine learning
Severe weather
Emergency system incidents
Predictions
Catalonia

ABSTRACT

In this work, Machine Learning (ML) techniques are used to develop tools capable of accurately predicting the impact of severe weather events. We use readily accessible predictors, including daily meteorological data, basic demographics, geographic and terrain features, along with the number of daily meteorological incidents reported to the emergency services. The model was built using disaggregated data from the 947 municipalities from the region of Catalonia between January 1, 2015 and June 30, 2021. Catalonia's region is situated in the northeastern part of Spain along the Mediterranean Basin, and frequently experiences storms with strong winds and intense rainfall. In 2020, such weather events resulted in 64 injuries and damages amounting to approximately 70 million USD. The ML-based model presented in this study shows a predictive capacity for extreme weather risk superior to that of the daily meteorological warning system of the Meteorological Service of Catalonia. In addition, the model has been used to estimate how urbanization modifies the extreme weather impact. Given the expected increase in the intensity, frequency, and duration of extreme weather events in the context of global warming, the methodology presented in this work could be helpful in developing tools to assist emergency service managers and policy makers in making rapid and effective decisions.

1. Introduction

Natural hazard events are pervasive phenomena, exerting profound and varied impacts on ecosystems, wildlife, and human infrastructure worldwide [1]. From earthquakes to thunderstorms, floods to droughts, these occurrences pose significant challenges, particularly in densely populated areas where the potential for widespread devastation is highest [2]. The consequences of these events spans multiple fields including health, economy, infrastructure, population dynamics, and agriculture, highlighting the critical need for accurate prediction and effective mitigation strategies [3]. Moreover, within the context of global warming, it is predicted that the frequency, duration, and intensity of extreme weather events, and consequently their associated impacts will increase in the near future [4]. Most recent estimations found the costs of extreme events attributable to climatic change to be US 143 billion per year [5].

The severe weather, coupled with population expansion and urbanization, underscore the pressing need for proactive disaster preparedness and response measures [6]. Disparities in mitigating human and economic losses from extreme weather are often

* Corresponding author.
E-mail addresses: jordi.iglesias.d@gmail.com (J. Iglesias), ildefonso.cuesta@urv.cat (I. Cuesta), clara.saluena@urv.cat (C. Salueña), jordi.sole@ub.edu (J. Solé), dprev@ufl.edu (D.O. Prevatt), alexandre.fabregat@urv.cat (A. Fabregat).

https://doi.org/10.1016/j.ijdrr.2024.104398
Received 12 September 2023; Received in revised form 12 March 2024; Accepted 12 March 2024
Available online 15 March 2024

linked to income disparities among regions [7]. Resilient public systems ensure effective emergency management through Emergency Response (ER) organizations, which coordinate firefighters, police, civil protection, forestry service, and others. In a context where there is a growing awareness of the need to improve responses to extreme weather events, this study introduces a machine learning-based methodology for obtaining accurate and operationally capable computational tools from readily available predictors.

Within the field of artificial intelligence, machine learning enables the elucidation of the relationship between a set of predictors and a dependent variable from large volumes of observations. Its penetration across all scientific domains, employing a wide array of learning methodologies such as neural networks, support vector machines, and decision trees, has facilitated the construction of statistical models for numerous applications, significantly contributing to scientific advancements in fields like genomics, astronomy, physics, and environmental sciences. In the field of environmental science and engineering, the utilization of machine learning for predictive purposes stands out as an active area of research. It enables the exploration of vast datasets derived from observations, satellite data, and/or outputs from other models in fields such as climate modeling, ecological prediction, and environmental risk assessment.

According to the literature review on ML applications for natural disasters in [8], while most of the works (around 40%) focus on the post-event response, only 9% aimed at developing predictive tools. Within the field of damage prediction and mitigation, different works have examined different types of natural disasters ranging from flood events [9] to earthquakes [10] as well as economic and societal impact due to extreme weather outages [11]. On the extreme weather front, considerable efforts have been devoted to enhancing forecasts of severe events such as storms or hurricanes. A notable early contribution is highlighted in [12], which aims to estimate the probability of a tornado occurrence at a specific spatial location within a defined time window. Typically, these studies leverage historical records of extreme weather events to design improved strategies for allocating emergency response resources or predicting specific types of extreme weather phenomena.

The goal of the present study is to develop a holistic prediction model for disasters caused by adverse weather, that includes all types of phenomena affecting the population and individuals, which requires the intervention of emergency response services, using incidents, which we define as any weather-related emergency, as a proxy of the impact of severe weather on population.

In the region of Catalonia, the present work domain of interest, the 112 Emergency Call and Management Center is the public ER responsible for coordinating emergency calls. It provides immediate assistance 24 h a day, 365 days a year, managing calls related to all sorts of emergencies. For each call, the 112 service geolocates the incident and classifies the emergency according to the type of incident. It's essential to understand that an increase in the number of calls does not correlate with an increase in the number of incidents. Instead, an incident can be triggered by the ER based on one or multiple calls. One of the key points of the study is to utilize simple and accessible predictors that capture meteorological conditions as well as the characteristics of municipalities susceptible to being affected by severe weather. In contrast to some previously referenced articles, the number of features and atmospheric data has been streamlined to create a straightforward model, that would be optimally situated if meteorological organizations utilize it based on their local or regional forecastings, which are far more accurate than global models. Consequently, these agencies could effectively notify the ER in their application area. One of the model's advantages is the municipal-level spatial resolution. In this case, a single model encompassing 947 municipalities in the Catalonia region has been used, covering an approximate area of 32,000 square kilometers. Each municipality has a set of predictors that do not vary temporally and are assumed to remain constant over the years of study such as urbanization percentage or river density, along with other predictors associated with meteorology that change daily. It should be noted that this model is constrained in its ability to discern the specific cause of an incident, whether it is wind, rain, or any other meteorological factor. Therefore, on a day with incidents, an accompanying interpretation of the input data is necessary.

In this study, the model is compared with the current risk prediction system utilized by the Catalan government, which is described in Section 3.2. It is important to note that the existing risk prediction system, considers meteorological risk and focuses solely on the impact of weather to define a risk zone. In contrast, the model presented in this study aims to assess the risk to the population, as it has been trained to predict incidents triggered by citizen reports, alerts from authorities, or other sources that necessitate action from the ER.

Furthermore, the study investigates the impact of weather forecast accuracy on the performance of the ML model while maintaining unchanged municipality features. This analysis utilizes ten-day weather forecasts from the WRF model of METEOCAT which is a public company affiliated with the Department of Climate Action, Food and Rural Agenda of the Government of Catalonia and it is responsible for managing weather observation and forecasting systems (for more information, visit: https://www.meteo.cat). Additional details about the WRF model forecasts used by METEOCAT can be found on their webpage: (METEOCAT). A case study is conducted for the storm Gloria, which occurred from January 15th to January 25th, 2020, marking the most significant extratropical storm in the NW Mediterranean zone in the past 40 years. Classified as a cyclone and blizzard, it resulted in 64 injuries and prompted over 7000 interventions by the fire brigade across Catalan territory. The storm caused over 70 million US dollars in damages. Research articles [13–15] extensively studied the storm. The predictions from the previously trained ML model with the predictors using METEOCAT's forecasts are compared.

Finally, the study investigates the role of municipality characteristics on the extreme weather risk resulting from increased percentage of urban areas at the expense of forested areas or intensified urbanization on agricultural land cover.

2. Materials and methods

2.1. Area of study

Catalonia is a region located in the northwestern part of the Iberian peninsula, spanning over approximately 32,000 square kilometers. It is home to a population of approximately 7,675,000 inhabitants (see Fig. 1). The region is renowned for having one

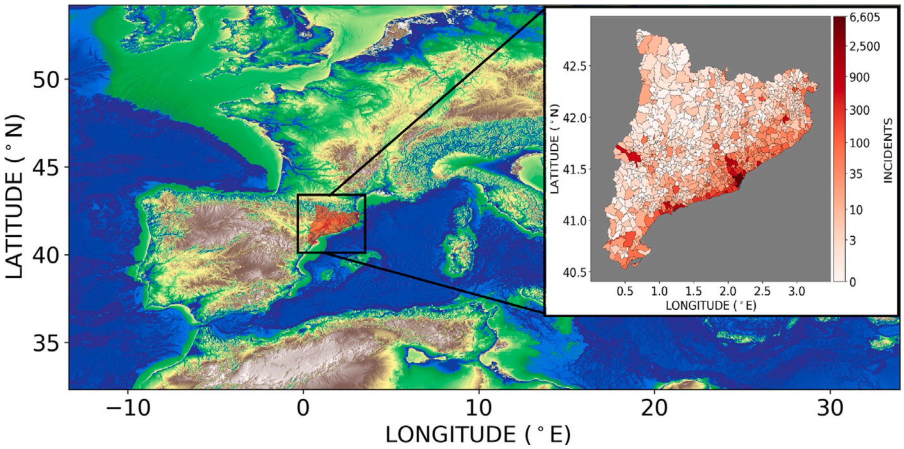


Fig. 1. Map of Europe with the location of Catalonia, color-coded to indicate the number of incidents between 2015 and 2021.

of the strongest and most diverse economies in the area, contributing a significant 19% share to the overall Spanish Gross Domestic Product. Its key economic sectors encompass manufacturing, tourism, services, and agriculture. Barcelona, the capital city, serves as a major economic center, hosting thriving industries such as automotive, chemicals, pharmaceuticals, textiles, and food processing.

Geographically, Catalonia boasts a varied landscape characterized by its borders along the Mediterranean coast to the east and the Pyrenees Mountains to the north. The central region showcases fertile plains and river valleys that sustain diverse agricultural activities, including wine and olive oil production. It experiences a Mediterranean climate, featuring mild winters and hot summers. The region encounters a wide range of weather patterns, occasionally experiencing heavy rainfall, particularly in the coastal and mountainous areas. Severe weather events, including storms with strong winds, can occur during specific seasons. The Meteorological Service of Catalonia (METEOCAT) plays a crucial role in providing accurate weather forecasts, monitoring the climate, and delivering meteorological information to support residents, businesses, and emergency services throughout the region.

The extreme weather patterns in Catalonia are characterized by several noteworthy phenomena. Strong easterly winds frequently accompany heavy rainfall and coastal flooding, especially in regions bordering the Mediterranean Sea. These conditions often lead to flooding along rivers that traverse densely populated areas, such as those along the Llobregat and Besòs rivers, which are home to millions of people. Additionally, Catalonia experiences summer droughts, intense snowfall in the Pyrenees and Prelitoral Ranges, impacting transportation and economic activities. Furthermore, the region is prone to forest fires, exacerbated by high temperatures and low humidity levels.

2.2. Data collection

In order to create a ML to predict the Emergency Response (ER) incidents, three different data types must be included. First, the municipalities characteristics, which are assumed to be static in the studied periods. This data includes municipality divisions shapefiles and information about rivers, population and extension, which were sourced from the Spanish and Catalan open data, *Iniciativa de datos abiertos del Gobierno de España*, *Centro Nacional de Información Geográfica (MINISTERIO DE TRANSPORTES, MOVILIDAD Y AGENDA URBANA)*.[1] Land usage data was obtained from the *Institut Cartogràfic i Geològic de Catalunya*.[2]

The second type of data is the historical dataset of incidents between January 1, 2015, to June 30, 2021, which will be used to train and validate the model. These data were obtained through a request to the *Centro de Atención y Gestión de Llamadas de Urgencia 112 Catalunya* a formal request is needed to get this data, in this references the databases of statics of municipalities are found; [16,17].

Finally, the third type of data for daily forecasts, the model was trained using meteorological forecasts for both, atmosphere and sea waves, to get wind speed, rainfall rate, temperature, snow depth and significant wave height. For wave forecasting data, the "Mediterranean Sea Waves Analysis and Forecast" [18] was used. For atmospheric forecasting, "CERRA sub-daily regional reanalysis data for Europe on single levels from 1984 to present" model [19] was employed. It is important to note that the CERRA model can be either reanalytical or forecasting, but for the purpose of training the model to predict future scenarios, weather forecast data was utilized. While data based on reanalysis offer higher accuracy and alignment with actual conditions, training the model on forecast data ensures its ability to predict incidents based on upcoming weather conditions. In the CERRA model, the data is obtained from a daily simulation that starts 6 h before the day in question. For example, to simulate a full 24-hour day, a prediction that starts at

1 https://www.mitma.gob.es
2 https://www.icgc.cat

Table 1
Predictors used in the ML model and their characteristics.

Predictor	Temporal variability	Statistic
Population density	Constant	–
Urban land cover percentage	Constant	–
Water cover percentage	Constant	–
River density	Constant	–
Forest land cover percentage	Constant	–
Agricultural land cover percentage	Constant	–
Temperature	Daily	Minimum
Temperature	Daily	Maximum
Significant wave height	Daily	Maximum
Rainfall rate	Daily	Maximum
Wind speed	Daily	Maximum
Snow depth	Daily	Maximum

6:00 PM the previous day and then the data used starts from 12:00 AM.[3] This approach allows for a balance between accuracy and having enough time for the development of atmospheric model variables in order to contribute properly to the forecasting outputs. Furthermore, this allocated time frame allows the Emergency Response (ER) services a time span to strategize and prepare for the forecasted incidents.

Table 1 presents an overview of the variables used in the study along with their temporal characteristics, which are either assumed to remain constant throughout the entire study period or vary on a daily basis. Some of the constant predictors are included to capture the density of built-up surfaces and population. We anticipate that densely populated and urbanized regions are more susceptible to a larger number of incidents. Similarly, the other constant features provide information on land cover and geophysical characteristics that may reflect distinct susceptibilities to extreme weather. Extreme daily predictors contain information regarding meteorological conditions. Land use/cover percentages are calculated by computing the fraction of each type within each municipality. River density is calculated using the total river length normalized by each municipality area. Additional dependence on the time of year is accounted by the sine and cosine of the month of the year. Due to its relevant role in impacting coastal areas, the maximum significant wave height has also been included. This predictor takes the zero value for all inland municipalities. Due to high values of correlation between accumulated precipitation and maximum daily rainfall rate, the latter has been selected. Despite the small values of snow depth over the temporal period considered in this work, the maximum value has been used to account for incidents at the highest altitudes, specially in terms of transportation network disruptions. These collection of constant predictors are shown in Fig. 2.

2.3. Data processing

The static data used include river density, population density, and percentage of land usage in: agriculture, forest, urbanized areas, and water coverage. The forecasts, which are raster data, were interpolated to a resolution of approximately 1.4×1 km and projected onto the WGS84 web Mercator projection. The data was then clipped to the boundaries of the municipality shapefiles. To obtain the predictors, a single value for each day, which is the most extreme in time and space, was used for each municipality.

The incident data is classified into several categories for ER organization (*Centre d'Atenció i Gestió de Trucades d'Urgència 112 Catalunya*[4]), one of which is meteorology. In 2.3.1 section is shown that the used model is a classifier with 5 different categories of predictions, these categories were chosen by doing the logarithm base ten of the maximum number of meteorology accidents in all datasets and dividing it into five equal distances, resulting into zero, one to five, six to thirty, thirty-one to one hundred sixty-five and more than one hundred sixty-five incidents.

After contacting the Catalan ER, they explained that sometimes a meteorology incident is classified as *other accidents*, *accidents* or *traffic* categories, being impossible to distinguish which incidents in these categories belong to meteorology. For example, a problem caused by meteorological conditions that affect roads could be classified under the *traffic* category, the category of *other accidents* includes weather-related disruptions to water, electricity, or gas supplies, as well as damage caused to animals. The *accidents* category refers to harm suffered by individuals in the sea, rivers, or building collapses, among other situations.

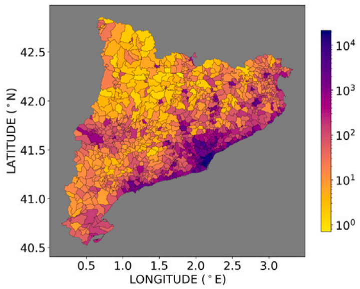
The boxplot shown in Fig. 3 was generated to compare the number of incidents of the last cited categories across the meteorology category, which revealed that incidents in these categories tend to increase when meteorological conditions are more severe, confirming the ER information.

To further explore the relationship between these three incident categories and meteorology. The total incidents were averaged and arranged into a table with five rows for the meteorology groups and five columns for the incident categories. Each average value in the table was then divided by the sum of all values in its corresponding row to obtain the table shown in Table 2.

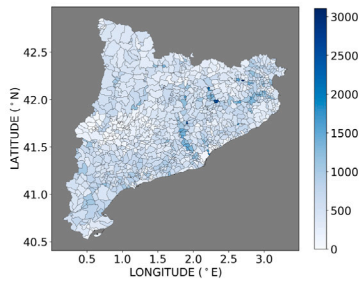
To estimate the total contribution of each incident category to the incidents, the number of meteorology incidents was counted for each municipality and day, and this count was added to a truncated value of the other categories, which was multiplied by the Table 2

[3] Additional information can be found in the documentation of CERRA: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-cerra-single-levels?tab=overview

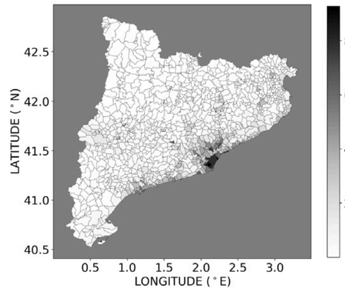
[4] https://112.gencat.cat/ca/inici/index.html



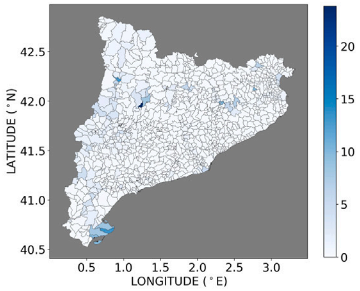
(a) Population density ($hab./km^2$), log. scale colours.



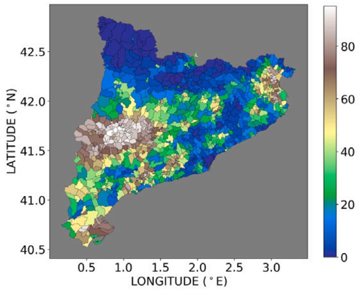
(b) River density ((m/km^2)).



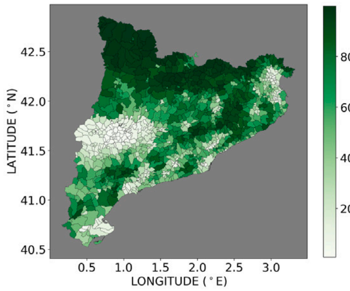
(c) Urban (%).



(d) Water (%).



(e) Agriculture (%).

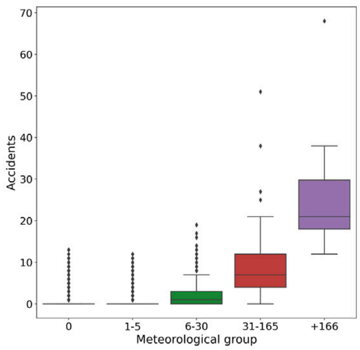


(f) Forest (%).

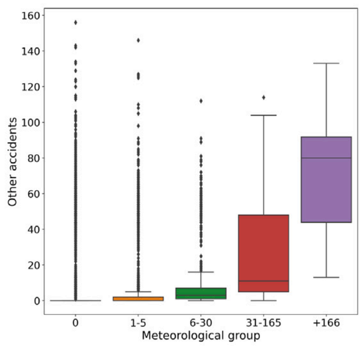
Fig. 2. Static data of the 947 municipalities.

Table 2
Table of scaling factors used into the different categories that are added to the meteorology incidents.

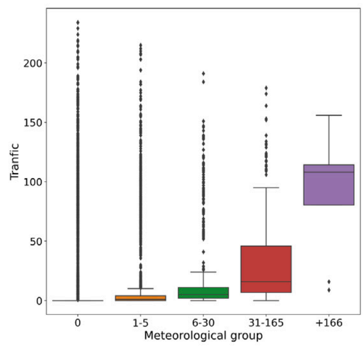
Meteorology group	Traffic	Other accidents	Accidents	Count meteorology	Count corrected
0	0.00	0.00	0.00	2 236 372	+0
1–5	0.04	0.03	0.01	9896	−186
6–30	0.10	0.09	0.05	806	+156
31–165	0.25	0.23	0.23	149	+25
+166	0.60	0.65	0.71	8	+5



(a) Accidents



(b) Other Accidents



(c) Traffic

Fig. 3. Box plot of incidents from categories that have the potential to be caused by meteorological factors according to the Emergency Response (ER), and their comparison based on the incident category corresponding to meteorology.

scaling factor. In general, the total number of incidents did not change drastically, as shown in this table, the "Count (meteorology)" column indicates the number of incidents in each meteorology category, while the "Count corrected" column represents the total contribution of the other incident categories. The first group did not undergo any modification, but some points in the second group were redistributed to the other groups.

Table 3
Final data distribution with training and validation.

Meteorology group	All data	Train data	Validation data
0	2 236 372	9480	344 276
1–5	9710	8462	1248
6–30	962	852	110
31–165	174	155	19
+165	13	11	2

Table 4
The top five models for regression and classification with their corresponding R-Squared and Root Mean Square Error, as well as Accuracy for classifiers. The model with the highlighted background was chosen for the results.

Regressors			Classifiers	
Model	**R-Squared**	**RMSE**	**Model**	**Accuracy**
Gradient Boosting Regressor	0.54	10.69	Random Forest Classifier	0.84
XGB Regressor	0.53	10.86	LGBM Classifier	0.84
Bagging Regressor	0.50	11.18	Bagging Classifier	0.83
Random Forest Regressor	0.45	11.77	Extra Trees Classifier	0.83
Extra Trees Regressor	0.36	12.62	Support Vector Classifier	0.82

The data used to train the model was selected from January 1, 2015, to July 30, 2021. For validation, a complete year had not been used in training (344,276 data points). The validating data year (2017), is the second year in terms of the severity of weather incidents, only surpassed by 2020 when the Gloria storm arrived in Catalonia. During the year 2017, Bruno Storm occurred. It was formed within a mobile trough,[5] that crossed the Atlantic towards Europe. Between the 24th and 25th of December, a closed vortex developed within this trough, located southwest of Ireland, approximately at 50° N-10° W. The trough and the low-pressure system then rapidly moved eastward towards Europe on the 26th and 27th, with the low-pressure system filling on the 28th. The low-pressure system and its associated fronts affected all countries in Western Europe, giving high wind velocities, for instance, 115 km/h in Reus[6] located at Tarragona province (South of Catalonia).

Hence, in terms of extreme weather, the year 2020 was used to train the model and show the most extreme case, while 2017 was used to validate the model because it contains all ranges of categories with extreme values. The present data exhibits a huge class imbalance with the class "0 incidents" being the most common for each of the 947 municipalities and around six years and a half of the period under study totaling into 2,236,372 data points as shown in Table 3. This imbalance is addressed using two strategies: Firstly, undersampling this most common class (0 incidents) with the same amount of data of the sum of the other classes. Secondly, "class weights" had been used to increase the relative importance of underrepresented categories.

All the data used was encoded into numeric values. To ensure consistency, all the data was normalized between zero and one. This normalization approach was found to offer several advantages after validating the model, compared to scaling the data to have a mean of zero and a standard deviation of one.

2.3.1. Model benchmark

The models and data have been processed using Python language modules. Table 4 summarizes the top models employed for regression and classification tasks, only the model with better accuracy was selected (Random forest classifier, with the highlighted background in the table. Although these models were not subjected initially to hyperparameter tuning, they have been used as a benchmark to select the most suitable for later tuning. The year 2017 was used for validation purposes, which shows that the regressors were unable to accurately predict the incidents. Further tests were conducted to manipulate the incident data using logarithms and root mean squares. This was done in order to reduce the skewness of the output. Additionally, predictors such as population density were also tested using a logarithmic transformation. However, in all cases, the results were significantly inferior to those obtained from the classification models.

2.3.2. Key hyperparameters

The hyperparameters allow improving the accuracy from 0.84 as seen in Table 4 to 0.87. Additionally, they enable modification of the prediction, ranging from conservative results with overprediction to scenarios where underprediction is more prevalent, allowing for the selection of the scenario that best fits the needs. The model with the best accuracy is the **Random Forest Classifier** from the *sklearn ensemble library*. The hyperparameters were tuned, and the best configuration was found to be 100 estimators, *log loss* criterion, a *max depth* of ten, *max features* equal to none, and class weight equal to *balanced subsample*. The remaining parameters were set to their default values. The *max depth* parameter affects the degree of over-prediction, with lower values leading to more over-prediction. The *balanced subsample* mode adjusts weights inversely proportional to class frequencies. To assess the stability of the model on the training data, five-fold cross-validation with stratification was performed using the *StratifiedKFold* function. The resulting accuracies were 0.79, 0.81, 0.80, 0.80, and 0.80. However, these tests only evaluate the variance of the training data, since the model is trained on all datasets except for 2017 which is used only for validation.

[5] A mobile trough in meteorology refers to a region of relatively low atmospheric pressure that is moving or shifting across an area.
[6] https://www.aemet.es/en/conocermas/borrascas/2017-2018/estudios_e_impactos/bruno

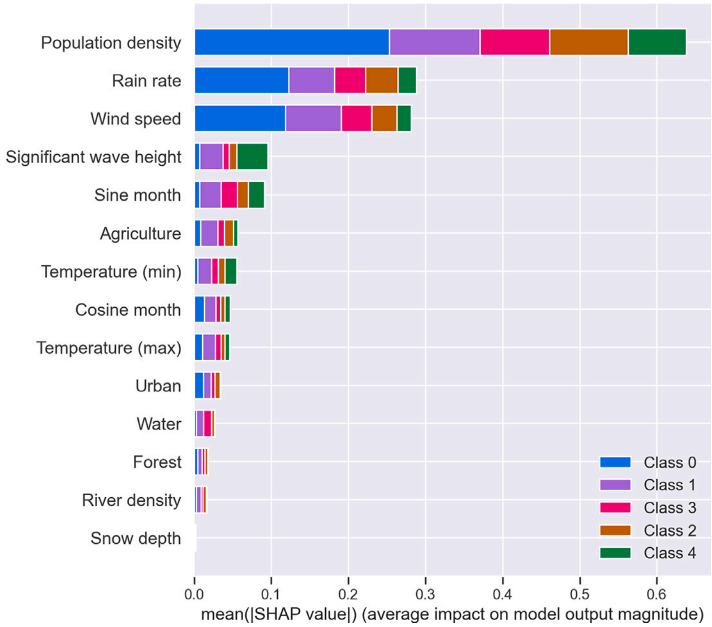


Fig. 4. Mean of SHapley Additive exPlanations (SHAP) values for each variable of the model.

2.4. Model features

The model obtained enables the prediction of incidents caused by severe weather that require emergency response. It operates at a municipal level with daily temporal resolution. The Random forest classifier can predict incident ranges categorized into intervals: 0, 1–5, 6–30, 31–165, and +166 incidents. The predictors utilized are based on municipality features, which can remain constant over time, and meteorological data, which varies daily.

3. Results and discussion

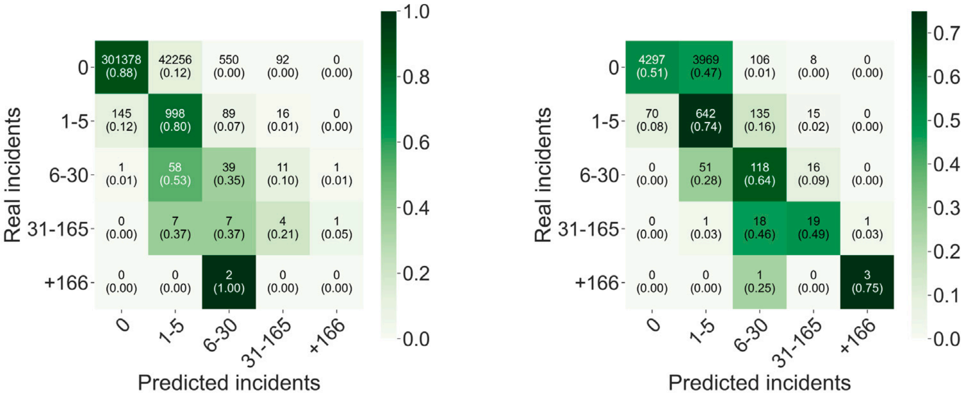
3.1. Prediction and validation

The SHapley Additive exPlanations (SHAP) analysis was performed and can be shown in Fig. 4, which differentiates the average impact on model output magnitude by each category. Population density is the most important one, which is logical for two main reasons. The first reason is the goal of the study, which predicts the effect of severe weather on the population. Zones with low density don't have an impact on the population. Therefore, an incident such as a tree falling in the forest, which no one notices or reports to emergency systems, won't be counted for the model and the data making a negative value of this predictor important for predicting in zero incidents class. The second reason is that severe damage is mainly caused in coastal zones. For example, Barcelona has the highest number of extreme incidents, and these coastal places have intense population densities.

This analysis also shows interesting effects for each class. For example, the significant wave height is not important in the class of zero incidents, but for the class of more than one hundred sixty-five incidents, it is the most important category after population density because of the previously explained facts. This figure shows the magnitude of SHAP, and for a deeper analysis, one for each category should be performed searching for the positive or negative effects of each variable for each category. However, this would be a lengthy explanation that could cause the reader to lose the scope of this work.

In Fig. 5a, the confusion matrix of the entire 2017 validation is displayed, which indicates good performance in the categories of 0 and 1–5 incidents. However, for the categories of six to thirty and more incidents, the prediction accuracy decreases. The *hyperparameters* of the model can be adjusted to correct the miss-prediction of high incident classes, but this may result in an increase in the number of over-predictions. To explain this behavior, one must understand the CERRA forecasting model. In general, a coarser mesh size tends to underestimate extreme rainfall and affects other prognostic variables. CERRA's documentation acknowledges this fact.

Furthermore, the model was tested from January 15th, 2020, to January 25th, 2020, which included the Gloria storm, using METEOCAT's forecasting of these events. As METEOCAT is the official Catalan government forecasting system, their expertise and experience in predicting the weather in the specific region of this study, along with the finer mesh resolution, result in a more accurate forecast of the storm. The data were pre-processed using the same methodology and used to validate the model, resulting in a more precise outcome. Fig. 5b shows the confusion matrix, where the points on the diagonal have a good tendency. However,



(a) 2017 validation with CERRA forecastings (b) METEOCAT Gloria storm

Fig. 5. The confusion matrix shows the real versus predicted categories, the number at the top is the total amount of numerical values and the number inside the parenthesis is the normalized one, the green color bar is also the normalized color.

Table 5
Classification report of RFC on left side and on the right side the RFC model using the 10 days of Gloria storm with inputs created with METEOCAT's forecastings.

Classification report								
2017 validation					Meteocat's validation			
Class	Precision	Recall	F1-Score	Support	Precision M	Recall M	F1-Score M	Support M
0	1.00	0.88	0.93	344 276	0.98	0.51	0.67	8380
1	0.02	0.80	0.04	1248	0.14	0.74	0.23	862
2	0.06	0.35	0.10	110	0.31	0.64	0.42	185
3	0.03	0.21	0.06	19	0.33	0.49	0.39	39
4	0.00	0.00	0.00	2	0.75	0.75	0.75	4
Accuracy			0.87	345 655			0.54	9470
Macro Avg	0.22	0.45	0.23	345 655	0.50	0.63	0.49	9470
Weighted Avg	1.00	0.87	0.93	345 655	0.89	0.54	0.63	9470

this validation shows lower accuracy, which can be explained because the 10 days of simulation of the storm had the most extreme weather of the last 40 years in the Catalonia region.

The metrics of the RFC model with 2017 validation are on the left side of Table 5, and the metrics of METEOCAT's simulation are on the right side. The metrics must be read in conjunction with the confusion matrix since the class imbalance is significant between classes 0 and +166, resulting in poor metric results. However, the confusion matrix shows that even if the prediction does not match the corresponding value, most of the time, it differs by only 1 or 2 adjacent categories.

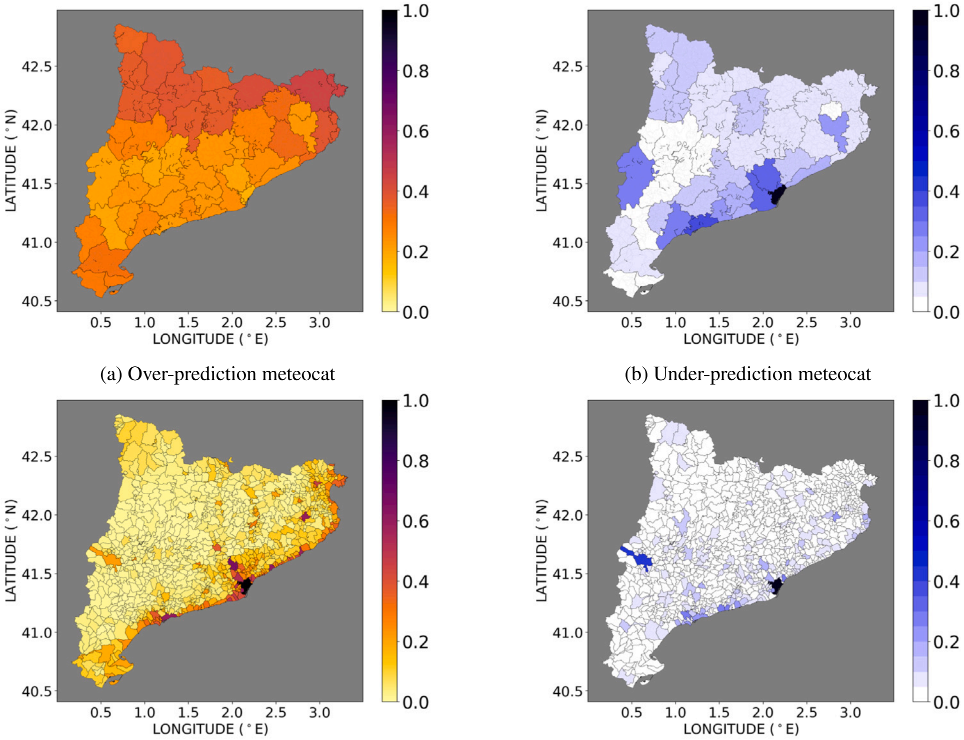
3.2. Comparison of ML with current meteorological risk prediction in Catalonia

As described earlier, the Random Forest Classifier (RFC) has demonstrated its ability to predict incidents based on meteorological forecasts as well as static data from municipalities, especially when atmospheric predictions are more precise. However, it is essential to compare the performance of this model with those currently in use as explained in the introduction section.

METEOCAT performs the meteorological risk danger system which includes different categories such as heat, cold, sea conditions, wind, accumulated snow, intensity, and accumulated rainfall. These categories are scored from 0 to 6 based on the level of danger for each countie, more information can be found on [20].

To make a comparison between the RFC and the risk system used by METEOCAT, all types of risk situations were aggregated by day and region and summed up. Since cold and heat cannot occur on the same day, the maximum value of these variables was used to create a logarithmic law based on 10, similar to the RFC incident classification. This logarithmic law was then divided into five value ranges, each corresponding to a category, to make it comparable to the incident system. Next, the categories of the risk system were subtracted from the incident categories for each day, and the results were separated into positive and negative values. A negative value indicates an over-prediction of the actual danger, while a positive value indicates an under-prediction. A value of 0 represents a perfect prediction of the real risk.

The absolute value of the difference between predicted and actual categories was calculated and summed up for each region. This process was repeated for both the RFC and METEOCAT models. Finally, the values were normalized between 0 and 1 in Fig. 6. The figures are comparable between the RFC and METEOCAT models, with different normalizations for under-prediction and over-prediction. The resolution of the present study model (RFC) is higher, in the case of METEOCAT it is in a county resolution. Each county is a tessellation of several municipalities, this is, each county is the results of combining several contiguous municipalities



(a) Over-prediction meteocat

(b) Under-prediction meteocat

(c) Over-prediction model

(d) Under-prediction model

Fig. 6. Maps that shows the comparison of RFC with the current risk prediction model.

that fit together without gaps or overlaps. This comprehensive coverage ensures spatial comparability across both models. High levels of over-prediction and under-prediction were observed in the Barcelona area for the RFC model due to the accumulation of extreme cases in that zone. However, in general, the RFC model has low levels of over-prediction and the METEOCAT model tends to over-predict for most regions as can be seen in the map. The RFC model exhibited very good and low values for under-prediction, indicating that the model rarely under-predicts.

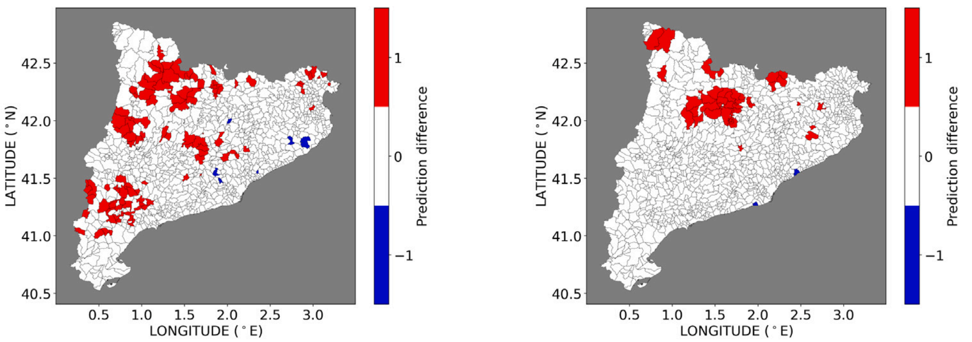
3.3. Municipality feature changes

The potential of predicting the impact of municipality changes on future incidents by separating variables and observing their effects lies with the RFC model. To test this potential, simulations were conducted on the distribution of all municipalities in Catalonia in 2017, using the CERRA model for forecasting data. Two cases were tested, on the one hand, the replacement of 10% of forest areas with urban areas and on the other hand, the replacement of 10% of agricultural zones with urban ones. The resulting changes are shown in Table 6, and several conclusions can be extracted.

The simulation results showed that the most severe scenario was observed when forest areas were replaced with urban areas. This meaningful statement can be explained physically because forests regulate temperature and humidity, creating a larger atmospheric boundary layer (ABL), which is lower than the ABL created by cities. However, in new urban constructions where tall buildings are lined up parallel to prevailing winds, wind flow between buildings can increase, resulting in a bigger wind speed. As seen in the data, wind speed is the dynamic variable that has the most significant impact on the creation of 1–5 incidents, and these cases are magnified, with 6350 more incidents in forest reduction and 3428 in agriculture reduction.

In very few cases, the urban increase decreased the severity of incidents from category 31–165 to 6–30, which is likely due to differences between villages and cities. The model is trained with the current locations, so if a place has a big forest area or agricultural area, the urban area tends to be small. Since most of the cases are villages, they are less prepared to handle heavy rains than cities. When a 10% of agriculture or forest area is converted into an urban area, the model can relate it to a city, notice that a city or village is not a variable of the model but statistically cities are more prepared for a rain rate that causes more incidents of that category. As a result, the incidents are predicted into one category lower.

These conclusions are possible explanations from the point of view of fluid dynamics, the model is trained with historical data and cannot evaluate the interactions between all pre-processed variables, the model does a prediction based on inputs and historic. Moreover, only the forest/agriculture to urban variables were changed, not the possible effect on the other variables like wind,



(a) High intensity weather 27/12/2017 (Bruno storm) (b) Medium intensity weather 2/6/2017

Fig. 7. Maps of category changes with an exchange of forest to urban features of the municipalities during high and mid-intensity storm days the values in red mean an increase of category and in blue a decrease of category.

Table 6
Table of changes of incidents predictions based on municipality feature changes.

Meteorology group	Forest to urban		Agriculture to urban	
	Increase	Decrease	Increase	Decrease
0	0	6336	0	3423
1–5	6350	0	3428	0
6–30	4	0	0	1
31–165	0	18	0	4
+165	0	0	0	0

temperature, etc. Nonetheless, it is reasonable to conclude that an increase in an urban area at the expense of agriculture or forest can have a negative impact on the amount and severity of incidents caused by severe weather.

Fig. 7, shows a comparison of two days, Fig. 7(a), is the map of a high-intensity storm (Bruno storm) that occurs on 27/12/2017, this is the most energetic storm of that year, and as explained before in wide municipalities this increment of forest to urban increases the incident category in a lot of municipalities, Fig. 7(b), is a day with a storm that is not strong enough to have a name, it also shows an increase of municipalities category of incidents.

4. Conclusions

The ML classification model used in this work can accurately predict the incidents occurrence and estimate the number of events in each municipality at a low computational cost. The model hyperparameters can be adjusted to obtain rough conservative predictions. However, it has a lower performance in predicting very strong storm events due to a lack of data availability and inaccurate forecasts. In particular, our Random Forest Classifier (RFC) algorithm provides an improvement over current risk prediction methods, which rely solely on CFD simulations and are focused on weather. As RFC uses characteristics of municipalities and statistics of past years can predict the impact on population. However, the quality of predictions is impacted by the low quality of prediction for precipitation and other atmospheric microphysics in the training data. The use of METEOCAT's forecastings has shown considerable improvements. Additionally, such models can be used to assess the impact of land changes, such as a 10% increase in urban areas instead of forest ones leading to a predicted increase of 6336 incidents category in a year.

The findings of this study could have important implications for improving the accuracy of severe weather incidents, enabling better preparation and response to extreme winds, rainfalls, etc. Moreover, this study shows the potential of ML techniques and thus aims to encourage the forecasting assessment institutions to use the data and combine the classic risk damage situation (which is based solely on the forecasts) with ML models, to provide the better preparedness for extreme weather impacts.

Software availability

Python 3.11.3 The scikit-learn version is 1.2.2. The matplotlib version is 3.5.2. The Seaborn version is 0.11.2. The geopandas version is 0.12.2. The shap version 0.41.0

Other software had been used (Matlab to interpolate Netcdf4 fields and R to generate raster of coma separated values from forecastings to shapefiles of geometries.)

CRediT authorship contribution statement

Jordi Iglesias: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Ildefonso Cuesta:** Data curation, Investigation, Supervision, Validation. **Clara Salueña:** Data curation, Investigation, Supervision, Validation. **Jordi Solé:** Data curation, Investigation, Supervision, Validation. **David O. Prevatt:** Supervision, Validation. **Alexandre Fabregat:** Data curation, Investigation, Methodology, Supervision, Validation, Writing – review & editing, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work has been funded by Spanish Ministerio de Ciencia, Innovación *y* Universidades through the grant PID2020-113303GB-C21 (MCIU/AEI/FEDER, UE), by the Generalitat de Catalunya through the grant 2017-SGR-1234 and by Universitat Rovira i Virgili through the grant 2019PMF-PIPF. The authors thank the Copernicus Marine Environment Monitoring Service (CMEMS) for data publicly available, and the Servei Meteorològic de Catalunya (METEOCAT) for additional supplied data and outputs of their models which were a crucial part of this work.

References

[1] O. Korup, J.J. Clague, Natural hazards, extreme events, and mountain topography, Quat. Sci. Rev. 28 (11–12) (2009) 977–990.
[2] R.M. Rauber, J.E. Walsh, D.J. Charlevoix, Severe and Hazardous Weather: An Introduction to High Impact Meteorology, Kendall Hunt Publishing Co., 2017.
[3] T. Miyoshi, G.-Y. Lien, S. Satoh, T. Ushio, K. Bessho, H. Tomita, S. Nishizawa, R. Yoshida, S.A. Adachi, J. Liao, et al., "Big data assimilation" toward post-petascale severe weather prediction: An overview and progress, Proc. IEEE 104 (11) (2016) 2155–2179.
[4] K.L. Ebi, J. Vanos, J.W. Baldwin, J.E. Bell, D.M. Hondula, N.A. Errett, K. Hayes, C.E. Reid, S. Saha, J. Spector, et al., Extreme weather and climate change: Population health and health system implications, Annu. Rev. Public Health 42 (1) (2021) 293–315.
[5] R. Newman, I. Noy, The global costs of extreme weather that are attributable to climate change, Nature Commun. 14 (1) (2023) 6103.
[6] P.A. Raschky, Institutions and the losses from natural disasters, Nat. Hazards Earth Syst. Sci. 8 (4) (2008) 627–634.
[7] A. Loretti, Y. Tegegn, Disasters in Africa: Old and new hazards and growing vulnerability, World Health Stat. Q. 1996; 49 (3/4): 179-184 (1997).
[8] V. Linardos, M. Drakaki, P. Tzionas, Y.L. Karnavas, Machine learning in disaster management: Recent developments in methods and applications, Mach. Learn. Knowl. Extract. 4 (2) (2022) 446–473.
[9] P. Prasad, V.J. Loveson, B. Das, M. Kotha, Novel ensemble machine learning models in flood susceptibility mapping, Geocarto Int. 37 (16) (2022) 4571–4593.
[10] M.S. Amin, H. Ahn, Earthquake disaster avoidance learning system using deep learning, Cogn. Syst. Res. 66 (2021) 221–235.
[11] P.L. Watson, A. Spaulding, M. Koukoula, E. Anagnostou, Improved quantitative prediction of power outages caused by extreme weather events, Weather Clim. Extremes 37 (2022) 100487.
[12] I. Adrianto, T.B. Trafalis, V. Lakshmanan, Support vector machines for spatiotemporal tornado prediction, Int. J. Gen. Syst. 38 (7) (2009) 759–776.
[13] A. Amores, M. Marcos, D.S. Carrió, L. Gómez-Pujol, Coastal impacts of Storm Gloria (January 2020) over the north-western Mediterranean, Nat. Hazards Earth Syst. Sci. 20 (7) (2020) 1955–1968.
[14] F.X. Roig-Munar, J.Á. Martín-Prieto, A. Rodríguez-Perea, B. Gelabert, Gloria storm effects on the coastal boulders East of Minorca (Balearic Islands), Environ. Earth Sci. 82 (7) (2023) 176.
[15] J. Iglesias, I. Cuesta, C. Salueña, J. Moré, J. Solé, Analysis and comparison of coupled and uncoupled simulations with the COAWST model during the Gloria Storm (January 2020) in the northwestern Mediterranean sea, Environ. Model. Softw. 169 (2023) 105830.
[16] Datos abiertos Gob. Esp., Datos abiernos del gobierno de esp, 2023, https://datos.gob.es/es/. (Accessed: 05 April 2023).
[17] Institut Cartogràfic i Geològic de Catalunya, Mapa de sòls de Catalunya 1:250.000 (MSC250m, 2023, https://www.icgc.cat/Administracio-i-empresa/Serveis/Sols/Mapa-de-sols-1-250.000. (Accessed: 05 April 2023).
[18] Copernicus Marine, Mediterranean sea waves analysis and forecast, 2023, https://data.marine.copernicus.eu/product/MEDSEA_ANALYSISFORECAST_WAV_006_017/description. (Accessed: 05 April 2023).
[19] Copernicus, CERRA sub-daily regional reanalysis data for Europe on single levels from 1984 to present, 2023, https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-cerra-single-levels?tab=form. (Accessed: 05 April 2023).
[20] METEOCAT danger, Danger meteorology situation, 2023, https://www.meteo.cat/wpweb/divulgacio/la-prediccio-meteorologica/situacio-meteorologica-de-perill/. (Accessed: 05 April 2023).