



# Rapid natural hazard extent estimation from twitter data: investigation for hurricane impact areas

Janine Florath<sup>1,2</sup> · Jocelyn Chanussot<sup>3</sup> · Sina Keller<sup>1</sup>

Received: 19 December 2022 / Accepted: 28 January 2024  
© The Author(s) 2024

## Abstract

Natural hazards have occurred more frequently in the past years and pose a severe risk to human life. Their extents and, thereby, the most heavily affected areas must be estimated at the earliest to limit damages or initiate rescue services. For such estimations, a widely available data source, which is comparatively responsive to short-time changes, is needed and provided by volunteered geographic information (VGI) data. Tropical cyclones are natural hazard events that can cause enormous spatially extended damage. In this study, we introduce Machine Learning approaches such as Extremely Randomized Tree (ET) and Geographically Weighted Regression for estimating hurricane-impacted regions from VGI data. In addition to the general approximate track extent estimation, we also evaluate the possibilities of temporal estimation of track development from VGI data. Different scenarios are evaluated, and we find that the results mainly depend on the choice of the geographical splits for training and test data for the underlying regression task. Suitable splits lead to  $R^2$  of 99% in the best cases with the ET model. The estimation results are satisfying when considering the temporal aspect and represent a use-case scenario. Such a combination of Machine Learning approaches and VGI is a simple and fast approach for early natural hazard estimation.

**Keywords** Machine learning · Spatial analysis · Temporal development · Estimation framework · Volunteered geographic Information data · Extremely randomized tree

---

✉ Janine Florath  
janine.florath@kit.edu

Jocelyn Chanussot  
jocelyn.chanussot@inria.fr

Sina Keller  
sina.keller@kit.edu

<sup>1</sup> Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology, Englerstrasse 7, 76131 Karlsruhe, Germany

<sup>2</sup> GIPSA-Lab, Université Grenoble Alpes, CNRS, Grenoble INP, 38000 Grenoble, France

<sup>3</sup> Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France

## 1 Introduction

Natural hazards and disasters have occurred more frequently and with a more vigorous intensity recently (Ward et al. 2020). Such hazards threaten humans' physical and socioeconomic well-being, nature and ecology, and infrastructures such as the road network (Songwathana 2018; Wisner et al. 2014). During hazards and their resulting critical situations, first responders, such as rescue services, have a crucial role in mitigating occurring events and any threats. One key piece of information is, for example, a rapid estimation of the hazardous location and extent. This estimation allows first responders to prepare for possible hazardous effects and to bring humans or critical infrastructures to safety (Hao and Wang 2020; Harrald 2006; Seaberg et al. 2017). Estimating the precise extent of a hazard is crucial but remains challenging due to limitations such as the absence of readily available real-time data and information.

Characteristics, impacts, locations, hazard-impacted areas, and the extent of different natural hazards vary highly (Lindell and Prater 2003; André 2012). Examples of hazards with significantly spatially extensive damages are tropical cyclones (TC) in the northeastern Pacific Ocean or the Atlantic Ocean (Cox et al. 2018; Walker et al. 2006). We refer to the latter as hurricanes. One of the latest, heavily media-covered natural hazards was Hurricane Ida affecting the mainland USA in 2021. This category four (wind speeds of 209 to 251 km/h) hurricane was the fifth-costliest tropical cyclone on record, with a damage cost of \$75 billion (NOAA-NCEI 2022).

When focusing on the extent estimation of a hurricane and the estimation of its track, hurricane prediction is conventionally employed. It uses weather data such as cloud-top temperatures and water vapor (ECMWF 2021; Birchfield 1960). This specific information is mainly extracted and surveyed based on satellite data (Zhang et al. 2019). However, such derived predictions usually have low accuracy, up to 150 km, and mostly around 50 km to 80 km errors shortly before landfall (Schultz et al. 2021; Cangialosi 2017; Bilskie et al. 2022). Concerning hurricane tracks, in general, specified cones of uncertainty are approximately 70 km for 24 h forecasts (NHC 2022a). The National Hurricane Center's (NHC) Tropical Cyclone Public Advisories (TCPAs) (NHC 2022c) contain information about when the hurricane's eye passed at which location approximately and achieved the highest locations accuracies with 46 km. The TCPAs are issued only every six hours. Therefore, to delimit the actual impact zones of possible destruction during or shortly after the hurricane passed that require help or interventions (e.g., by rescue services), such predictions are neither precise nor provide the necessary information in real-time. More precise information concerning the actual (not predicted) hurricane track is only provided after the hurricane events (Knapp et al. 2018). The best track data usually take a few months to a year or more after the hurricane (Knapp et al. 2010) to be finalized. Particularly in the context of the first response, obtaining this precise information about the hurricane's impact zone, cannot be delayed for such a long time. Unlike track predictions, which offer valuable guidance for preparedness before the hurricane's occurrence, the actual impact zone estimation should provide real-time information necessary for timely disaster response.

Therefore, we introduce an approach to estimate hurricane impact areas in near real-time during or shortly after hurricane passage. It is essential to emphasize that our approach is not intended for predicting newly emerging hurricanes. Instead, it is designed to retrospectively leverage data from one hurricane for which we already have information about its impact at location 1 during timestep 1 to estimate the impacted area for location 2 at timestep 2. To deal with the identified gap of near real-time,

higher-accuracy hurricane impact estimation, one must consider new (data-driven) approaches with alternative data sources.

One example of widely available data that is comparatively responsive to short-time changes is volunteered geographic information (VGI) data (Goodchild 2007; De Albuquerque et al. 2015). VGI data include all geospatial data generated by nonprofessionals. Though many data sources exist in the category of VGI, we focus on Twitter (now X) data, which has proven a valuable source of real-time information and insights (e.g., Wang et al. (2016); Dittrich (2016); Imran et al. (2016)). While not all so-called tweets are georeferenced, a sufficiently large number of them are. Therefore, they provide a substantial and diverse source of geospatial information for our analysis.

However, most studies do not include VGI data for a detailed spatial analysis regarding natural hazard estimation. Instead, two different topics are addressed: first, detecting natural hazards per se, as presented by, e.g., Benson et al. (2011); Dittrich and Lucas (2014); second, text messages are used for retrieving specific information on damage and hazard severity analysis as given in, e.g., Wang et al. (2016); Kumar et al. (2014); Fohringer et al. (2015).

To conclude, using and analyzing VGI data for spatial analysis for a more precise and real-time natural hazard extent estimation has not been conducted frequently (Guan and Chen 2014; Wang et al. 2016).

In this study, we investigate and evaluate the possibilities and limitations of tweets for the spatial analysis of natural hazards. This approach includes Machine Learning regression, which estimates the distance between the geolocation VGI points (tweets) and the hurricanes' tracks purely data-driven. Estimating these Euclidean distances can help localize areas more heavily affected by hurricanes as located closer to the hurricane eye. Furthermore, from these distances, the hurricane track course could be inferred explicitly if relevant for disaster management agencies. The underlying task, spatial analysis of natural hazards (here: hurricane track) includes two aspects to be addressed: spatial hazard estimation and temporal hazard estimation.

- *Spatial Hazard Estimation (Extent Estimation)*: First, the feasibility of hurricane track estimation from VGI data is evaluated in general by training and testing several Machine Learning (ML) approaches to estimate the distance of respective points to the reference track. We investigate whether or not and in which accuracy range our proposed approaches can estimate a natural hazard event from Twitter data.
- *Temporal Hazard Estimation (Development Estimation)*: The temporal aspect plays a role in several natural hazard scenarios. For example, hurricanes are agile natural hazards with a fast movement rate compared to hazards with lower degrees of temporal development, e.g., wildfires or floods.

Estimating the distance of VGI points to an unknown track in the second timestep can be a real-time estimation of affected zones of possible destruction during or shortly after the hurricane has passed.

We shortly describe the study area, data basis, and methodology in Sect. 2. Next, the results (see Sect. 3) of the regression are explained. Finally, we discuss the results (see Sect. 4) and conclude the study with a resume (see Sect. 5).

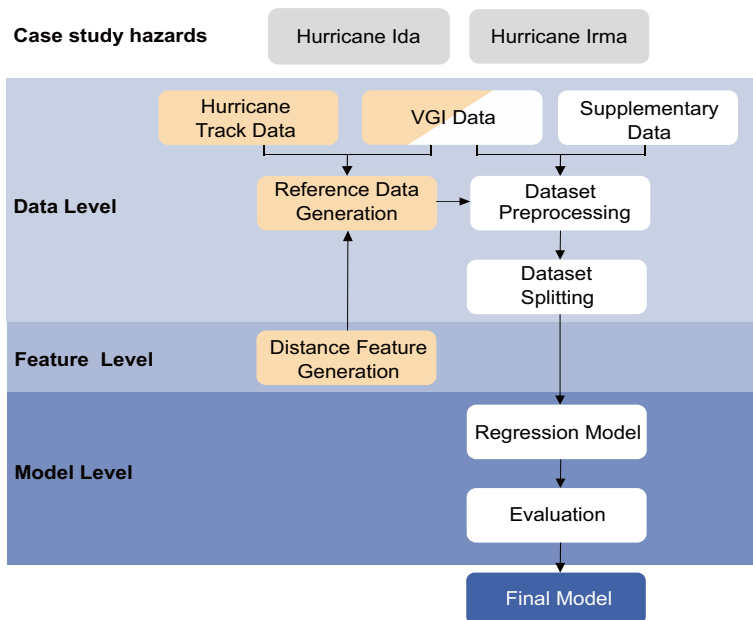
## 2 Datasets and methodology

Since the spatial analysis of hurricane tracks based on VGI data is a regression task in our case, we conduct an appropriate workflow. Figure 1 shows our applied regression framework structured into different levels. First, the data basis is described for the case study hazards covering the input data, namely hurricane track data, VGI data, and supplementary data (see Sect. 2.1). Next, the generated dataset is described in Sect. 2.2, including the reference generation in the feature level in Fig. 1. Note that we refer to the combination of input features and desired output data as a datapoint. On the data level, the generated datasets are split (see Sect. 2.3), which is necessary for the regression's training and evaluation. Finally, included in the model level of the framework, selected regression models, their optimization, and the model evaluation metrics are stated in Sect. 2.4.

### 2.1 Input data

Since we aim to develop a generic approach, including the model for hurricane track mapping, the study area selection has to meet two main criteria. First, the study area should include regions characterized by strong hurricane events in recent years. Second, the regions should show the existence of VGI data posting activities.

We rely on three types of input data listed below and described in further detail in the following subsections:



**Fig. 1** Visualization of the regression framework for natural hazard estimation divided into the data level, the feature level, and the model level. Adopted from Florath and Keller (2022). (VGI: volunteered geographic information)

1. Firstly, hurricane track data are used (Sect. 2.1.1). They are acquired from the NHC of the National Oceanic and Atmospheric Administration (NOAA) (NHC 2022b). These data are used to generate the reference data, which is necessary for the model's training and test (see Sect. 2.3 for more details).
2. Secondly, we employ VGI data in the form of Twitter data (Sect. 2.1.2). The information about the tweets' locations is the main feature used to estimate hurricane tracks.
3. Finally, we include other supplementary data (Sect. 2.1.3), like population density. These data are used as an additional input feature since different underlying data can affect the occurrence and the accumulation of VGI data in a specific region.

### 2.1.1 Hurricane data

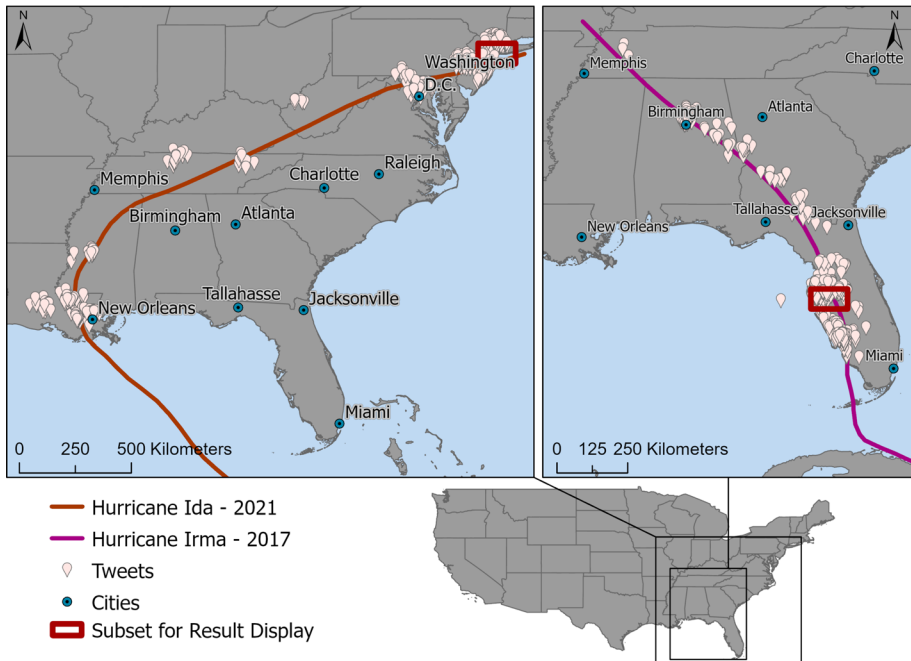
We rely on two exemplary hurricanes in the region of the mainland USA as case study hazards. The considered hurricanes are hurricanes Ida 2021 and Irma 2017. Their data are applied for the regression approaches' training, validation, and test. Hurricane Ida formed on August 26, 2021, and dissipated on September 4, 2021. It made landfall on the USA coast in Louisiana on August 29, 2021. Hurricane Ida caused destruction in many states of the USA and became the second-most damaging and intense hurricane to strike the USA state of Louisiana on record, behind Hurricane Katrina (NOAA-NCEI 2022). Hurricane Irma formed on August 30, 2017, and made its first landfall on Cudjoe Key, Florida, USA, on September 10, prior to another landfall on Marco Island, Florida, later that day. It dissipated on September 13, 2017, and became the sixth costliest USA Atlantic hurricane (NOAA-NCEI 2022).

Each hurricane's data are provided in vector files, showing the track of the hurricane (Fig. 2). These vector data provide the calculated best track of the hurricane usually generated after the season when forecasters gather all the available information from different sources and datasets (Knapp et al. 2010). We use these data representatively to develop our approach for hurricane estimation since they come with higher level accuracies. These would be replaced for other data in near real-time application scenarios of our approach (see Sect. 4.1.1).

Furthermore, to evaluate the temporal aspect of the hurricane development, we need information on the time-dependent location of the hurricane. We use the NHC's TCPAs (NHC 2022c) to extract information about when the hurricane's eye passed at which location approximately. The TCPAs are issued every six hours and provide the actual hurricane's eye geographic location-specific time. We employ these data to distinguish which tweets are during-hazard or post-hazard at a specific location for a given time. We choose the respective geolocations day-wise for our analysis of the temporal aspect of hurricane track estimation.

### 2.1.2 VGI data: Twitter tweets

We use Twitter data as VGI data to benefit from the high usage and easy access. The selected tweet data are available on Twitter's download API accessed via Python. We rely on the direct location extraction method, which can obtain the most accurate locations. Locations can be directly extracted from metadata obtained with the text data when accessing VGI. Twitter delivers a JSON file that provides `coordinates` and/or a `place` field where the tweets are created (X Corp. 2023). However, we employ only the `coordinates` given in the metadata to achieve the highest possible accuracy for our natural hazard geolocalization from the tweets.



**Fig. 2** Visualization of the selected hurricanes in the USA with their hurricane tracks (colored lines) and the respective collected Twitter data locations (colored markers). Furthermore, the subsets used for the later visualization of the results are displayed (colored boxes). Data basis: © 2018 GADM. Projection: WSG84

These coordinates originate from the GPS sensor of the users' mobile devices in combination with other positioning techniques like Wi-Fi. With these sensors, a reasonable scenario can achieve location accuracy of 2 m to 100 m (Dittrich 2016).

We search for tweets including the term *hurricane* and the particular hurricane's name, e.g., *Ida*, for locations close to the hurricane track. A point location is chosen for spatial constraint on the predicted hurricane track center line of ca. 46 km accuracy, and the maximum search radius of 25 mi is set. Temporally, we restrict the search to the days from the hurricanes' landfall on the USA coast to their dissipation (over the USA mainland). Furthermore, we extract the tweets temporally iteratively, choosing only those that were posted during- or post-disaster at the specific location. Overall we collected 1,375 messages with coordinates. Note that in the following, we refer to tweets or tweet points as the geographical location point of the respective tweet messages, not the message itself. Figure 2 visualizes the geolocations of the VGI data and the investigated hurricane tracks used for the regression approaches' training, validation, and testing. Detailed information about the distribution of the training, evaluation, and test data subset concerning the different regions is given in Sect. 2.3. Note that the VGI data appear in rather distinct clusters, mainly depending on the location of populous cities.

### 2.1.3 Supplementary data

The added supplementary data are included as they can be more explanatory for the occurrence of tweet points than the distance to the hurricane track itself. For example, more VGI

data occur in areas with higher population density. Different data affecting tweet data, e.g., population data, have been investigated in several studies (Jiang et al. 2019; Wang et al. 2016). In our case, the supplementary data are population density, altitude, slope, aspect, distance to the nearest road, and Digital Divide Index (DDI) (Forati and Ghose 2022). A systematic relationship between population density and Twitter use has been reported by Arthur and Williams (2019). Altitude and slope are additional factors influencing tweet density and may have interaction effects with population density. By including them separately, we allow the model to capture potential nonlinear relationships or interactions that might be missed when they are combined in a single feature. Slope and aspect (orientation of slope, measured clockwise in degrees from 0 to 360) are calculated based on the altitude using the software ArcGIS (ESRI Inc. 2020). Aspect can influence temperature, vegetation, and potentially the desirability of locations for various activities, which may, in turn, be related to tweet density. The distance to the nearest road can also influence tweet density through more people mostly staying near roads and not staying in entirely secluded regions. The distance is calculated from an OSM road feature dataset using ArcGIS. The DDI (Gallardo 2020) is an index that measures the physical access and adoption of broadband infrastructure and the socioeconomic characteristics that limit their use. It comprises two scores: the infrastructure/adoption (INFA) and the socioeconomic (SE) scores. The DDI ranges from 0 to 100, where 100 indicates the highest digital divide. The complementary data are pulled or derived from different OSM and non-OSM sources, as displayed in Table 1.

## 2.2 Reference and dataset generation

In this section, we describe the reference data generation and the combination of our input data (see Sect. 2.1) to create a complete dataset that can be used for a regression task. For our study's aim, it is indispensable to have reference data that differentiate the distance from the tweet points to the natural hazard event. Therefore, we create a respective dataset with all datapoints containing the input features (attributes) and the corresponding labels of the distance from the tweet point to the hazard (hurricane track). We calculate the geodesic distance of each tweet point to the nearest natural hazard (hurricane track) point in km for each tweet point. This distance serves as the label for the regression task and is the value estimated by the regression approaches.

**Table 1** Supplementary data with their sources and years of creation

Data	Data Source	Year
Population Density	Worldpop.org (2022)	2020
Altitude	Hijmans (2022) - CGIAR SRTM (3 s resolution)	2018
Slope	Derived from altitude	2018
Aspect	Derived from altitude	2018
Distance to nearest road	Hijmans (2022) - Digital Chart of the World	Unknown
DDI	Gallardo (2020)	2019

CGIAR SRTM: Consultative Group on International Agricultural Research - Shuttle Radar Topography Mission, DDI: Digital Divide Index

In the following, we summarize the steps to generate the input features from the input data: First, we preprocess the Twitter data locations to be displayed as point features in a Geographic Information System (GIS) environment, as displayed in Fig. 2. Each point possesses its  $x$  and  $y$  geographical coordinates, which should be used as input features for our regression model. Due to the significant spatial extent of our study area, the geographical coordinates exhibit substantial variability. This wide range of values in the input feature might impede a regression model's ability to form generalized patterns based on location. We convert our geographical coordinates into Universal Transverse Mercator (UTM) coordinates to address this issue. In this coordinate system, the Northings and Eastings are expressed in meters and maintain consistent values within each zone. The UTM coordinates exhibit reduced spatial variation compared to geographical coordinates. These two UTM coordinates are the first two input features for the Machine Learning (ML) models. Furthermore, we retain the posting time information for the tweets. This information is not used as an input feature but is relevant for the temporal hazard evaluation, especially for partitioning datapoints in during- and post-hazard tweets. The tweet's text information is discarded.

Next, we combine the tweets' information with the supplementary data. As a result, we obtain five (six with DDI) new input features that are added to the dataset. All supplementary data are used as general input features for our regression models, except the DDI, which is only used in models when indicated (as *DDI*, see Sect. 3.1). In conclusion, the dataset contains seven (eight with DDI) features, the distance label, and the date column for 1375 datapoints.

### 2.3 Dataset preparation

After generating the reference data and the features, we prepare the dataset and split it for the regression task (see Fig. 1). An independent splitting of the subsets in training, validation, and test datasets is necessary when evaluating any model's regression performance. We investigate different splits.

- *Spatial Hazard Estimation (Extent Estimation)*: We first use standard splitting with a ratio of 60 : 20 : 20. Standard ML guidelines are followed with the chosen split ratio (see, e.g., Kattenborn et al. (2021)) to evaluate the influences of the supplementary data and the model choice on the feasibility of our spatial analysis approach. However, since we work with geographically clustered data, the sets obtained from this random splitting procedure may not accurately represent our dataset. Therefore, to evaluate and potentially avoid difficulties resulting from a random split, we test several other splits for the spatial hazard estimation that consider the geographical distribution of the datapoints. For a general regression feasibility evaluation, these splits comprise a *geographically-balanced (GB) split* and a *distance-balanced (DB) split*. A GB split ensures that datapoints of each cluster (and therefore region) (see Fig. 2) are included in each set. A DB split creates buffer zones around the track to ensure that sample datapoints of each distance range and from each track side are included in each set.
- *Temporal Hazard Estimation (Development Estimation)*: We investigate spatially and temporally different splits that represent different possible real-life scenarios. *Stationary* hereby, refers to using datapoints of the same geographical area in timestep 1 (training) and in timestep 2 (testing). Note that this is rather useful for more stationary hazards like fires, which cover approximately the same area in timesteps 1 and 2. On the



contrary, *nonstationary* refers to using datapoints of different geographical areas passed by the hurricane (timestep 1 for training and timestep 2 for testing). The respective splits are: a) a *temporal-stationary (TS) split* that aims to estimate the natural hazard over time in a stationary area, b) a *temporal-nonstationary (TN) split* that aims to estimate the natural hazard development over time, c) a *temporal-nonstationary (TN)-all split* analogous to (b) but additionally including past hurricane (Irma) datapoints, and d) a *temporal-nonstationary (TN)-2 split* analogous to (b) but splitting into less training datapoints and more datapoints for testing of the models.

In these last splits, the number of datapoints in the training and test sets of the temporal splits is defined by the number of tweets in the respective selected spatial areas. Contrary to the predefined split ratios for the baseline random split (e.g., 60 : 20 : 20), the training-test ratios vary for the temporal split according to available tweets.

Table 2 summarizes the number of datapoints and dates of tweets included in the respective train and test sets for the different splits. Furthermore, we establish validation datasets for hyperparameter optimization by further splitting the initial training set into two sets with a ratio of 75 : 25 randomly.

### 2.4 Estimation approaches

This subsection describes the model level (see Fig. 1), including the ML models to solve the underlying regression task, as well as their optimization and evaluation. Estimating hurricane tracks from tweet data can be done using regression of the nearest distance from the tweets to the track. Since we have several input features and are trying to depict a rather complex connection with our output label, we need to employ a sophisticated regression approach. Therefore, we apply two different ML models to evaluate their estimation performances on the different splits and datasets (see Sects. 2.3 and 2.1).

We use an Extremely Randomized Tree regressor (ET). ET is applied as a tree-based regression model and is associated with decision trees (DTs). Generally, they include a root and a leave node linked by branches. During the training of DTs, the data of the respective dataset are split at every branch. These splits generate subsets, which correlate highly to the input features. Compared to a Random Forest, ET relies on a random split, which reduces variance more (Geurts et al. 2006). The ET regression model is implemented with the Python package `scikit-learn` (Pedregosa et al. 2011). However, an ET can only incorporate very different geographical variances by coordinates

**Table 2** Number of datapoints and posting dates of included tweets for the different temporal splits

Split	Train/Val		Test	
	N°DP	Dates	N°DP	Dates
TS	96	Sep 1, 2021	188	Sep 2, 2021
TN	527	Aug 29 - Sep 2, 2021	36	Sep 2, 2021
TN-all	1286	Aug 29 - Sep 2, 2021 + Sep 10 -Sep 13, 2017	36	Sep 2, 2021
TN-2	1146	Aug 29 - Sep 2, 2021 + Sep 10 -Sep 13, 2017	83	Sep 2, 2021

Val: Validation, N°DP: number of datapoints, TS: temporal-stationary, TN: temporal-nonstationary

given as input features and might not be able to model the local relationships between the coordinates. Therefore, we select GWR as a different approach (Brunsdon et al. 1996). GWR considers nonstationary input features by incorporating features within each target value's neighborhood. It can, therefore, link the local relationships between the features and the label. In geographical applications, GWR has been used for modeling the existence of potentially complex spatial relationships in a relatively simple and effective way (Páez et al. 2011; Wu et al. 2021). The model is applied as implemented in ArcGIS (ESRI Inc. 2020).

For the ET model, we obtain the hyperparameters by a grid search. We demonstrate feature importance exemplary for one ET application. The significance of a feature is determined by calculating the (normalized) overall improvement in the criterion due to that feature. A concept also referred to as Gini importance. For the GWR approach, hyperparameter tuning is performed manually through iterative experimentation using ArcGIS. The test dataset is not used for the training procedure in both model approaches. Table 5 in Sect. A summarizes the respective hyperparameter settings. Note that for the temporally varying splits, we choose different hyperparameters.

For the evaluation of the models' regression performances and the comparison of the different results, we rely on several metrics. The coefficient of determination  $R^2$ , the root mean squared error (RMSE), and the mean absolute error (MAE) are metrics that are usually applied in regression problems. Furthermore, we employ the maximum error (ME). Table 6 in Sect. B summarizes the respective evaluation metrics.

## 2.5 Postprocessing

For the regression task, the distance from the tweet point to the hurricane track was used as a substitute value to enable the deployment as a data label. Since only the distance is estimated from the approach, but no directional information is given, trilateration is necessary to obtain the hurricane-impacted zone from the estimated data. We conduct the following steps to obtain the impact zone:

1. Buffer the tweet points with their respective estimated distance.
2. Calculate the intersections of the buffer circles.
3. Calculate the kernel density of the intersection points. This step is based on the assumption that where more of the buffers intersect, the higher the chance that this area is actually a hazard impact zone.

## 3 Results

We present the regression results in the following sections. First, Sect. 3.1 presents the overall estimation performance of the applied models on different splits for hazard estimation without consideration of temporal aspects. Finally, Sect. 3.2 shows the results of evaluating the regression models on the datasets considering temporal aspects. For the visualization of the respective results, we focus on the selected study area subsets (as demonstrated in Sect. 2.1) to allow the recognition of details for better understanding. The total area of the hurricane passage could not be visualized due to its size.

### 3.1 Spatial hazard estimation results

The approaches on the selected splits trained without a temporal aspect (see Sect. 2.3) generally achieve high scores. Table 3 shows the estimation results of the applied models on the respective test sets. ET model on the baseline split and the GB split show the best regression results with an  $R^2$ -score of >93% and even >99% for the GB split. ET model’s results on the DB split follow with  $R^2 = 86\%$ . The GWR produces minor accurate results with an  $R^2 = 82\%$ . When considering RSME and MAE, the ET model on the GB split achieves outstanding regression results between 1.2 km and 1.7 km. The models’ values on other splits and other models’ values range from 6.4 km to 10.5 km for the RMSE and 4 km to 5.5 km for the MAE, with a slightly stronger differing value for the GWR model. With a range of 18.6 km to 130.5 km, the ME is very different for all approaches. Particularly the ET model on the GB split outperforms the model on the other splits with 18.6 km, while GWR achieves only 130.5 km. The ET, on the other splits, achieves median values of 30.2 km to 69.6 km.

Considering the feature importance of the attributes when trained with the ET model and a baseline split (Fig. 3), the tweet locations (x and y coordinates) are the main important features. The DDI also has significant importance, closely followed by altitude. On the contrary, we do not detect a significant importance of population, slope, aspect, and distance to roads as features for the hazard estimation.

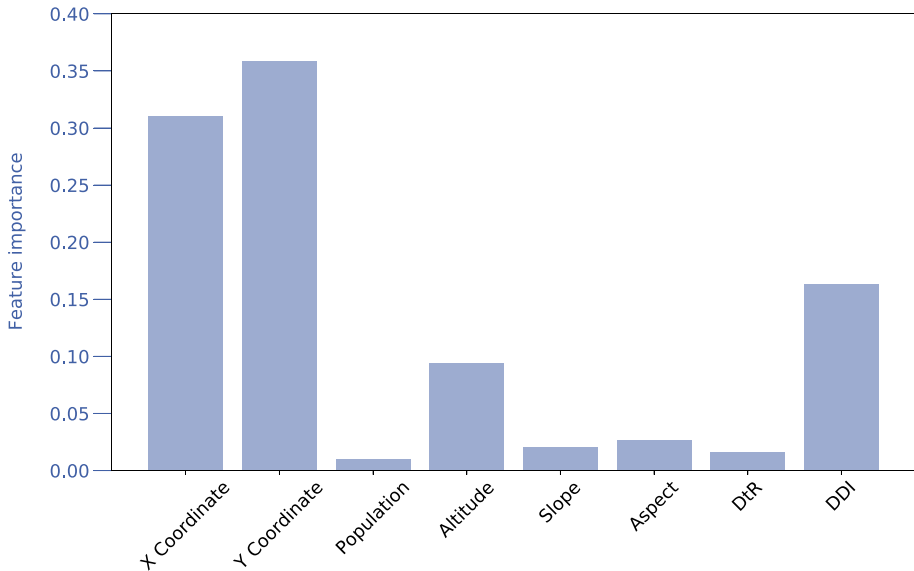
We visualize our approach’s geographical accuracy in Fig. 4 for the baseline split with the GWR model. The displayed subset consists of a regional area in Florida, the USA. Very few estimated test datapoints show deviations from their expected distance compared to their surrounding points of about 15 km. Most deviating test datapoints differ from their expected distance compared to their surrounding datapoints by only about 5 km. In the eastern part of the subset, a significant part of the test datapoints is estimated to be of a distance of  $\leq 30$  km from the track, while they should be  $\leq 50$  km from the track according to the training datapoints. When postprocessing the estimated distance from the tweet point to track to derive the actual impact area, we obtain distinct areas characterized by varying degrees of likelihood for the impact area’s presence. Overall, the estimated impact area corresponds very well with the hurricane track reference, especially for the central area of the subset. Toward the edges of the investigated subset, the impact area is not estimated with high likelihood.

**Table 3** Regression metrics of all models estimated on the test dataset and compared to the reference data in %

Model	Split	Reference	$R^2$ in %	RMSE in km	MAE in km	ME in km
ET	Baseline	Sect. 2.3	93.2	6.5	4.0	57.4
DDI	Baseline	Sect. 2.2	90.2	7.8	5.2	69.6
<b>ET</b>	<b>GB</b>	Sect. 2.3	<b>99.5</b>	<b>1.8</b>	<b>1.2</b>	<b>18.7</b>
ET	DB	Sect. 2.3	86.4	7.1	4.8	30.2
GWR	Baseline	Sect. 2.4	82.4	10.5	5.9	130.5

The bold values represent the best results, respectively.

$R^2$ : coefficient of determination, RMSE: root mean square error, MAE: mean absolute error, ME: maximum error, ET: Extremely Randomized Tree regressor, DDI: Digital Divide Index, GWR: Geographically Weighted Regression, GB: geographically balanced, DB: distance balanced



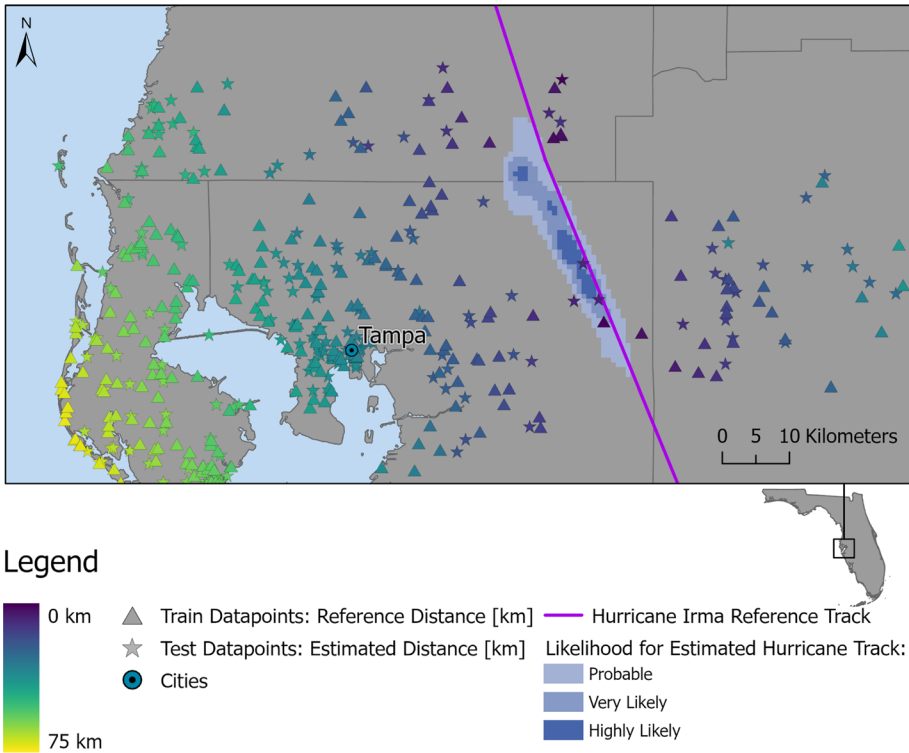
**Fig. 3** Exemplary feature importance for extremely randomized tree (ET). (DtR: Distance to Roads, DDI: Digital Development Index)

### 3.2 Temporal hazard estimation results

The approaches on the selected splits trained considering the temporal aspect (see Sect. 2.3) achieve medium scores. Table 4 shows the estimation results of the applied models on the respective test sets.

ET model shows the best  $R^2$  results with a value of about 70% on the TS split, while it performs concerning the other metrics better on the TN-all split. When considering RSME, MAE, and ME, ET achieves outstanding results on the TN-all split (RSME: 8.5 km, MAE: 6.4 km, and ME: 22.8 km). The other models' values range from 11 km to 19 km for the RMSE and 8.9 km to 15.6 km for the MAE. With a range of 22.8 km to 80.9 km, the ME is very different for all models. The GWR produces less accurate results on the selected scenarios than the ET model, with an  $R^2$  of 58% on the TS split, followed by an  $R^2 = 33\%$  on the TN split. In general, the models achieve better results on the TS split ( $R^2 = 70\%$  and  $58\%$ ) than on the TN split ( $R^2 = 44\%$  and  $R^2 = 33\%$ ). Furthermore, the model performs better on the TN-all split with an  $R^2 = 67\%$  than on the TN-2 split ( $R^2 = 42\%$ ).

We apply the ET model exemplary for TN split to visualize our model's geographical accuracy in Fig. 5, analogous to Fig. 4. The displayed subset consists of a regional area in N.Y., USA. The TN split can be easily distinguished as train datapoints lie in the western subset area, while test datapoints are located in the eastern area of the subset. For the test datapoints the hurricane track is not known at the timestep of estimation. Most of the estimated test datapoints' distances corresponds well with the closest actual train datapoint's distances. We see no major miscalculations of test points except for one test point closest to the hurricane track. It shows a deviation from the expected distance compared to its surrounding points of about 20 km. Additionally, in this visualization, we explicitly display the estimated distances from the test datapoints to the yet to be estimated hurricane track. This presentation aims to enhance the



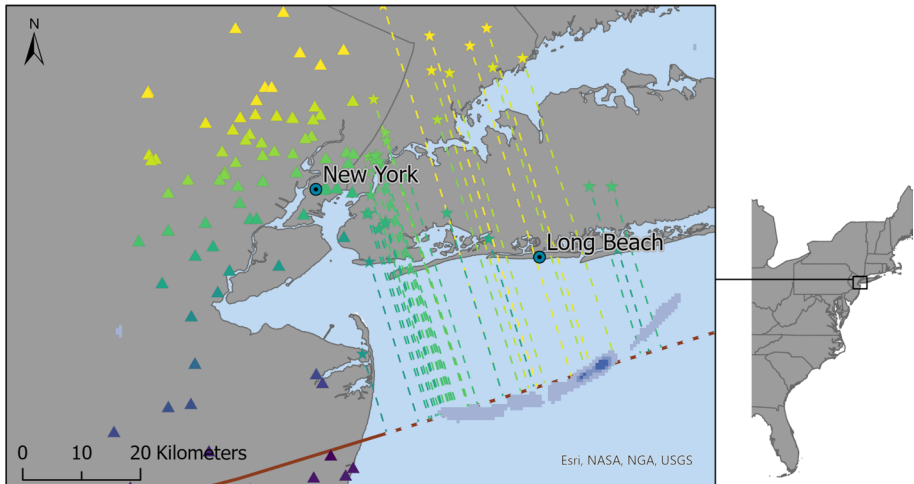
**Fig. 4** Visualization of the estimated distances of test datapoints to the hurricane Irma track of the baseline split based on the Geographically Weighted Regression (GWR). The estimated distances are displayed by color-coding of respective points. The estimated impact area obtained from the postprocessing of the estimated distances is displayed with its respective likelihood values. Data basis: © 2018 GADM. Projection: WSG84

**Table 4** Regression of all models estimated on the test dataset and compared to the reference data in %

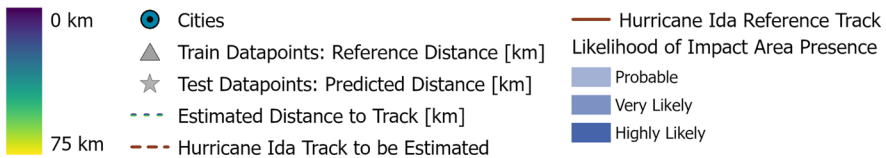
Model	Split	Reference	$R^2$ in %	RMSE in km	MAE in km	ME in km
<b>ET</b>	<b>TS</b>	Sect. 2.3	<b>69.9</b>	14.9	11.5	44.3
GWR	TS	Sect. 2.4	58.3	17.5	11.6	80.9
ET	TN	Sect. 2.3	43.9	11.0	8.9	29.3
GWR	TN	Sect. 2.4	32.9	12.0	9.6	30.4
<b>ET</b>	<b>TN-all</b>	Sect. 2.3	66.8	<b>8.5</b>	<b>6.4</b>	<b>22.8</b>
ET	TN-2	Sect. 2.3	42.2	19.0	15.6	42.0

The bold values represent the best results, respectively

$R^2$ : coefficient of determination, RMSE: root mean square error, MAE: mean absolute error, ME: maximum error, ET: Extremely Randomized Tree regressor, GWR: Geographically Weighted Regression, TS: temporal-stationary, TN: temporal-nonstationary



### Legend



**Fig. 5** Visualization of the estimated distance of test datapoints to hurricane track of the temporal-nonstationary (TN) split based on the Extremely Randomized Tree (ET) for hurricane Ida including reference distance of train datapoints. The estimated distances are displayed by color-coding of respective points. Furthermore, an explicit representation of estimated distances between test datapoints and track is given. The estimated impact area from the postprocessing of the estimated distances is displayed with its respective likelihood values. Data basis: © 2018 GADM. Projection: WSG84

comprehension of the factual meaning of the estimated values. When postprocessing the estimated distance from the tweet point to track to derive the actual impact area, we obtain distinct areas characterized by varying degrees of likelihood for the impact area's presence. The test datapoints' estimated distance gives a good representation of the probable hurricane course and the probable, more heavily impacted areas closer to the hurricane track.

## 4 Discussion

This section discusses the presented study and the achieved estimation results. First, we evaluate the dataset we created for the estimation task of the distance from tweet points to a natural hazard (see Sects. 2.1 and 2.2). Then, we discuss the results of different models and splits as stated in Sect. 2.3, according to the different results in the same order as shown in Sect. 3 (see Sects. 3.1 and 3.2).

## 4.1 Dataset evaluation

The ML approaches require a sufficient amount of data (see Sect. 2.1) to properly train on the one hand. On the other hand, the ML models require appropriate reference data to solve the task of estimating the distance from tweet points to the natural hazard area. This section discusses the hurricane data employed as a reference, Twitter, and supplementary data.

### 4.1.1 Hurricane data evaluation

Since reference data are lacking but are needed to solve regression tasks, we need to generate reference data. Reference data are generated in the form of geodesic distances from tweet points to the nearest natural hazard (hurricane track) point. We rely on an NHC-provided track for generating reference data. This track data is the most appropriate and accurate to represent the proper hurricane track (see Sect. 2.1). The accuracy is given at 1.8 km (Landsea and Franklin 2013). However, these tracks are only provided after a hurricane, sometimes not until one year after the hurricane (Knapp et al. 2010). Therefore, other reference data would be needed for a near real-time application approach. Such other data should be available in real-time and should include:

- Tracks from previous hurricanes, which are already available from archives. These would be employed similarly to the track data of hurricane Irma in the hurricane Ida analysis (compare Sect. 3.2).
- Actual impact zones of the hurricane under investigation from earlier timesteps. These should be impact zones for location 1 that the hurricane already passed at timestep 1 to allow the estimating of the hurricane at location 2 at timestep 2. These could, e.g., be delineated from remote sensing data that captured the impact zone of the hurricane shortly after it passed location 1 or ground observations from individuals, emergency responders, or agencies or numerical weather models that can retrospectively simulate past weather events from measured weather data.

These impact zones provide rough visual estimates that approximate the hurricane track, which could be used for further methodology in real-time applications.

In general, any regression results provided by the models can only be as accurate as the reference data themselves.

An additional challenging aspect for the temporal approach of spatial analyzes of natural hazards from VGI data for the demonstrated proof of concept is the accurate choice of tweets created at the time of the hurricane passing in a specific area. Temporal information on which date these tweet locations were affected by the hurricane is extracted from the TCPAs (see Sect. 2.1.1), which are only approximate. However, in a real-time application, this difficulty would not arise, as we would rely on all available tweets prior to the time of the investigation.

### 4.1.2 Twitter data evaluation

We rely on the available Twitter data with provided coordinates. These data have a high location accuracy of several meters (Dittrich 2016). Our dataset primarily consists of tweets related to Hurricane Irma in 2017, constituting about 56% of our dataset, while

tweets with coordinates for Hurricane Ida in 2021 make up only 44% (compare Fig. 2). This trend reflects the general decline in geotagged tweets in the past years (Kruspe et al. 2021). To enhance datasets, considering tweets with filled `place` fields in addition to `coordinates` fields could be valuable. Data gaps are evident in areas with challenging terrain or lower population density, limiting the use of VGI data in these regions. The discussed transformation from geographical to UTM coordinates, only partially addresses the challenge of diverse geographical coordinates across our study area.

### 4.1.3 Supplementary data evaluation

In addition to the tweets and the hurricane track data, we employ supplementary data (see Sect. 2.1.3). Concerning the various input features, the VGI location data combined with the supplementary data of population, altitude, slope, aspect, and distance to roads represent the basis for the regression of the distance from VGI points to natural hazards. The DDI feature was tested supplementary in one approach. This setup was chosen based on the significance of the DDI feature (Fig. 3), which exhibited the highest feature importance, apart from the `x` and `y` coordinates. On the contrary, we do not detect a major importance of population, slope, aspect, and distance to roads as features for the hazard estimation compared to other studies, e.g., (Jiang et al. 2019). This could be attributed to the fact that the extraction of all these features is conducted for the single tweet points. Selecting these data values and comparing them to tweet occurrence in other spatial formats (e.g., census-tract-wide or zip-code area-wide) could result in other outcomes. Implicitly, the features influencing tweet occurrence do influence the natural hazard estimation. Therefore, we conclude that the applied input features are sufficient to solve the regression task.

## 4.2 Spatial hazard estimation discussion

Our results (see Fig. 4 in Sect. 3.1), indicate that it is indeed possible to estimate the distance from VGI data to a hurricane track with a high degree of accuracy, as demonstrated in Table 3. We conclude that the underlying regression task is feasible based on the provided input features extracted from the tweet locations and the supplementary data. The findings from different investigated setups (Table 3) can be summarized as follows:

- The DDI data do not significantly improve the regression accuracy. While it may possess relatively high feature importance, it does not inherently require consideration as an additional input feature.
- The GB approach achieves the best results in all metrics, as spatially well-distributed datapoints are considered as training input.
- The ET is the best-performing regression model. It can estimate the distance from tweet points to hazard well. The GWR generally has worse generalizing capabilities in our regression task than the ET model.

Given the exceptional performance of the ET model, particularly with the baseline or GB split, our results suggest its viability for practical applications. Our employed postprocessing of the estimated distance from tweet points to the hurricane track is suitable for the delineation of distinct regions characterized by different probabilities of the impact area's presence. Comparing the estimated impact areas with the hurricane track confirms the high accuracy of the estimation approach, particularly in the central region of the subset.



However, as moving toward the periphery of the investigated subset, the likelihood of accurately estimating the impact area decreases due to higher test point density distributions in this area. This leads to the characteristic runout toward the edges of the investigated subset. Combining such partially estimated impact areas from various high tweet density hubs could be used subsequently to interpolate from the total hurricane track length impact area.

### 4.3 Temporal hazard estimation discussion

In our second objective, we explored the feasibility of estimating the distance of VGI points to an unknown track in a second time step when trained on a known track in the first time step. The four investigated splits (Table 4) can be summarized as follows:

- The models achieve better results on the TS split than on the TN split, as the geographical variations between the training and test set are minor.  
However, the TN RMSE, MAE, and ME are slightly better since the TN test set is much smaller. The higher the number of test points, the more it crystallizes when the model does not correctly estimate some points.
- The results on the TN-all split reveal that the increase in training datapoints allows the model to achieve much better accuracies.
- The TN-2 split in Table 4 leads to lower results than the TN-all split, as it has a slightly lower number of datapoints in the training set but a higher number of datapoints in the test set.

Overall, the accuracy of estimating the distance from tweet points to the natural hazard area depends on the selected temporal split. Despite a comparatively high approach sensitivity to geographical changes as investigated with the different splits, estimating approximate affected areas regionally next to the geographical training area provides satisfactory results. In the temporal distinct estimation case, our employed postprocessing of the estimated distance from tweet points to the hurricane track is suitable for the delineation of the impact area's presence. In the investigated subset, the circular shape of the estimated area can be explained by the distribution of the test datapoints to the track. These are all located on the track's northern side, leading to a less suitable trilateration arrangement. Overall, with errors of approximately 9 km, our track estimation results from tweet datapoints compare favorably with weather data track predictions (Regnier 2008; Cangialosi 2017). The results demonstrate that the natural hazard impact area estimation from during-hazard tweet point locations trained on past hurricane data is feasible.

## 5 Conclusions

This study proposes and presents an approach for natural hazard-impacted area estimation based on ML approaches and VGI data in the use-case of hurricane track impact area estimation. The results of the two investigated research questions summarize as follows:

- *Spatial Hazard Estimation (Extent Estimation)*: All selected regression approaches with distinct splits achieve satisfying regression results for the general extent estimation with, for example, an  $R^2$ -value of  $> 82\%$ .  
The ET model is the best-performing model with an  $R^2$ -value of  $> 99\%$  on the GB split.

- *Temporal Hazard Estimation (Development Estimation)*: For the temporal investigation of the natural hazard estimation, the selected ML models combined with the different splits achieve medium regression results with, for example, an  $R^2$ -value of 33% to 70%. An appropriate choice of splits for geographically suitable information content immensely influence the models' performance.

Our presented approach can be considered an initial approach toward natural hazard and hazard development estimation from VGI data. This approach demonstrates high accuracy compared to traditional weather data forecasts (e.g., Cangialosi 2017; Bilskie et al. 2022). It is applicable to various types of natural hazards and can provide timely information, making it valuable for disaster management and first responders.

Moreover, this methodology is versatile and can be adapted for other forms of VGI data, which can be accessed through various platforms, including Instagram, Facebook, or even police reports (under the conditions that they contain posting/hazard location, date, and time). Thus, potential challenges related to specific social media platforms or API access can be mitigated by sourcing data from alternative platforms.

In the future, the transferability of this approach for impact zone estimation should be tested on newly emerging hurricanes. Much more training data from many past hurricanes would be required to do so to account for more possible spatial and sociodemographic constellations. This would enhance the dispensability of information about the hurricane's previously passed locations. Furthermore, VGI data could be combined with the available remote sensing data to enhance our proposed approach further (e.g., Cervone et al. 2016; Bruneau et al. 2021). Additionally, one significant contribution of VGI and an advantage compared to remote sensing data is the text messages from which information about the hazard event could be extracted (e.g., Wang et al. 2016; Kumar et al. 2014). A spatial analysis, including text information, could be conducted to map the approximate extent of the hazard and the extent of differently damaged zones. In further studies, other methodological approaches like Recurrent Neural Networks (RNNs) that are designed for sequential data, could also be investigated. These would also have the ability to capture information from previous time steps.

## Appendix A: Hyperparameters

**Table 5** Hyperparameter setup for the ML models

Model	Split	Package	Hyperparameter setup
ET (Geurts et al. 2006)	Baseline, GB, DB, TN-all, TN-2	Scikit-learn	<code>n_estimators = 100;max_depth = 12;min_samples_leaf = 1;min_samples_split = 2</code>
ET (Geurts et al. 2006)	TS, TN	Scikit-learn	<code>n_estimators = 5;max_depth = 5;min_samples_leaf = 1;min_samples_split = 6</code>
GWR (Brunsdon et al. 1996)	Baseline, TS, TN	ArcGIS	<code>number_of_neighbors = 30</code>

The models are implemented in `scikit-learn` (Pedregosa et al. 2011) and ArcGIS

ET: Extremely Randomized Tree, GB: geographically balanced, DB: distance balanced, TN: temporally nonstationary, TS: temporally stationary

## Appendix B: Evaluation metrics

**Table 6** Evaluation metrics for regression approaches

Metric	Formula	Description
$R^2$	$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$	Represents the proportion of variance that has been explained by the independent variables in the model
RMSE	$RMSE = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{n}}$	The square root of the mean squared error (MSE) measures in the same units as the response variable, in contrast to the MSE
MAE	$MAE = \frac{\sum y_i - \hat{y}_i }{n}$	The average of the absolute differences between the predicted values and the actual values in a dataset
ME	$ME = \text{Max}( y_i - \hat{y}_i )$	It is the worst case error between the predicted values and the actual values in a dataset

$i$  represents a respective datapoint,  $y_i$  is the true label and  $\hat{y}_i$  is the estimated label of the  $i$ -th datapoint.  $\bar{y}$  is the mean of the observed data and  $n$  is the number of estimated datapoints.

$R^2$ : coefficient of determination, RMSE: root mean squared error, MAE: mean absolute error, ME: maximum error

**Acknowledgements** We thank Stefan Hinz, head of the Institute of Photogrammetry and Remote Sensing at the Karlsruhe Institute of Technology, for the partial funding of this work. We acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

**Author Contributions** All authors prepared the methodological concept of this study, the original draft, as well as the editing of the manuscript. Janine Florath designed the software, curated the data, performed the investigation, formal analysis, and validation. All authors contributed to the visualization of the data and results. Sina Keller and Jocelyn Chanussot initialized the research and provided didactic and methodological inputs. All authors have read and agreed to the published version of the manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This work has been partially supported by MIAI@Grenoble Alpes (ANR-19-P3IA-0003) and the AXA Research Fund.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- André G (2012) Natural hazard mapping across the world. A comparative study between a social approach and an economic approach to vulnerability
- Arthur R, Williams HT (2019) Scaling laws in geo-located Twitter data. *PloS one* 14(7):e0218454
- Benson E, Haghighi A, Barzilay R (2011) Event discovery in social media feeds. In: Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies, pp. 389–398
- Bilskie MV, Asher TG, Miller PW, Fleming JG, Hagen SC, Luettich RA Jr (2022) Real-time simulated storm surge predictions during Hurricane Michael (2018). *Weather Forecast* 37(7):1085–1102
- Birchfield GE (1960) Numerical prediction of hurricane movement with the use of a fine grid. *J Atmosph Sci* 17(4):406–414
- Bruneau P, Brangbour E, Marchand-Maillet S, Hostache R, Chini M, Pelich RM, Matgen P, Tamisier T (2021) Measuring the impact of natural hazards with citizen science: the case of flooded area estimation using twitter. *Remote Sens* 13(6):1153
- Brunsdon C, Fotheringham SA, Charlton ME (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. *Geograph Anal* 28(4):281–298
- Cangialosi JP (2017) National Hurricane Center Forecast report: 2017 hurricane season
- Cervone G, Sava E, Huang Q, Schnebele E, Harrison J, Waters N (2016) Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 boulder flood case study. *Int J Remote Sens* 37(1):100–124
- Cox TS, Hoi CS, Leung CK, Marofke CR (2018) An accurate model for hurricane trajectory prediction. In: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), vol 2, pp. 534–539. IEEE
- De Albuquerque JP, Herfort B, Brenning A, Zipf A (2015) A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *Int J Geograph Inform Sci* 29(4):667–689
- Dittrich A (2016) Real-time event analysis and spatial information extraction from text using social media data. PHD Thesis, Karlsruhe Institute of Technology, Karlsruhe

- Dittrich A, Lucas C (2014) Is this Twitter event a disaster? In: AGILE'2014 international conference on geographic information science, connecting a digital europe through location and place, Castellón. AGILE Digital Editions
- ECMWF European Centre For Medium-Range Weather Forecasts. (2021) IFS Documentation
- ESRI Inc. (2020) ArcGIS Pro 2.6.1. Redlands, CA: Esri Inc. Software
- Florath J, Keller S (2022) Supervised machine learning approaches on multispectral remote sensing data for a combined detection of fire and burned area. *Remote Sens* 14(3):657
- Fohringer J, Dransch D, Kreibich H, Schröter K (2015) Social media as an information source for rapid flood inundation mapping. *Natl Hazards Earth Syst Sci* 15(12):2725–2738
- Forati AM, Ghose R (2022) Examining community vulnerabilities through multi-scale geospatial analysis of social media activity during hurricane Irma. *Int J Disast Risk Red* 68:102701
- Gallardo R (2020) Digital Divide Index
- Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. *Mach Learn* 63(1):3–42
- Goodchild MF (2007) Citizens as sensors: the world of volunteered geography. *Geo J* 69(4):211–221
- Guan X, Chen C (2014) Using social media data to understand and assess disasters. *Natural Hazards* 74(2):837–850
- Hao H, Wang Y (2020) Leveraging multimodal social media data for rapid disaster damage assessment. *Int J Disast Risk Reduct* 51:101760
- Harrald JR (2006) Agility and discipline: critical success factors for disaster response. *Ann Am Acad Polit Soc Sci* 604(1):256–272
- Hijmans R (2022) DIVA-GIS
- Imran M, Mitra P, Castillo C (2016) Twitter as a lifeline: Human-annotated twitter corpora for NLP of crisis-related messages. arXiv preprint [arXiv:1605.05894](https://arxiv.org/abs/1605.05894)
- Jiang Y, Li Z, Ye X (2019) Understanding demographic and socioeconomic biases of geotagged twitter users at the county level. *Cartogr Geograph Inform Sci* 46(3):228–242
- Kattenborn T, Leitloff J, Schiefer F, Hinz S (2021) Review on convolutional neural networks (CNN) in vegetation remote sensing. *ISPRS J Photogr Remote Sens* 173:24–49
- Knapp KR, Diamond HJ, Kossin JP, Kruk MC, Schreck CJ (2018) International best track archive for climate stewardship (IBTrACS) project, version 4
- Knapp KR, Kruk MC, Levinson DH, Diamond HJ, Neumann CJ (2010) The international best track archive for climate stewardship (IBTrACS) unifying tropical cyclone data. *Bull Am Meteorol Soc* 91(3):363–376
- Kruspe A, Häberle M, Hoffmann EJ, Rode-Hasinger S, Abdulahhad K, Zhu X (2021) Changes in Twitter geolocations: Insights and suggestions for future usage. arXiv preprint [arXiv:2108.12251](https://arxiv.org/abs/2108.12251)
- Kumar A, Jiang M, Fang Y (2014) Where not to go? Detecting road hazards using Twitter. In: Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval, pp. 1223–1226
- Landsea CW, Franklin JL (2013) Atlantic hurricane database uncertainty and presentation of a new database format. *Monthly Weather Rev* 141(10):3576–3592
- Lindell MK, Prater CS (2003) Assessing community impacts of natural disasters. *Natl Hazards Rev* 4(4):176–185
- NHC National Hurricane Center. (2022a) Definition of the NHC Track Forecast Cone
- NHC National Hurricane Center. (2022b) NHC Data in GIS Formats
- NHC National Hurricane Center. (2022c) Tropical cyclone advisory archive
- NOAA-NCEI National Centers for Environmental Information. (2022) U.S. Billion-Dollar Weather and Climate Disasters
- Páez A, Farber S, Wheeler D (2011) A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. *Environ Plann A* 43(12):2992–3010
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine Learning in python. *J Mach Learn Res* 12:2825–2830
- Regnier E (2008) Public evacuation decisions and hurricane track uncertainty. *Manag Sci* 54(1):16–28
- Schultz MG, Betancourt C, Gong B, Kleinert F, Langguth M, Leufen LH, Mozaffari A, Stadler S (2021) Can deep learning beat numerical weather prediction? *Philosoph Trans Royal Soc A* 379(2194):20200097
- Seaberg D, Devine L, Zhuang J (2017) A review of game theory applications in natural disaster management research. *Natural Hazards* 89(3):1461–1483
- Songwathana K (2018) The relationship between natural disaster and economic development: a panel data analysis. *Proc Eng* 212:1068–1074

- Walker ND, Haag A, Balasubramanian S, Leben R, Van Heerden I, Kemp P, Mashriqui H (2006) Hurricane prediction. *Oceanography* 19(2):24
- Wang Y, Wang T, Ye X, Zhu J, Lee J (2016) Using social media for emergency response and urban sustainability: a case study of the 2012 Beijing rainstorm. *Sustainability* 8(1):25
- Wang Z, Ye X, Tsou MH (2016) Spatial, temporal, and content analysis of Twitter for wildfire hazards. *Natural Hazards* 83(1):523–540
- Ward PJ, Blauhut V, Bloemendaal N, Daniell JE, de Ruiter MC, Duncan MJ, Emberson R, Jenkins SF, Kirschbaum D, Kunz M, Mohr S, Muis S, Riddell GA, Schäfer A, Stanley T, Veldkamp TI, Winsemius HC (2020) Natural hazard risk assessments at the global scale. *Natural Hazards Earth Syst Sci* 20(4):1069–1096
- Wisner B, Blaikie P, Cannon T, Davis I (2014) *At risk: natural hazards, people's vulnerability and disasters*. Routledge, Taylor
- worldpop.org. (2022) Worldpop
- Wu J, He J, Christakos G (2021) *Quantitative analysis and modeling of earth and environmental data: space-time and spacetime data considerations*. Elsevier
- X Corp. (2023) Twitter API
- Zhang F, Minamide M, Nystrom RG, Chen X, Lin SJ, Harris LM (2019) Improving harvey forecasts with next-generation weather satellites: advanced hurricane analysis and prediction with assimilation of GOES-R all-sky radiances. *Bull Am Meteorol Soc* 100(7):1217–1222

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.