International Journal of Geo-Information

Article

Road Accessibility during Natural Hazards Based on Volunteered Geographic Information Data and Network Analysis

Janine Florath [1,2,*], Jocelyn Chanussot [3] and Sina Keller [1]

1 Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany; sina.keller@kit.edu
2 GIPSA-Lab, Université Grenoble Alpes, CNRS, Grenoble INP, 38000 Grenoble, France
3 Inria Center, Université Grenoble Alpes, CNRS, Grenoble INP, LJK, 38000 Grenoble, France; jocelyn.chanussot@inria.fr
* Correspondence: janine.florath@kit.edu

Abstract: Natural hazards can present a significant risk to road infrastructure. This infrastructure is a fundamental component of the transportation infrastructure, with significant importance. During emergencies, society heavily relies on the functionality of the road infrastructure to facilitate evacuation and access to emergency facilities. This study introduces a versatile, multi-scale framework designed to analyze accessibility within road networks during natural hazard scenarios. The first module of the framework focuses on assessing the influence of natural hazards on road infrastructure to identify damaged or blocked road segments and intersections. It relies on near real-time information, often provided by citizen science through Volunteered Geographic Information (VGI) data and Natural Language Processing (NLP) of VGI texts. The second module conducts network analysis based on freely available Open Street Map (OSM) data, differentiating between intact and degraded road networks. Four accessibility measures are employed: betweenness centrality, closeness centrality, a free-flow assumption index, and a novel alternative routing assumption measure considering congestion scenarios. The study showcases its framework through an exemplary application in California, the United States, considering different hazard scenarios, where degraded roads and connected roads impacted by the hazard can be identified. The road extraction methodology allows the extraction of 75% to 100% of the impacted roads mentioned in VGI text messages for the respective case studies. In addition to the directly extracted impacted roads, constructing the degraded network also involves finding road segments that overlap with hazard impact zones, as these are at risk of being impacted. Conducting the network analysis with the four different measures on the intact and degraded network, changes in network accessibility due to the impacts of hazards can be identified. The results show that using each measure is justified, as each measure could demonstrate the accessibility change. However, their combination and comparison provide valuable insights. In conclusion, this study successfully addresses the challenges of developing a generic, complete framework from impact extraction to network analysis independently of the scale and characteristics of road network types.

Keywords: road network accessibility; volunteered geographic information; natural language processing; natural hazards; wildfires; floods; disaster risk management; evacuation; critical road infrastructure; geographic information

check for updates

Citation: Florath, J.; Chanussot, J.; Keller, S. Road Accessibility during Natural Hazards Based on Volunteered Geographic Information Data and Network Analysis *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 107. https://doi.org/10.3390/ijgi13040107

Academic Editors: Wei Huang and Wolfgang Kainz

Received: 12 January 2024
Revised: 8 March 2024
Accepted: 17 March 2024
Published: 22 March 2024

1. Introduction

In recent times, there has been a noticeable increase in the frequency and severity of natural hazards [1]. For example, from 1962 to 2011, 34% of investigated stations in the United States (US) reflect an increasing frequency and 13% an increasing magnitude of flood events [2]. Wildfire frequency is four times as high and total burned area six times as high in the period between 1986 and 2005 compared to the period of 1970 to 1986 in the US [3]. These events pose a significant threat to both the physical and socio-economic well-being of individuals and natural ecosystems [4,5]. Critical infrastructures, consisting of a fusion of technical and organizational elements essential for sustaining societal operations, are also susceptible to the impact of these natural hazards [6,7]. The road network, as such critical infrastructure, assumes a crucial role in the transportation system. It frequently suffers damage due to natural hazards (for example, see [8,9]). Therefore, it is critical to investigate the functionality of critical road infrastructure before, during, and after disasters.

Many studies employ detailed road network analysis, considering degraded network scenarios. Few studies also employ road network analysis in a natural hazard context (e.g., [8–11]). However, these studies usually concentrate on a single case study hazard in a single location and have limited applicability to other contexts. As it is especially difficult to access complex degraded road datasets after a natural hazard, proposed advanced models or simulations are often not transferable. Consequently, there is a lack of a generic framework that can tackle the challenge of restricted global data availability and directly analyze the impact of natural hazards on critical road infrastructure.

For such analysis, near real-time information about natural hazard locations or, more specifically, its impact on road infrastructure, needs to be investigated. Such information is usually available from responsible agencies like firefighters or police. Still, recently, more studies have focused on including citizen science [12] in the form of Volunteered Geographic Information (VGI) data. VGI encompasses geospatial data that originates from non-professional sources [13,14]. In the present study, our focus lies on former Twitter, now X, data. Twitter data, commonly called "tweets," frequently contain georeferenced, short messages that serve as a valuable resource for discerning information regarding the impact of hazards on road infrastructure [15]. VGI has been used in various studies for extracting location information from text [16,17]. Extracting place names from text, in general, is conducted using geoparsing. Geoparsing is converting text descriptions of places into geographic identifiers, like coordinates. It consists of two steps:

1. Extracting place names from text; and
2. Geocoding the extracted place texts into geographic identifiers.

The first step (1) is mainly conducted with Named Entity Recognition (NER) or Regular Expression (RegEx) (e.g., [18–20]), which are Natural Language Processing (NLP) methods. New studies focus on improving general place extraction from texts by using various techniques like convolutional neural networks for text analysis [21] and transformer models [22]. Some studies also focus specifically on improving road names' extractions [23]. Secondly (2), only a few studies focus on the improvement of the place names' geocoding. However, it is equally important, as a better place name extraction is only valuable for geographic applications if all the newly extracted place names can also be allocated to a geographic location. Currently, the main geocoding methods employed are Google Place Application Programming Interface (API) [18,19,24], Yahoo Placemaker [25], and ArcGIS geocoding [26], which are not freely available. Recently, more studies have also used open-source geodatabases (e.g., [27,28]). However, not all APIs and databases offer precise geoparsing at road level.

VGI has been used in a few studies in road analysis contexts. It has generally been applied mainly for road location approximation from Twitter posting location [29] or text indicators [30]. Furthermore, VGI has been used for event detection near roads (e.g., traffic event detection) [26,30,31]. These studies, however, do not focus on detailed road intersections or road segment locations. Using VGI text for specific road place extraction, among other traffic event-related named entities, has been described in Ünsal [32]. Vallejos et al. [33] precisely geocode traffic incident locations recognized from NER. Yu and Li [23] and Gelernter and Balaji [34] focus on road extraction, extracting place names from text data, but do not conduct the geoparsing step. In summary, geoparsing has improved in the past years when focusing on road name extraction, but gaps exist in specific geographical road location extraction. VGI data have yet to be studied for intersection or road segment information extraction for geographical applications. As a result, the accuracy requirements for such applications have yet to be adequately addressed.

Road network analysis studies the relationships and interactions between entities (nodes and edges) in a road network, utilizing tools to calculate paths within the network. A path refers to the route between nodes in a network. It is a series of interconnected edges that connect one node to another. The paths within a network can be computed using various performance measures as a basis. For example, the distance corresponds to the calculation of the shortest paths, or the travel time corresponds to the computation of the fastest paths [35,36]. Further performance measures are throughput and capacity [37,38], topological measures [39,40], economic measures [41], and accessibility [42]. One important measure defining road network performance is accessibility [42]. Accessibility is defined as the ease with which a location can be reached. However, there are various measures of how accessibility can be determined. Accessibility indices [36,43] have been applied in a few studies. These metrics mostly rely on connectivity information but require more detailed traffic and travel demand data for more complex network analysis (e.g., [37,44]). Obtaining daily or average travel demand data is often difficult [45,46]. This difficulty is intensified in hazard situations that deviate from average daily traffic scenarios. Even when average travel demand data are available, they may not accurately reflect travel patterns in hazardous situations. As a result, metrics that do not rely on other traffic data are more useful in such cases.

Utilizing accessibility-based metrics enables fast and generic analysis of the hazard-impacted road network. Such metrics measure the connectivity between locations, which is a direct road network performance measure. Many methods rely on graph theory (e.g., [47,48]). Common accessibility methods include the use of various measures of connectivity (e.g., beta index, association number, alpha index, gamma index) and accessibility indices (Shimbel index and nodal degree) [49–51]. For network analysis, the graph theoretical measures betweenness and closeness centrality are often used [52–54]. They provide a fast and overall picture of accessibility in a road network but do not include specific scenarios that could be important to natural hazard planning and emergency response. Antunes et al. [55] and Santos et al. [56] present accessibility-maximization approaches in a critical road infrastructure context. Accessibility-based performance measures are also proposed by Chen et al. [57].

Otherwise, accessibility indices can be adapted to situational requirements, like an immediate emergency response (evacuations or emergency facility access [36]) or general connectivity investigations (remoteness of places [43]). As a basis serves the calculation of shortest paths, according to Dijkstra [58]. However, here, different metrics can be used again, like the distance along graph edges or the travel time along graph edges [36,59], for calculating the shortest or fastest path, respectively.

Applying graph theory for road network analysis, different criteria apart from the simple shortest paths can be valuable in general [60,61] and for accessibility measures [50]. The use of k-shortest paths has been investigated, e.g., in Chondrogiannis et al. [62] and Bader et al. [60]. However, these criteria have not been considered for alternative road network accessibility analysis measures in a hazard context.

Road Network Analysis incorporating text data as a data source for road network degradation analysis has rarely been studied. Bai et al. [63] collected data on geological disaster road blockages, which emergency responders reported to the responsible agency for geohazard accident duration prediction. Keller and Atzl [64] map extreme precipitation impacts on road infrastructure using weather-related traffic reports.

VGI data, in different formats, have been used for various analyses like road network extraction from images [65,66] or road network analysis, where VGI data serve as road network information [67,68]. However, they have rarely been used for the analysis of the effects of natural hazards on road networks. Tzavella et al. [69] use weather-related traffic reports for network analysis of fire brigade accessibility. Schnebele et al. [70,71] use VGI combined with other data sources to classify potential road damages and road assessment.

However, openly available road information data as input data for a generic road network analysis approach has generally not been used yet, let alone in the context of a natural hazard-impacted road network.

To deal with this identified gap, a generic framework, that is complete from the impacted road data analysis during the hazard to the network analysis, is introduced. It is composed of two modules and is applicable to any natural hazard on any scale. Figure 1 summarizes the framework's workflow.

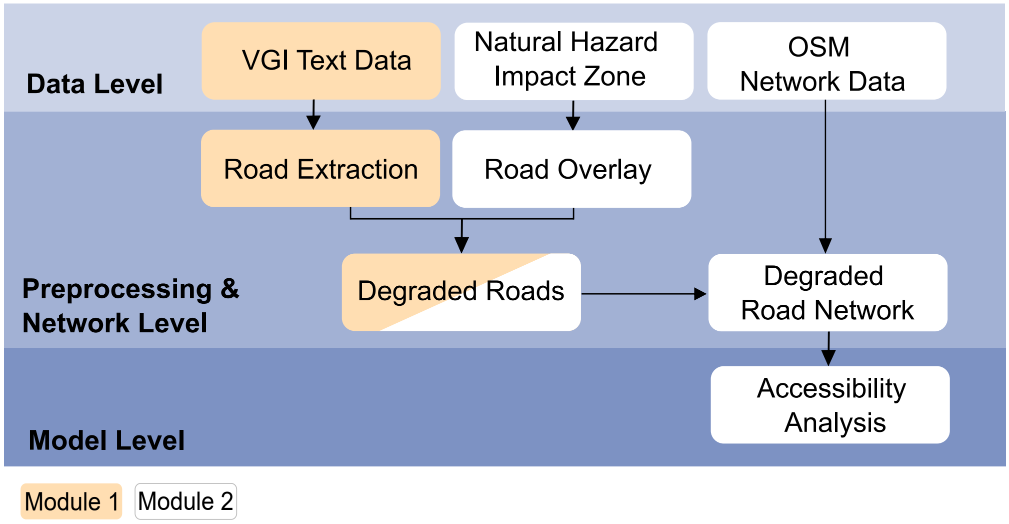

Figure 1. Visualization of the framework for network accessibility analysis during natural hazards. The first module focusing on improving intersection or road segment extraction using Volunteered Geographic Information (VGI) is highlighted in orange. The following abbreviation is used: OSM—Open Street Map.

The framework first evaluates the natural hazard's influence on the road infrastructure to assess where, when, and to what extent the roads are impacted and the network is degraded. To accomplish the first task of natural hazards' impact estimation on road infrastructure, a methodology is developed to extract road-related data such as road intersections and segments obtained from tweet texts that contain information regarding damage or blockages on the roads. Furthermore, we rely on impacted areas' information.

Only after this first step of the analysis of the hazard's impact on the road, can the road network analysis be conducted. Just like the first module, the network analysis relies only on freely and openly available road data in the form of Open Street Map (OSM) data. We conduct a network analysis, taking into account the intact road network versus the degraded network. The degraded network is constructed from the intact network by subducting the road segments and road intersections that were identified as damaged or impassible in Step 1.

We rely on four different methods to analyze the accessibility change due to the hazards' impact. First, we apply betweenness centrality and closeness centrality (e.g., [53,72]). These theoretical accessibility measures, based on graph theory, are employed for the first time in an accessibility change estimation using the comparison of measures on an intact and a degraded road network. Additionally, we employ an accessibility index based on a free-flow assumption, as adapted from Guth [36], for analyzing shelter accessibility within a network. Moreover, a novel measure of accessibility based on an alternative routing assumption will be developed. Contrary to existing studies that focus on emergency facility accessibility (e.g., [36,73,74]), we investigate a use-case scenario focusing on the accessibility analysis of evacuation shelters. The focus of this study on shelter accessibility is based on the understanding that emergency facility accessibility analysis may overlook critical considerations in emergency response, like the occurrence of traffic congestion. Emergency facility accessibility analysis assumes that emergency vehicles have priority and can pass through traffic without significant obstacles (free-flow assumption). However, it is equally crucial to consider the accessibility of places for non-emergency vehicles. During crises, such as natural hazards, people in affected areas may face significant congestion and transportation challenges. These conditions can hinder their ability to reach places directly. By examining shelter accessibility, this study aims to address the occurrence of anticipated congestion scenarios.

Each method is applied to the data from before and during the hazard, and the differences in accessibility are evaluated.

The main objectives of this paper are summarized in the following:

1. Development of a workflow for the extraction of natural hazard-impacted roads from freely and openly available data.
2. Evaluation of considered measures for the task of road network accessibility change analysis based on the extracted degraded road network through the application and adaption of existing network measures.
3. Development of an alternative routing assumption accessibility measure for road network accessibility change analysis in congested scenarios implemented with free-to-use and worldwide available data and open-source software.
4. Development of a generic framework that is complete, from the impacted road data extraction during the hazard to the network analysis using different accessibility measures.

Results are presented for an exemplary application for different hazard scenarios of different geographic scales in California, United States (US).

2. Case Studies and Data

In this study, we rely on specific cases to investigate the impact of natural hazards on road networks. Two different hazards are chosen to demonstrate the applicability of our developed approaches.

Since we aimed to develop a generic approach that could be applied to various natural hazard scenarios, we relied on two case studies subject to various underlying factors that influence their characteristics. The first case study is the flooding in 2022/23 in the San Francisco Bay Area, California (CA), US. From 31 December 2022, heavy precipitation events led to widespread flooding in large parts of CA, persisting until 25 March 2023. The flooding resulted in evacuation orders for 6000 people, 200,000 homes without electricity, and 22 fatalities. The city of Oakland, located within the San Francisco Bay Area, documented unprecedented 24-h rainfall records during this event. We apply the first module, which focuses on the extraction of degraded roads, to cover the entirety of the San Francisco Bay Area affected by the flood. However, due to the large size of this region, the application of the network analysis is restricted exclusively to a subset within this area, specifically the city of Oakland.

The second case study is the Bobcat fire in 2020 in the Los Angeles National Forest, CA, US. The fire ignited on 6 September 2020 and was one of the significant incidents in the 2020 wildfire season. It was only fully contained on 8 December 2020 and burned a total area of 46,900 ha. Figure 2 visualizes the locations of these two natural hazards. These two case studies vary on underlying factors like hazard type, duration, location, and general geographic influences. For example, the flood occurred in an urban area with a complex, dense, and well-developed road network. On the other hand, the Bobcat fire occurred in a forested mountainous region characterized by a sparse and less developed road network in the rugged terrain.

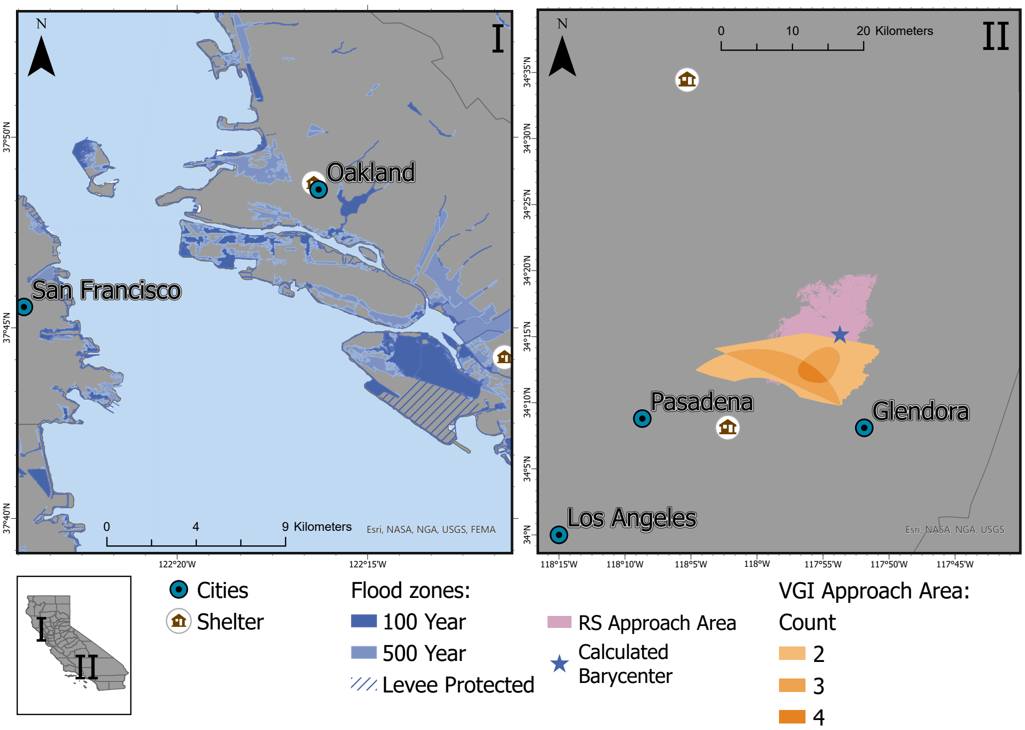

Figure 2. Visualization of the selected case study areas in California, United States, for the selected natural hazards: the flood in the San Francisco Bay Area (I) in 2022/23 (**left**) and the Bobcat fire near Los Angeles (II) in 2020 (**right**). The fire area is displayed as extracted for 6 September 2020 to 10 September 2020. Data basis: © 2018 GADM. Projection: WGS84.

Datasets

We use various geoinformation data in the course of this study.

Information about closed roads or deteriorated road conditions during a natural hazard is usually available from responsible agencies like firefighters or police. Still, recently, more studies have focused on including citizen science [12] in the form of VGI data. This section employs Twitter data as VGI data for impacted road extraction. Twitter is used in recent studies due to its widespread usage and easy accessibility [15,75,76]. The specific Twitter data are acquired using Python programming from Twitter's download API. The API maintains identical functionality after the transition to X within the context of the Twitter developer platform. The tweet data containing relevant keywords about the hazard, for the time when the natural hazard occurs, are extracted for the application scenario. Next, only tweets containing relevant keywords about infrastructure are selected. We extracted the tweet location, date, and text. We exclude:

- Messages of identical content if a specific road segment is closed for longer and the exact message is repeated for update purposes.
- Messages that are road-related but related to a hazard other than the investigated one (e.g., car fires in the investigation region).

For the construction of the degraded network after the hazard, we use the degraded roads extracted from VGI. However, in some geographical regions, only the extraction of a limited number of degraded roads may be feasible. To enhance the assessment of our network analysis approach, we aim to incorporate a more extensive set of deteriorated roads. We intend to utilize datasets that enable inferences regarding road degradation, particularly for roads not mentioned in VGI and, consequently, not extracted with our methodology. To achieve this, we leverage general information about the extent of hazard zones. Roads intersecting with hazard zones will be classified as degraded, too. Therefore, we employ datasets that represent the potential hazard extents.

For the Bay Area floods, this is a 100/500-year flood hazard map provided by the Association of Bay Area Governments for the San Francisco Bay Area. These Federal Emergency Management Agency's flood hazard zones are based on historical data about regional flooding. One-hundred-year floodplains are areas with a 1% (1 in 100) annual chance of flooding that are likely to be flooded at least 0.15 m. Five-hundred-year floodplains are areas with a 0.2% (1 in 500) annual chance of flooding. The map includes the 100- and 500-year floodplains designated by the agency and potential floodplains currently protected by levees. The used map was updated on 1 July 2022.

For the Bobcat wildfire, the considered dataset for the areal information of the hazard is the fire extent as extracted from our methodology proposed in Florath and Keller [77] and Florath et al. [78] for the mapping of natural hazard areas from VGI and remote sensing data. A combined fire and burned area detection approach using a 1D-convolutional neural network (CNN) is applied to estimate the natural hazard impact zone with Sentinel-2 satellite images [77]. Furthermore, the fire hazard area is approximated using information about the wildfire contained in VGI (Twitter) data texts. Two primary methods are used: approximate barycenter calculation and approximate location estimation from tweet datapoints [78]. This approximate location estimation approach uses several methods for fire area estimation from VGI texts and counts the overlaps of single estimates to provide a higher reliability measure for the actual hazard extent [78]. Higher counts refer to a higher certainty of the estimation. Figure 2 displays the respective natural hazard areas.

OSM data are used for road network analysis as they are freely and openly available worldwide, though with varying regional quality and completeness. However, data can be accessed for any location using a single platform instead of navigating through different governmental or proprietary databases in search of authoritative data from various regional authorities separately. The OSM project is widely recognized as the most popular and prominent VGI mapping initiative [79]. Launched in 2004, its primary objective is to generate and offer freely accessible geographic data. For a road network accessibility analysis, road networks of OSM data are employed. OSM street networks are represented as graphs, which are mathematical representations of networks [80]. Generally, a graph, denoted as G, comprises a collection of nodes, represented by the set N, connected by edges, represented by the set E. In a street network, intersections and dead-ends are depicted as nodes, whereas edges represent the connecting street segments. An edge establishes a connection between two nodes or, in the case of self-loops, within a single node. In a directed graph, each edge indicates a specific direction, pointing from one node to another. In a street network, directed edges represent the driving direction. Furthermore, street network graphs allow for parallel edges representing several lanes [47].

Connected graph models, preferably without errors in the data, are necessary for correct road network analysis. Contrary to most authoritative road data, OSM data do not contain road names in full detail. On the other hand, many additional tags are available for each road segment, adding attributes like travel speed or maximum number of lanes to the edges [36]. OSMnx [81] is a Python package that allows users to retrieve geospatial data from OSM. These data are subjected to a cleaning process, followed by creating graph-theoretic models [82]. OSMnx retrieves speed limits from OSM data where available and allows the imputation of missing speed data (e.g., using default speeds based on the functional class of the road). Furthermore, it allows us to calculate travel times for all edges.

Furthermore, we use authoritative road data from government sources of the respective application scenario region. These are the San Francisco Bay Region Roadways by the Metropolitan Transportation Commission, the transportation planning, financing, and coordinating agency for the nine San Francisco Bay Area counties. For the Bobcat wildfire, these authoritative road data are roadways, which extract cartographic information from the U.S. Census Bureau's Master Address File Database (MTDB) for the affected counties. These road data contain very detailed road naming, which is essential for the road extraction task (see Section 3.1). We compare the edges of the OSM network data to the authoritative road data to evaluate the OSM data's accuracy. Between the OSM and the administrative data source, 6% of the road data are different. This means that 6% of the road segments in the datasets do not match or align with each other. When visually comparing OSM road data and Maxar World Imagery, OSM data mostly deviate on private property grounds, e.g., harbor area and open cast mine land. The reason is that private roads are not mapped in OSM data. However, this absence of private road data does not pose a problem for the road network analysis with OSM, as the focus lies on the analysis of public roads. Therefore, OSM data are suitable for road network analysis applications due to their accuracy and completeness of relevant public spaces and infrastructure.

We select reasonable regions of interest for the two case studies and extract the graph's nodes and edges within these regions. The selected region for the flood case study is the city of Oakland. For the Bobcat case study, the selected region is the region within the bounding box of 118.2063881° W 34.5815616° N (top left corner) and 117.6357117° W 34.1330156° N (bottom right corner). This selection covers the Angeles National Forest and the surrounding urban areas (North: Palmdale, South: Arcadia, Azusa, parts of Pasadena). The reason for these selections is the availability of a reasonable number of emergency shelters (two) that can be reached in this region, which is essential for our network accessibility measures. Table 1 displays the number of edges and nodes selected for the respective case studies. The road network of the selected area for the Bobcat wildfire has many nodes and edges compared to the Oakland case study. Some preprocessing steps are necessary to deal with this high number of edges and nodes (see Section 3.2.1).

Additionally, some of our network analysis approaches need specified destination points within the network. In our study case, the destination points are emergency shelters (see Figure 2). Data about the location of emergency shelters are obtained from the respective responsible agency, the city of Oakland administration (oaklandca.gov, accessed on 1 March 2024), for the flood hazard. The reason for selecting the subset of the city of Oakland is the availability of a reasonable number of emergency shelters (two) that can be reached within this region. Table 1 displays the selected shelters' information and the number of edges and nodes set for the two case studies.

Table 1. Data overview for the respective natural hazards. Note that road tweets are collected for the totality of the San Francisco Bay Area, while network and shelter data focus on the subset of the city of Oakland. # is the total amount of collected road tweets, and # of edges and nodes are the respective numbers.

Case Study	# Tweets	# Edges	# Nodes	Shelter Locations
Oakland (Bay Area) flood	32	22.289	8.472	St. Vincent de Paul, Ira Jinkins Community Center
Bobcat fire	24	36.634	14.319	Palmdale High School, Santa Anita Park

3. Methodology

3.1. Degraded Road Extraction

We consider blocked road information. Road authorities often post this information in emergency or hazard cases. We search tweets that mention such information or tweets posted by the responsible agencies. Tweets by road agencies or other agencies posting about road conditions mostly have a similar structure, mentioning a major road affected between certain intersections with that road. Considering this structure, we conduct our methodology. The road extraction can be divided into three tasks:

- Text processing;
- Geoparsing;
- Geographic information system (GIS) processing.

Figure 3 displays the process.

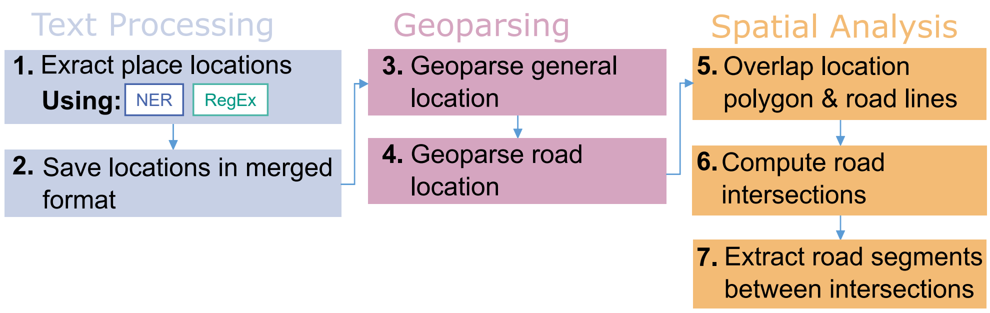

Figure 3. Visualization of Twitter data on the road extraction framework during natural hazards.

The tasks are conducted as follows for each tweet text:

1. From the tweets' texts, we first extract place locations via NER (e.g., *Oakland, CA*). Furthermore, we extract road locations via RegEx implemented for roads (e.g., *Angeles Crest Hwy*).
2. We save the extracted locations of one tweet text in a merged format, saving (1) the general location and (2) all extracted road names separately.
3. We geoparse the general location by searching the location name in the geodatabase and extracting the associated place polygon. This general area is used to search for the road names only in a specific area in the following steps, which is vital as specific road names appear in different neighborhoods or cities.
4. We geoparse the road names by post-processing the road names and searching for their names in the road database. Thus, we extract road lines in a GIS.
5. We overlap the extracted road lines with a general location polygon to ensure we extract the correct road of the specified neighborhood.
6. We then compute the intersections of the extracted road lines if several road names have been mentioned in the text (e.g., *Atherton Ave and US101*) and obtain intersection points.
7. Finally, to find affected road segments as mentioned in the tweets' texts (if several road names have been mentioned), we extract the road segments between the extracted intersection points.

The process is displayed for the exemplary tweet, "At 1:07 PM PST, 1E Novato Dept of Highways reports FLOOD. Flooding on WB37 between Atherton Ave and US101", including schematic displays of the respective steps in Figure 4. The extracted roads represent the degraded roads that cannot be used for routing in the following network analysis Section 3.2.

3.2. Network Analysis

Four methods are investigated and compared for road network accessibility analysis in a hazard scenario. All applied methods are based on graph theoretical considerations. The four applied methods are betweenness and closeness centrality, a free-flow assumption accessibility index, and an alternative routing assumption accessibility measure. The deployed network accessibility measures are presented in Section 3.2.2.

3.2.1. Basics and Preprocessing

To carry out the accessibility analysis, we use a road graph network as the basis (compare Section 2). Few preprocessing steps are necessary to use the network data for such analysis. In the following, we present the preprocessing steps and the general basics of the network analysis.

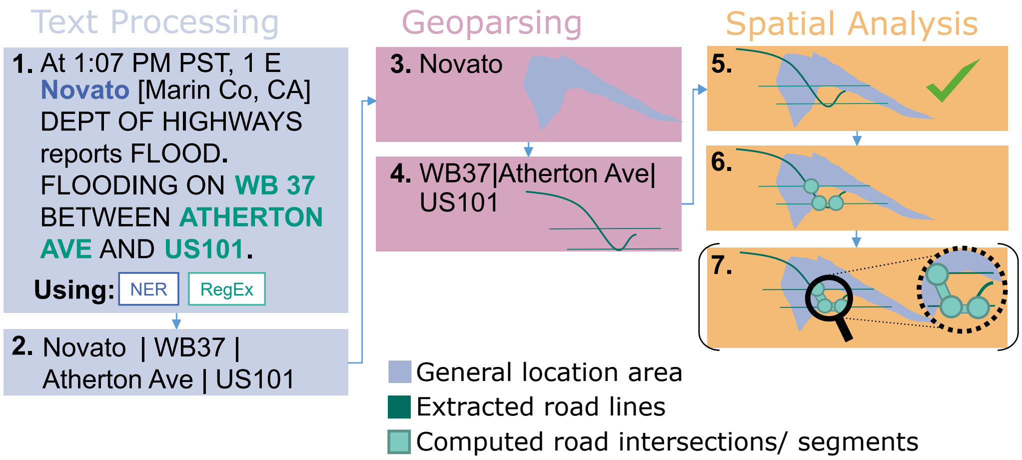

Figure 4. Visualization of the detailed road extraction framework during natural hazards from Twitter data starting from an exemplary tweet text.

Network:

The network analysis examines the accessibility of the road network, concentrating solely on the accessibility of locations within the road network. The graph nodes in the road network serve as the points of origin and destination for the analysis, referred to as starting and endpoints. Using cleaned graphs containing only relevant nodes representing road intersections is crucial; the conducted OSMnx preprocessing assures this for the employed OSM road data (compare Section 2). As the nodes are the starting points, the accessibility measures are calculated and stored for each network node. To estimate and visualize the accessibility of each edge as well, the edges' accessibility is calculated as the minimum value of its two connecting nodes, as an edge can be only as accessible as the node with the lowest accessibility among the nodes it is connected to.

Starting and Endpoints:

The starting and endpoints for calculating our accessibility measures depend on the specific network being analyzed and the selected measure. In general, the starting and end points of the centrality measures are all network nodes. The starting point can be any node within the network, and the endpoint is each of the other nodes in the network. Furthermore, the starting points for the index measures are all nodes within the network, too. However, these measures do not calculate the accessibility to each node (like the centrality measures) but only to a few specified endpoints. In our study case, the destination points are emergency shelters (see Section 2).

Data Thinning:

The starting and endpoints for the calculation of our accessibility measures depend on the specific network. Generally, all nodes are used as starting points. However, as the selected network for the Bobcat fire contains many edges and nodes, and each node would serve as a starting point for the measures of accessibility calculation, the calculation would require too much computation time. Therefore, we preprocess the network data and thin the node data used as starting points. We consider only nodes from roads of high hierarchy levels, not including nodes of residential roads as starting points. This technique aims to reduce the density of data points, allowing for a more manageable dataset while preserving crucial spatial information. Note that the accessibility calculation does consider all nodes and edges to find the shortest paths; only the number of starting points is reduced.

Degraded Road Network:

The degraded network is constructed from the uncorrupted network by eliminating the degraded segments (edges) and degraded intersections (nodes) as extracted in Section 3.1.

To estimate potentially degraded roads for our analysis, we eliminate roads (nodes and edges) overlapping with the respective hazard area as displayed in Section 2. Nodes and edges that are eliminated are considered *degraded*. When degraded nodes and edges are removed, parts of the network might become *disconnected* when they no longer connect to the rest of the network. In this case, we keep the leading network with the most connected components for the subsequent analysis. Degraded and disconnected nodes are considered non-accessible and, therefore, an accessibility measure is not calculated for them.

Hazard Impact:

To evaluate the impact of natural hazards on road network accessibility, all accessibility measures for the uncorrupted and the degraded network are calculated. Finally, to assess the network accessibility change, the difference in network accessibility before (uncorrupted network) and during (degraded network) the hazard is calculated. For each measure M for each node v, the impact is calculated as:

$$M_{\text{impact},v} = M_{\text{before},v} - M_{\text{after},v} = M_{\text{intact},v} - M_{\text{degraded},v} \quad (1)$$

3.2.2. Accessibility Measure Definition

Betweenness and closeness centrality indicate accessibility for a network in a general, holistic view. They are selected as measures representative of general measures of graph structure widely used in network theory.

Betweenness centrality (BN) [83] is a measure that quantifies the importance of a node within a network based on the number of shortest paths that pass through that node. BN of a node v is the sum of the fraction of all-pairs shortest paths that pass through v:

$$BN(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}, \quad (2)$$

where V is the set of nodes, $\sigma(s,t)$ is the number of shortest (s,t)-paths considering the travel time, and $\sigma(s,t|v)$ is the number of those paths passing through some node v other than s,t [84]. A node with a high BN value acts as a bridge or connector within the network. A node with a low BN value has limited influence on the overall network structure.

Closeness centrality (CN) [85] is a measure that quantifies how quickly a node can be accessed from other nodes in a network. It calculates the average shortest path distance from a node v to all other nodes in the network:

$$CN(v) = \frac{n-1}{\sum_{w=1}^{n-1} d(w,v)}, \quad (3)$$

where $d(w,v)$ is the shortest path distance between w and v considering the travel time, and $n-1$ is the number of nodes reachable from v. A node with a high CN value can reach other nodes in the network more quickly. On average, a node with a low CN value takes longer to reach other nodes in the network. All centrality values are finally normalized to the range [0, 1], fitted on the value range before the hazard, and transformed into the value range before and after.

Free-flow assumption accessibility indices have been used, e.g., for emergency response (EFAI [36]) or general connectivity investigations (ARIA [43]). For our evacuation scenario, we slightly adapted the EFAI towards reaching emergency shelters instead of emergency facilities. An emergency facility or shelter accessibility index can consider specific facilities and critical locations to provide a more targeted assessment of accessibility in the hazard-affected areas compared to global measures. Therefore, the Shelter Accessibility Index (SAI) for an origin node v is defined as:

$$SAI(v) = \sum_{S} \frac{t_{vS}}{t_S}. \quad (4)$$

t_{vS} is the travel time from origin v to the nearest shelter S, and $\bar{t_S}$ is the mean travel time of all origins to the nearest shelter S. The EFAI and ARIA define the index with a fixed threshold of three for the ratio. This threshold is employed to remove the effects of extreme values. EFAI and ARIA have predominantly been used to analyze overland road networks and focus primarily on the hierarchy of roads ranging from motorways to tertiary roads, excluding residential roads. Our study area, in contrast, includes all levels of road hierarchy. For example, residential roads contribute significantly to the total road count. This inclusion tends to reduce the mean travel time. However, in rural, mountainous regions, the travel times can be notably higher than the mean. When divided by relatively small mean values, destinations in these regions result in comparatively high index values.

Consequently, extreme values might be much more prevalent in some application scenarios, which makes it necessary to consider these values, too. Therefore, contrary to the EFAI and ARIA, we do not define the SAI with a fixed threshold of three for the ratio. A node with a low SAI value is more accessible, as it indicates that its travel time to the nearest shelter is relatively short compared to the average travel time. On the contrary, a node with a high SAI value is less accessible. Note that this measure operates inversely compared to all the other measures, where a higher value denotes greater accessibility. However, the existing and the adapted shelter indices are considered to be a free-flow scenario. These free-flow assumptions might be critical, especially for evacuation scenarios where many people leave the exact location at the same time. As many people are prone to taking the same routes (the fastest one), this can lead to congestion. These, in turn, lead to other routes being faster than the routes that are usually the fastest in free-flow cases but are not considered in the index calculation.

Therefore, an alternative routing assumption accessibility (ARAA) measure is advantageous. Since a free-flow assumption is probably wrong in evacuation scenarios, it would be preferable to correctly consider travel demand and capacity data to account for congested roads [44,86]. However, obtaining travel demand data is difficult, especially in a hazard-case scenario (compare Section 1). Therefore, we develop an approach that is independent of travel demand and considers routing alternatives to the respective shortest paths. The measure ARAA considers a selected number of k shortest paths. We assume that people would be using paths that are (a) still short (fast) compared to the shortest path, but (b) considerably different (few overlaps) to the shortest path, following Kondo et al. [50]. The accessibility measure is, therefore, the sum of the number of common nodes between the k shortest selected paths. In contrast to other approaches that try to find several non-overlapping shortest paths [50], we adapt the underlying assumption to develop an approach that uses the amount of overlapping as an accessibility measure. The fewer the found paths between two points overlapping partially, the more these points are accessible and vice versa.

We use the sum of the number of common nodes between the selected paths to measure the amount of overlapping of paths. The ARAA for each origin node v is therefore calculated as:

$$ARAA(v) = \frac{1}{\sum_{i,j=1}^{k} |Set(path_i) \cap Set(path_j)|} \tag{5}$$

for $path_i$, $path_j \in paths_{vS}(t)$, $i \neq j$, where k is the number of calculated shortest paths between origin node v to the nearest shelter S considering the travel time. A node with a high ARAA value is more accessible as it has more paths leading to a shelter that are not (partially) overlapping. A node with a low ARAA value is less accessible as it has fewer paths leading to a shelter that are not (partially) overlapping.

3.3. Evaluation

Evaluating extracted impacted roads is challenging, as no datasets concerning affected roads and network accessibility during the investigated hazards exist.

For the road extraction task, we evaluate according to its three separate functions: First, we evaluate how good the text processing is, meaning how many mentioned road names and intersections the methodology can extract from the texts. We manually count how many text messages our approach is missing and not using to extract roads. Secondly, we evaluate how many of the roads are geoparsed correctly, accordingly.

The evaluation of the changes to road network accessibility due to the hazard is equally challenging, as no ground truth (reference) datasets concerning observed network accessibility exist in general; results could only be compared to other studies' results. However, there are no reference data, since we use specific datasets for our study regions that have not been investigated. Therefore, we compare our four selected evaluation methods. Three of the methods employed have been applied in different contexts in different scenarios (see Section 1) and can be used to evaluate our newly developed approach. The established BN and CN have been extensively tested in numerous studies, making them valuable reference points for our assessment [52–54]. Additionally, our adapted SAI measure, which builds upon the EFAI measure, has undergone extensive evaluations of other accessibility index measures in previous studies [36,43]. By comparing the outcomes of these measures in our case study scenarios with those obtained using the established BN, CN, and adapted SAI measures, we can ascertain whether our developed measures findings confirm or contradict their results, thereby providing validation of our approach. However, it is essential to acknowledge that a direct comparison of outcomes, especially for complex network alterations, from these measures is difficult, given their distinct conceptualizations of accessibility stemming from divergent underlying definitions. As a result, their comparability remains limited.

4. Results

4.1. Degraded Road Extraction

As explained in Section 3.3, the road extraction consists of three tasks we evaluate separately. We collected 32 (for the complete Bay Area during the flood) and 2 (for the Bobcat wildfire) non-identical road-related and hazard-related tweets. Table 2 displays the number of text messages from which we can extract road information.

For the Bay Area floods, from a total of 32 messages that contain road location information, four messages have data that are not accurate enough for road information extraction (e.g., *flooding in all lanes on Interstate 580 EB*). This message would allow us to extract the road name Interstate 580. However, since the highway passes through the total San Francisco Bay Area region and from the text, we do not know which road segment(s) is/are meant. Therefore, these messages are considered too unspecific. Furthermore, the methodology fails on another four messages in the text processing or geoparsing steps. The name extraction fails due to different naming conventions, e.g., *Camino Pablo*. In this example, the English RegEx and NER do not consider Camino (Spanish) a road name. The geoparsing fails because the extracted road location names cannot be linked with a geographical location (e.g., *Bryant onramp*) as they are not named in our database. We have obtained 24 messages from which road extraction is possible. For the Bobcat wildfire, we have obtained a total of 303 messages that contain road location information. However, most texts contain road information that does not consider degradation from the investigated fire but, e.g., general road closures. Since we want to concentrate on roads degraded from impact from hazards, we have chosen only texts associated with fire. Furthermore, the degradation of the same road is reported several times if the hazard lasts several days. After removing messages with identical information, we retain two messages associated with the Bobcat fire. From these two messages, road extraction is possible.

Table 2. Summary of the number (#) of road names that are extracted correctly.

	# With Road Locations	# Too Coarse Road Location	# Failed	# Extracted
Bay Area flood	32	4	4	24
Bobcat wildfire	2	0	0	2

With our road extraction method, we obtain three types of information about roads, depending on the information available in the text:

1. Only a single road name is mentioned, which means we can extract the total road segment as a line feature.
2. Only an intersection of two roads is mentioned, so we can extract this intersection as a point feature.
3. Several roads and/or specific intersections of several roads are mentioned, which means we can extract the road segments between the said intersections as a line feature. Therefore, this type of information offers the most precise impacted road segment extraction.

Figure 5 displays the different extracted roads for the application scenario of the Oakland flood. We observe that, from most text information data, we obtained road segments and corresponding intersection locations. For one example, we only extract a single intersection information, which does not allow us to extract a corresponding affected road segment. The extracted roads are distributed all over the investigated areas.

Furthermore, we observe that major roads are mostly being extracted, as these are talked about in tweets from road agencies.

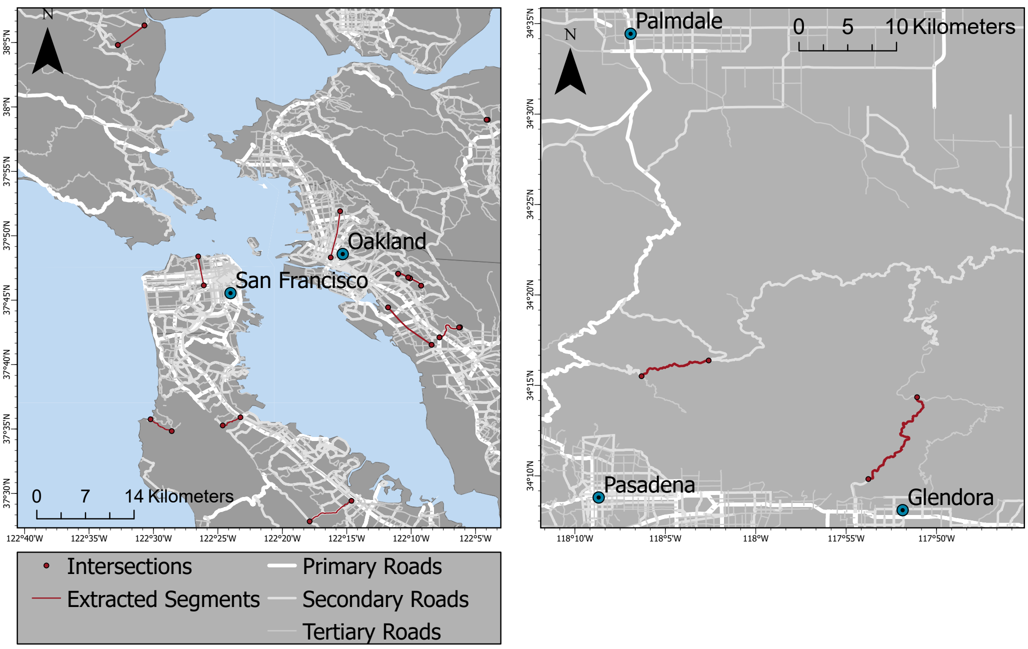

Figure 5. Visualization of extracted degraded roads from VGI text information for the flood in the San Francisco Bay Area in 2022/23 (**left**) and the Bobcat fire in 2020 (**right**). Data basis: © 2018 GADM. Projection: WGS84.

4.2. Degraded Road Network Construction

In this section, we present the construction of the degraded network. In addition to our extracted degraded roads (compare Section 4.1), we include roads that overlap with the respective natural hazard extent into the degraded road dataset. We use the hazard impact areas as described in Section 2. However, due to the substantial size of the San Fransisco Bay Area region, we restrict the application of the degraded network analysis exclusively to a subset within this area, the city of Oakland.

Figure 6 displays the road edges and the degraded nodes and degraded edges for the respective case study areas. We construct the degraded network for the network analysis from the intact network, removing the degraded nodes and edges according to extraction from road text information and from overlapping the natural hazard extent. We observe that, for the flood hazard, many more roads are considered degraded due to the location of the hazard in an urban area with a higher road density in general. Due to the high node and edge density, most nodes stay connected even after removing degraded nodes and edges. Several points in the south (near Oakland Airport) became disconnected due to the removal of the degraded road from the network.

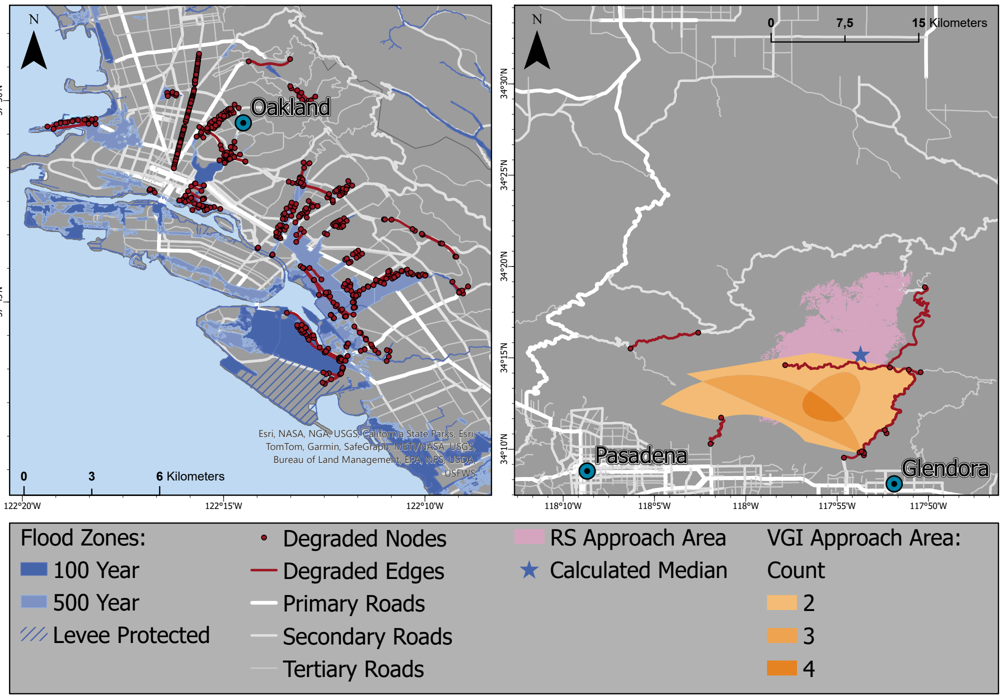

Figure 6. Visualization of the road network and its degradation for the flood in Oakland in the San Francisco Bay Area in 2022/23 (**left**) and the Bobcat fire in 2020 (**right**). Degraded roads consist of road segments extracted as degraded from text data directly, and roads that overlap general hazard areas and are therefore considered degraded implicitly. Data basis: © 2018 GADM. Projection: WGS84.

For the fire hazard, a smaller number of roads are degraded as the hazard is located in a rural area. However, we see that major roads, which are essential for the connections to several places in the mountains, are degraded. A few nodes in the center close to the fire became disconnected due to the degraded roads that are the only ones leading to those nodes. Other connections remain intact but require detours around the degraded roads.

4.3. Network Analysis

In this section, we present the results of the network analysis. We present the results of the four accessibility measures for the network edges because the visualization of the node accessibility values' becomes less effective when numerous nodes overlap. The display of the edge accessibility values allows a more straightforward interpretation of the degraded and less accessible road sections, facilitating the assessment of the hazard's impact on the road infrastructure.

4.3.1. Oakland Floods

In the following, we present the results of the four measures applied to the intact and degraded road network for the Oakland study region. Furthermore, the change in network accessibility from before (intact road network) and during/after (degraded road network) a flood is presented for the four measures. The change is called the hazard's impact on road infrastructure in the following. We visualize the road network accessibility and the hazard's impact estimated by the four measures in a road map (see Figure 7). The values of these measures were grouped into five classes defined by quintiles. For the visualization of the after-hazard accessibility, the quintiles of the before-hazard accessibility are used to show possible changes.

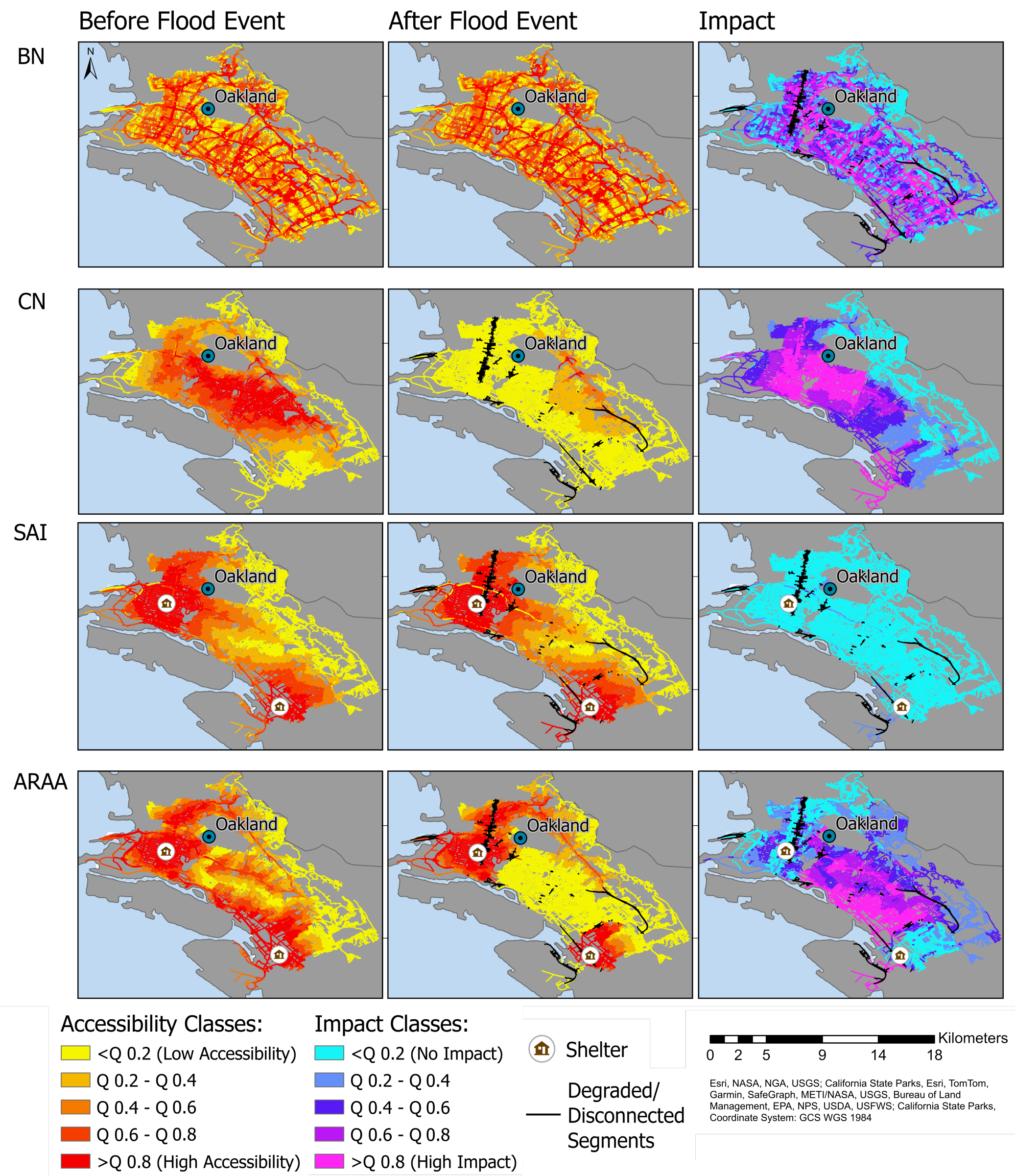

Figure 7. Visualization of the betweenness centrality (BN), row 1, closeness centrality (CN), row 2, shelter accessibility index (SAI), row 3, and alternative routing assumption accessibility (ARAA), row 4, measure for the flood in the San Francisco Bay Area displayed for the city of Oakland and the intact network before the hazard, the degraded network during/after the hazard and the change in accessibility between the two for the network edges. Quintiles for each measure are given in the text. Note that the SAI impact values are visualized with different intervals (interval boundaries I1: 0.1, I2: 2.45, I3: 4.89, I4: 7.33). Data basis: © 2018 GADM. Projection: WGS84.

Figure 7, row 1, shows the results for the BN accessibility measure for the road network before and during/after the hazard ($Q_{0.2}$: 0.0053, $Q_{0.4}$: 0.0130, $Q_{0.6}$: 0.0300, $Q_{0.8}$: 0.0760) and the accessibility change due to the hazard impact ($Q_{0.2}$: 0.00016, $Q_{0.4}$: 0.00057, $Q_{0.6}$: 0.00170, $Q_{0.8}$: 0.00730, range $[0, 0.12]$) for the Oakland application scenario. This measure highlights major traffic arteries as highly accessible, while roads in residential neighborhoods are less accessible. As several significant arteries become less accessible due to the impact of the hazard, we see a shift of accessibility towards other arteries. Some arteries were moderately accessible before the hazard and became highly accessible after the hazard. These become more important for the general traffic flow, as the previously important ones are less accessible. Roads in urban neighborhoods remain lowly accessible. Consequently, we see less change in neighborhood roads and, generally, to the east of the area, where accessibility has not changed much, as no significant arteries were present before or after the hazard. The BN computation performed on the Oakland dataset, which comprises approximately 8500 nodes, exhibits a computational time of 29 min and 33 s utilizing a standard notebook equipped with an Intel Core i7-10875H CPU featuring eight cores.

Figure 7 row 2, shows the results for the CN measure for the road network before and during/after the hazard ($Q_{0.2}$: 0.64, $Q_{0.4}$: 0.73, $Q_{0.6}$: 0.79, $Q_{0.8}$: 0.83) and the accessibility change due to the hazard impact ($Q_{0.2}$: 0.0004, $Q_{0.4}$: 0.0005, $Q_{0.6}$: 0.0007, $Q_{0.8}$: 0.0009, range $[0, 0.001]$) for the application scenario. Before the hazard, using this measure, edges in the center of the study area were much more accessible, while edges at the area's borders were less accessible. For large parts of the area, accessibility is very high, especially in the center of the study area. During/after the hazard, large parts of the area are less accessible overall, especially in the center-west. The edges in the center-east remain the most accessible. Looking at the change, we see a pattern where especially edges near the center-east area show a higher change to less accessibility. Accessibility of edges in the east remains about the same. We can see which neighborhoods are clearly affected by the impact of the hazard on the road network and, therefore, the change in accessibility.

Figure 7, row 3, shows the results for the SAI measure for the road network before and during/after the hazard ($Q_{0.2}$: 1.3, $Q_{0.4}$: 1.1, $Q_{0.6}$: 0.9, $Q_{0.8}$: 0.6, range $[0, 10]$) and the accessibility change due to the hazard impact ($Q_{0.2}$: 0.012, $Q_{0.4}$: 0.018, $Q_{0.6}$: 0.022, $Q_{0.8}$: 0.027, range $[0, 9.86]$) for the application scenario. Note that for this measure, the value range is inversed, where a high SAI value represents low accessibility and a low SAI value represents high accessibility. Using the shelters as destination points, we see a significantly higher accessibility of the edges close to the shelters. For the intact road network, the center of the study area is also comparatively highly accessible. However, the center north is more accessible because it is more connected to the northern shelter area, mainly by a major road that runs to the south. The eastern study area is the least accessible. For the degraded network, we see almost the same pattern of accessibility overall, as places near the shelters, whose paths are not obstructed, are still very accessible. However, several significant arteries, such as Interstate 580 and 880 (Points A and B), become much less accessible in specific segments near degraded roads. The area near the Oakland Airport (Point C) obtains higher accessibility values. The distribution of the SAI values is extremely right-skewed (median: 0.02) with 95.2% of the values lying below the value of 0.1, meaning that the accessibility of most road segments has not changed. The first interval (interval boundary I1) is defined as "*no change*" with values below 0.1. Four further classes are formed with equal interval boundaries (I2: 2.45, I3: 4.89, I4: 7.33). Only for a few roads, changes in accessibility have been observed around points A, B, and C and a few locations close to degraded edges. Exemplary, we added a close-up visualization for a subset of the Oakland investigation region for the SAI measure (Appendix A).

Figure 7, row 4, shows the results for the ARAA measure for the road network before and during/after the hazard ($Q_{0.2}$: 5.05, $Q_{0.4}$: 6.09, $Q_{0.6}$: 7.24, $Q_{0.8}$: 9.25, range $[0, 55]$) and the accessibility change due to the hazard impact ($Q_{0.2}$: 0.00, $Q_{0.4}$: 0.34, $Q_{0.6}$: 2.11, $Q_{0.8}$: 3.98, range $[0, 39]$) for the application scenario. Similar to the SAI measure results, we see the most accessible places near the shelter points. Accessibility declines quickly when moving away from these locations. Major traffic arteries and connecting segments are shown with slightly higher accessibility values compared to purely residential neighborhood roads. The least accessible areas are some in the center, in the east, and the southeast of the study area. In the degraded network, the areas nearest to shelter locations maintain high accessibility, as degraded roads do not need to be used to reach shelters. The areas at the center become less accessible. Regarding the change in network accessibility, we observe that the areas close to the shelter locations that already had good accessibility before and after the hazard experienced the least amount of change. The center of the study area between the two shelters is most affected by the impact of the hazard, while the central-east and northwest show moderate changes in accessibility.

4.3.2. Bobcat Wildfire

In the following, we will present the results of the four measures applied to the intact Angeles National Forest road network before the fire and the degraded road network affected by the fire. Furthermore, the change in network accessibility from before and during the fire is subsequently presented for the four measures. Accessibility and impact value visualizations are conducted as described in Section 4.3.1.

Figure 8, row 1, shows the results for the BN accessibility measure for the road network before and during/after the hazard ($Q_{0.2}$: 0.0003, $Q_{0.4}$: 0.0008, $Q_{0.6}$: 0.0021, $Q_{0.8}$: 0.0081) and the accessibility change due to the hazard impact ($Q_{0.2}$: 0.00022, $Q_{0.4}$: 0.00064, $Q_{0.6}$: 0.00120, $Q_{0.8}$: 0.00330, range $[0, 0.002]$) for the application scenario. Similar to the Oakland application scenario results, the Bobcat application scenario results reveal that this metric highlights major traffic arteries as having high accessibility. On the other hand, roads within residential neighborhoods and rural mountainous regions, which do not serve as significant city connectors, exhibit lower accessibility levels. The wildfire event does not significantly affect the condition of the road network. However, three road segments on both sides of the degraded segment of the Angeles Crest Highway (Hwy) (Point A) experience degradation, resulting in decreased accessibility. A minor alteration in accessibility is observed for a few road segments situated in the southern part of the Angeles National Forest (around San Gabriel Canyon Road, Point B) and the urban area close to degraded edges. Figure 8, row 2, shows the results for the CN accessibility measure for the road network before and during/after the hazard ($Q_{0.2}$: 0.63, $Q_{0.4}$: 0.67, $Q_{0.6}$: 0.76, $Q_{0.8}$: 0.80) and the accessibility change due to the hazard impact ($Q_{0.2}$: 0.000010, $Q_{0.4}$: 0.000025, $Q_{0.6}$: 0.000059, $Q_{0.8}$: 0.00027, range $[0, 0.00042]$) for the application scenario. For the intact road network, a cluster characterized by high accessibility is situated in the southwest region. In contrast, the accessibility in the northern region is comparatively lower according to this metric. The accessibility of connecting road segments between the south and north regions is moderate, depending on the road's proximity to the respective clusters. In the degraded network, for road segments crossing the Angeles National Forest and linked to degraded segments, a decrease in accessibility exists. Looking at the change, the Angeles Crest Hwy (Point A) and connecting edges in the Angeles National Forest are significantly affected along its length. A few connecting road segments leading to the Angeles Crest Hwy and a few mountain roads connected to the San Gabriel Canyon Road (Point B) are also impacted.

Figure 8, row 3, shows the results for the SAI measure for the road network before, during, and after the hazard ($Q_{0.2}$: 1.49, $Q_{0.4}$: 1.03, $Q_{0.6}$: 0.81, $Q_{0.8}$: 0.53, range $[0, 4.79]$) and the accessibility change due to the hazard impact ($Q_{0.2}$: 0.0012, $Q_{0.4}$: 0.0018, $Q_{0.6}$: 0.0023, $Q_{0.8}$: 0.0033, range $[0, 1.63]$) for the application scenario. Note that for this measure the value range is inversed. A high SAI value represents low accessibility and a low SAI value represents high accessibility. For this measure, we rely on the thinned node data (as described in Section 3.2.1), which does not include residential roads' nodes and edges as starting points. Residential neighborhood roads in the northeastern study area are not displayed. Using the shelters as destination points for this measure, we see significantly higher accessibility of the edges near these shelter locations. For the intact road network, edges in a relatively large radius around the shelters exhibit notably high accessibility levels. The accessibility levels decline as one moves farther away from these shelter points, particularly in the mountainous areas. For the degraded network, the analysis reveals no visualized fluctuations in accessibility. Nonetheless, when assessing the visual representation of the change based on the suitable tailored intervals "*no change*" (interval boundary I1: 0.1) and equal intervals on the remaining values (I2: 0.58, I3: 0.93, I4: 1.27) (explanation see Section 4.3.1), it becomes apparent that there is a marginal decrease in accessibility along the Angeles Crest Hwy (Point A) and its connecting edges. Moreover, the connecting route to the San Gabriel Canyon Road (Point B) in the Angeles National Forest exhibits a substantial decrease in accessibility.

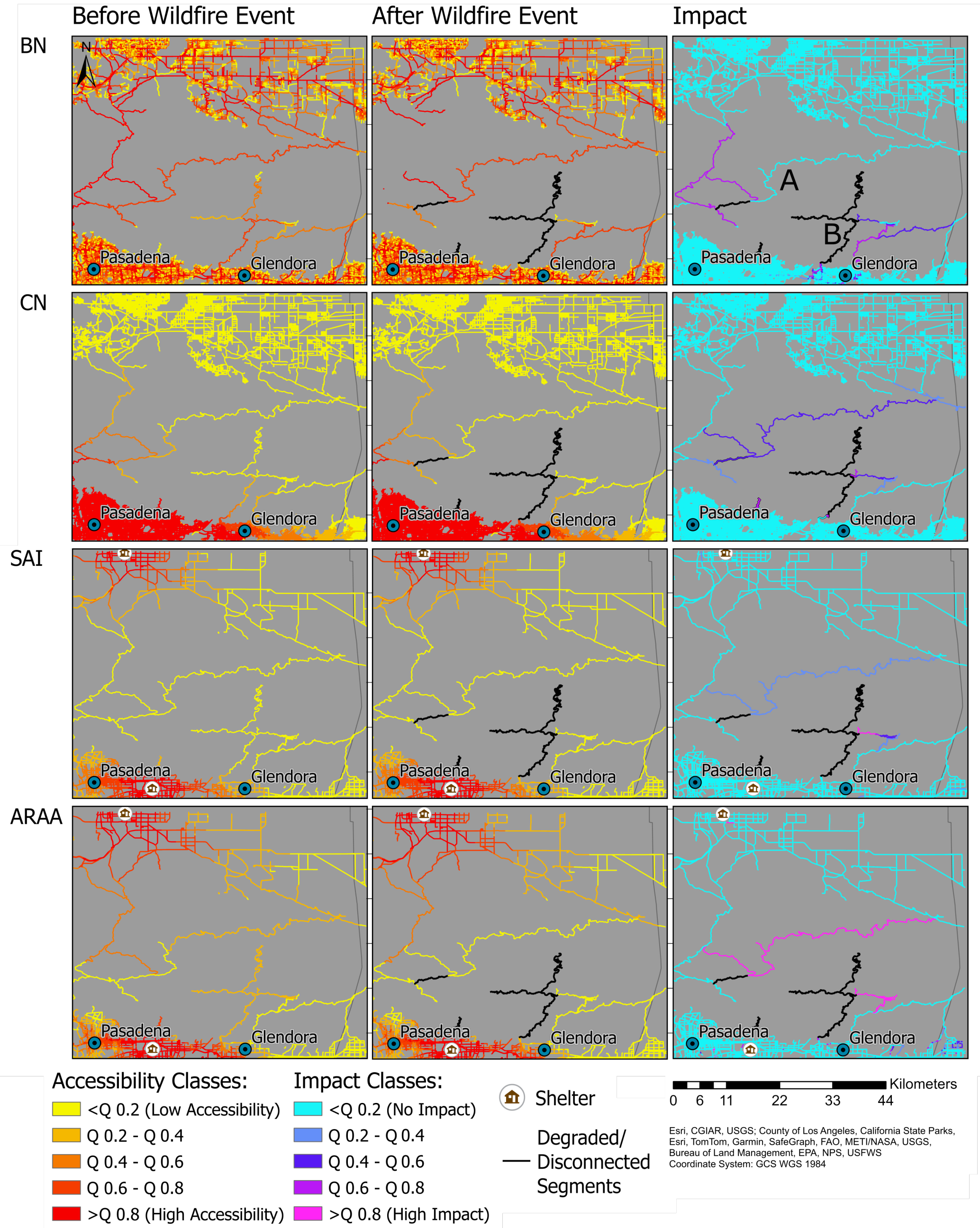

Figure 8. Visualization of the betweenness centrality (BN), row 1, closeness centrality (CN), row 2, shelter accessibility index (SAI), row 3, and alternative routing assumption accessibility (ARAA), row 4, measure for the Bobcat wildfire displayed for the intact network before the hazard, the degraded network during/after the hazard and the change in accessibility between the two for the network edges. Quintiles for each measure are given in the text. Note that the SAI impact values are visualized with different intervals (interval boundaries I1: 0.1, I2: 0.58, I3: 0.93, I4: 1.27). Data basis: © 2018 GADM. Projection: WGS84.

Figure 8, row 4, shows the results for the ARAA measure for the road network before, during, and after the hazard ($Q_{0.2}$: 4.44, $Q_{0.4}$: 5.92, $Q_{0.6}$: 7.20, $Q_{0.8}$: 9.34, range $[0, 91]$) and the accessibility change due to the hazard impact ($Q_{0.2}$: 0.000001, $Q_{0.4}$: 0.039, $Q_{0.6}$: 0.045,

$Q_{0.8}$: 0.267, range $[0, 2.31]$) for the application scenario. For this measure, we also rely on the thinned node data, which do not include residential roads' nodes and edges as starting points and are, therefore, not displayed. Using the shelters as destination points for this measure, we see significantly higher accessibility of the edges near these shelter locations. Edges in the mountainous areas still have comparatively high accessibility values as more alternative routes to shelters can be found here, moving towards both valley sides. In contrast, urban areas, e.g., in the south-east, mostly connected to only one shelter without alternative routes, are least accessible. For the degraded network, it is again challenging to visualize the slight accessibility decrease effectively. The impact visualization shows that the Angeles Crest Hwy (Point A) and connecting edges to degraded edges overall are highly affected.

5. Discussion

The road information extraction methodology was conducted for two different hazard events: the San Francisco Bay Area floods and the Bobcat wildfire. For both case studies, we have a low number of non-repeating texts about the road conditions during the hazards. However, these few texts allow the extraction of relevant impacted road segments for this hazard. We encountered a few challenges in accurately extracting road information from the messages. We found that a few messages needed to be more accurate for road information extraction, containing only general road names without intersection locations. These messages are not suitable for our analysis. Additionally, we encountered issues with the name extraction step that failed due to different naming conventions and the geoparsing step that failed due to location names needing to be linkable to geographical locations. Furthermore, when only a single intersection location is extractable, determining the corresponding affected road segment is impossible. However, a substantial number of roads that are impacted by the hazard can be extracted with a high level of detail (intersection-wise). Due to the extraction from authority data, which verifies and often updates the data compared to other tweet data, these roads can be extracted with a high level of certainty.

Our presented methodological approach is the first approach towards precise road place extraction, including name extraction and geocoding, which has not been conducted before. Though Yu et al. [87] focused on precise road name extraction, they do not focus on geocoding. Other studies do not focus on road extraction specifically and often use geocoding with nonspecific place names, e.g., [18,22]. A drawback of the developed methodology is that extracted roads primarily consisted of major roads, as they were the ones prominently discussed in general tweets and from road agencies. Information about smaller impacted roads cannot be gained from these data sources.

From the network accessibility analysis findings in Section 4.3.1, we conclude that the flood hazard heavily impacts the road network of the Oakland study region. The degradation of many major roads is mentioned in Twitter messages texted during the 2022/23 flood. Additionally, when including minor roads that are most likely influenced as derived from the inundation zones map, large portions of the road network became degraded due to the hazard. The degradation influences the network accessibility largely. Furthermore, the overall road network accessibility changes heavily compared to the intact and the degraded road network. Large portions of the total area experience changes due to the impact of the hazard. This results from having many impacted roads in a comparatively small urban area, where impacts on major roads that traverse the city largely influence the whole road network.

From the findings in Section 4.3.2, we concluded that the road network is not heavily impacted by the wildfire hazard for the Bobcat wildfire application scenario. Though significant roads that lead in the Angeles National Forest are impacted, these do not influence the overall network degradation as there are no connecting roads to other major roads. A segment of the Angeles Crest Highway is degraded, but this closure does not cause an overall degradation, as a bypass is available for this segment. The degradation only slightly impacts the total Angeles Crest Highway accessibility. Therefore, the road network accessibility overall changes only slightly compared to the intact and the degraded road network. The overall accessibility remains high due to the numerous connections between cities south and north of the Angeles National Forest that do not necessitate travel through the mountains. Additionally, accessibility within the respective city regions is unaffected.

Comparing the outcomes of accessibility measures applied to the Bobcat wildfire study with the Oakland flood application scenario reveals significant differences. In the Oakland application scenario context, the developed measures show that the flood hazard heavily impacts the road network overall. Furthermore, the developed measures show divergent results. On the contrary, within the Bobcat application scenario, the measures mostly agree. Furthermore, they consistently identify a few edges as the most impacted ones. These results can be linked to the difference in network characteristics (e.g., urban vs. regional, rural, crossing mountains) and the quantity and spatial distribution of degraded edges within each network for the two application scenarios. Due to the excellent alignment of the accessibility measures in the Bobcat application scenario, the impacted edges can be identified, and the impacts of the hazard on the road infrastructure can be evaluated with high certainty. Therefore, the developed framework offers valuable insights into road accessibility for first responders in hazard scenario route planning.

Four accessibility measures have been investigated. Table 3 shows a short summary of the advantages and drawbacks of each measure. The BN gives a good overview of general network accessibility by highlighting important and highly accessible traffic arteries. In the context of the degradation of the road network due to a hazard, this measure balances the accessibility degradation of former major arteries with the accessibility amelioration of new major arteries. Therefore, this measure is valuable for route planning to select major arteries that can provide fast access to edges in intact road networks and degraded networks. Here, displaying whether a positive or negative change happened to the different edges would be valuable. However, the measure's drawback is its tendency to ignore absolute accessibility. It highlights new arteries as highly accessible, even though they may still be less accessible than the former major arteries of the intact network.

Using the CN, it is easily distinguishable which edges in the network are able to reach other edges easily. For network degradation, the measure displays neighborhoods that become less accessible. Overall, the accessibilities in all areas are similar and less scattered and heterogeneous than for the BN. This helps to investigate which regions might have more difficulty being reached by emergency vehicles or accessing other places in the case of an evacuation. The drawback of this method is the major dependence of the accessibility value on the selected investigation area. Edges in the center of the selected area are, by default, more accessible. This measure computes a universal picture, with almost all roads experiencing a drop in accessibility due to the degraded roads. Therefore, the measure is suitable to display changes in accessibility accurately but needs to be used with caution for absolute accessibility analyses.

Our BN and CN findings correspond with other studies that used those measures for network accessibility analysis (e.g., [52–54]). These measures provide valuable insights into networks' structural characteristics and efficiency, and are fast and easy to apply to any network. BN and CN are global measures that capture the overall importance of nodes in the network very well. However, they heavily rely on the network structure. Due to natural hazards, the degraded road network has an altered network structure. Therefore, BN and CN values may change significantly (BN for major arteries, CN overall, compare Section 4.3.1), even for minor disruptions. Minor changes can completely alter the selected paths, making different nodes more important from a theoretical perspective. Viewed from a practical point of view during a natural hazard, it is important to recognize that certain regions or specific roads that hold significance in the intact network may retain their critical importance even when assigned lower accessibility values in a degraded network.

Accessibility indices can be adapted to situational requirements, like immediate emergency response (evacuations or emergency facility access [36]) or general connectivity investigations (remoteness of places [43]). These indices can consider specific facilities and critical locations to provide a more targeted accessibility assessment during a natural hazard. Compared to the previous measures, the SAI is adaptable to situational requirements for route planning investigations to provide a more targeted assessment of the accessibility of regions. However, the inverse definition of the SAI (like EFAI and ARIA) could be more optimal, and this measure is less straightforward and more challenging to interpret. Adapting the SAI to avoid using a fixed threshold (see Section 3.2.2) enhances its usability for networks with all hierarchies of roads, especially when compared to EFAI and ARIA [43,88].

However, the SAI still follows an extremely right-skewed distribution and is not optimal for the visualization of accessibility and especially its changes. Only a few major arteries (Points A and B in Figure 7, row 3) and few other roads, especially in the South (around Point C), experience changes. Additionally, due to the definition of this measure, using the mean travel time value, roads near the Oakland Airport (Point C in Figure 7, row 3) become more accessible compared to the mean, as the mean value becomes larger overall. Furthermore, considering only one shortest path, the measure can be seen as too optimistic in evaluating the impact on accessibility. In reality, when evacuation orders are given, and all people decide to choose the depicted shortest path, the accessibility of this path would drop rapidly. This would lead to congested scenarios where free-flow measures cannot estimate accessibility.

The ARAA can be adapted to situational requirements as much as the SAI, e.g., for shelter or emergency facility accessibility. With the ARAA, major traffic arteries, as well as regional clusters of accessibility, can be easily detected. In the context of the degradation of road networks, the ARAA can give clear statements about areas of change in accessibility. For the impact, the ARAA shows a very detailed and differentiated accessibility value pattern for the impacted regions (center of study region). Few edges closest to the shelter locations are considered highly accessible when assuming congestion.

The ARAA can be adjusted to simulate congestion severity by augmenting the number of k paths to be computed. Furthermore, due to the use of alternative routes, it is a more realistic measure for congested scenarios than free-flow measures [88]. Therefore, this measure is valuable for route planning in hazard scenarios compared to other accessibility-based measures [57]. Additionally, it is valuable compared to methods that rely on more complex and often non-available data like traffic data (e.g., [44,86]). However, due to the higher number of paths to be calculated, one drawback of the ARAA is the longer calculation time.

We employ a comparison between our four selected methods as a means of evaluation, as three of the used methods have been applied in different contexts in different scenarios (see Section 1) and have been extensively tested. They are state-of-the-art methods and can be used to evaluate our newly developed approach. Comparing the outcomes of the measures in our Oakland flood case study scenario shows that evaluating the employed accessibility measures is problematic. The measures show contradicting results for the accessibility of specific road segments. This is because the measures focus on a network's different attributes, e.g., a node's connection function or reachability to any other node. Furthermore, they consider different scenarios, e.g., general accessibility or evacuation. In this case study with a high number of degraded roads in complex patterns, direct comparison of outcomes from these measures is unfeasible, given their distinct definitions of *accessibility* stemming from divergent attributes and application uses. As a result, their comparability needs to be improved. Nevertheless, the measures complement each other by providing divergent perspectives on road network accessibility.

By comparing the outcomes of the measures in our Bobcat wildfire case study scenario, evaluating the employed accessibility measures is feasible. Notably, the SAI and ARAA measures exhibit consistent findings, with the same impacted road segments only with slight variations in magnitude. Similarly, the CN measure also presents confirmatory signals, identifying one additional impacted road segment, further reinforcing the overall coherence of our analysis. Despite its inherently different definition, the BN measure yields similar results.

These contribute to a more comprehensive understanding of the accessibility dynamics within the road network during natural hazards and enable a nuanced evaluation of the network's vulnerabilities in hazard management. However, for evaluating the developed network accessibility analysis, the comparison to more detailed network analysis models, for example, for a defined hazard where traffic information is available, would be favorable. For example, including road capacity data in congestion scenarios using detailed modeling [89,90] can give more detailed insights into network accessibility. The complexity of modeled parameters concerning network accessibility could be increased to a high degree according to the desired accuracy and detailing of accessibility analysis.

Furthermore, a straightforward model for evacuation scenarios is used in this case: we assume that people are evacuated to two shelters. In reality, evacuation scenarios are much more complex, incorporating, e.g., scenarios where people choose their evacuation destination or where roads are blocked to regulate traffic [91]. Such scenarios are defined, e.g., in the *Sea to Sky Multimodal Evacuation Plan* developed by the District of Squamish and the Resort Municipality of Whistler [92]. However, in our framework, the destinations can easily be adapted according to the necessities and each authority's evacuation plans.

Table 3. Advantages and disadvantages of different accessibility measures.

Measure	Advantages	Disadvantages
Betweenness	+ Fastest connections easily distinguishable + In hazard case newly fastest connections easily distinguishable	- No differentiation for absolute accessibility
Closeness	+ Area-wide not punctual accessibility values + Good for degradation impact on accessibility	- Accessibility depends on the choice of network borders - Not for absolute accessibility estimation
SAI	+ Adapted to situational requirements + Most important road degradation highlighted	- Reduced interpretability - Too optimistic (shortest path does not consider possible congestions)
ARAA	+ Adapted to situational requirements + Less optimistic (realistic for congestions) + Important arteries visible	- Calculation time

The presented study addresses the gap in the existing literature related to road network analysis in a natural hazard context. While some studies have considered degraded network scenarios, they often focus on single case studies in specific locations, limiting their applicability to other contexts. Additionally, accessing complex degraded road datasets after a natural hazard is challenging, which hinders the transferability of proposed advanced models or simulations.

The framework tackles the major challenges for a holistic hazard-impacted road network analysis. The framework is complete, from the hazard data acquisition to the road network accessibility analysis. The framework ensures that all necessary steps are integrated by covering the entire process chain, eliminating potential gaps. The framework uses only freely available data and does not rely on restricted access data or data that might not be accessible during a natural hazard crisis. This characteristic holds for both modules, using Twitter data for road degradation extraction and OSM data as a network basis. RS data from Sentinel-2, as also described in Florath et al. [93], and authoritative data are also openly accessible. Furthermore, the framework is applicable in near-real time. One advantage of the data used is that they are primarily available near real-time during or shortly after a natural hazard event. A rapid understanding of network accessibility can save precious time for emergency planning, which is crucial for the affected population. Another significant aspect is the framework's independence from traffic data. Obtaining daily or average travel demand data is often difficult [45,46], and this difficulty is intensified in hazardous situations. Even when average travel demand data are available, they may not accurately reflect travel patterns in hazardous situations. As a result, our framework that does not rely on other traffic data is more useful in such situations. Moreover, the framework is generic and applicable independently of the scale and location of the investigated road network. The framework enhances the applicability and transferability of the results across different contexts and hazards, whether dealing with an urban or rural setting or a regional or local network. The framework's broad applicability eliminates the need to reinvent methodologies for each unique scenario, saving time and effort. Overall, the framework is suitable for rapidly estimating network changes due to the impact of a natural hazard.

6. Conclusions

The generic framework introduced in this study for the network accessibility analysis during natural hazards consists of two major modules: degraded road extraction from VGI and network analysis of degraded road networks.

The road extraction methodology allows the extraction of impacted roads from text messages of VGI data. Overall, our road information extraction methodology provided valuable insights into the impact of hazards on road networks. While facing particular challenges, it still demonstrated its potential in capturing relevant road information from social media messages during hazard events. However, future improvements could be made to address the limitations encountered during name extraction and geoparsing. Improvements could include, e.g., RegEx patterns for other languages that could extract Spanish road naming conventions or the integration of different databases, including highway ramp names for geoparsing. Additionally, as the developed methodology is based on general text assessment, it could be transferred to other textual information sources, such as police reports documenting road issues. This expansion could improve the accuracy and increase the amount of extracted information regarding impacted roads.

Different measures of accessibility change are employed to assess the network's performance before and during the hazard, all of which are valid for road network impact analysis. Two standard accessibility measures ([83,85]) are benchmarks for evaluating and contrasting our developed approaches. Specifically, we introduced an alternative routing assumption accessibility measure that considers anticipated congestion scenarios. All employed measures achieve the objective of estimating road network accessibility without using additional traffic data. Therefore, they are all suitable for road network accessibility analysis independently of restricted global data availability on hazard-induced road network impacts. In total, the use of each measure is justified, as each measure could demonstrate the accessibility change due to the degradation of the network. However, their combination and comparison give valuable insights. In summary, this framework offers a quick and straightforward assessment of the impact of natural hazards on road networks. It is an initial step for disaster risk management planning shortly after a hazard event. The proposed approaches are particularly valuable due to their applicability using only degraded road network information without any additional data sources.

In conclusion, this study successfully addresses the challenges of developing a generic, complete framework from impact extraction to network analysis independently of scale and characteristics of road network types. The proposed generic framework offers a valuable contribution to road network analysis in a natural hazard context. The exemplary applications in different hazard scenarios demonstrate the framework's versatility and effectiveness in assessing road network performance during hazardous events. This study provides a foundation for further research and practical applications in emergency response and disaster management.

The integration of the developed framework with other components could be investigated in the future. Due to its modular structure, the hazard impact estimation on road infrastructure or network analysis could be included. As demonstrated, external impact zone data allow for a more thorough analysis of the impacts on road infrastructure before module 2 is applied. Further, more detailed approaches for network analysis, considering additional data sources like travel demand data, could be added to module 1.

Author Contributions: Conceptualization, Janine Florath, Jocelyn Chanussot and Sina Keller; methodology, Janine Florath, Jocelyn Chanussot and Sina Keller; software, Janine Florath; validation, Janine Florath; formal analysis, Janine Florath; investigation, Janine Florath, Jocelyn Chanussot and Sina Keller; resources, Janine Florath; data curation, Janine Florath; writing—original draft preparation, Janine Florath; writing—review and editing, Janine Florath and Sina Keller; visualization, Janine Florath and Sina Keller; supervision, Jocelyn Chanussot and Sina Keller. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the AXA Research Fund and by MIAI@Grenoble Alpes (ANR-19-P3IA-0003).

Data Availability Statement: The data supporting these findings are available from the corresponding author upon reasonable request.

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

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Close-Up Visualization

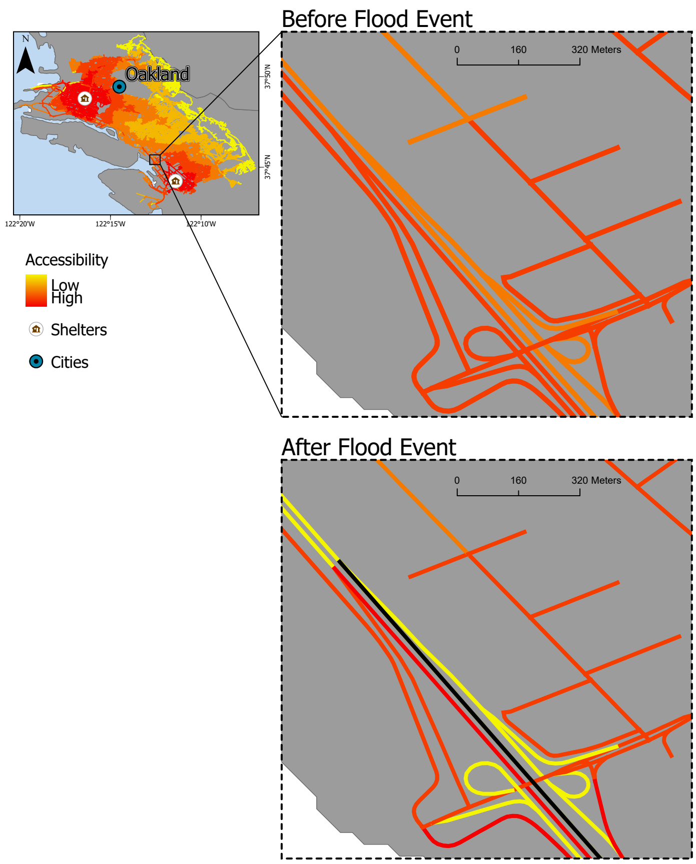

Figure A1. Visualization of the shelter accessibility index (SAI) measure for the San Francisco Bay Area flood displayed for a subset of the Oakland area for the intact network before the hazard and the degraded network after the hazard. Data basis: © 2018 GADM. Projection: WGS84.

References

1. Ward, P.J.; Blauhut, V.; Bloemendaal, N.; Daniell, J.E.; de Ruiter, M.C.; Duncan, M.J.; Emberson, R.; Jenkins, S.F.; Kirschbaum, D.; Kunz, M.; et al. Natural hazard risk assessments at the global scale. *Nat. Hazards Earth Syst. Sci.* **2020**, *20*, 1069–1096. [CrossRef]
2. Mallakpour, I.; Villarini, G. The changing nature of flooding across the central United States. *Nat. Clim. Chang.* **2015**, *5*, 250–254. [CrossRef]
3. Westerling, A.L.; Hidalgo, H.G.; Cayan, D.R.; Swetnam, T.W. Warming and earlier spring increase western US forest wildfire activity. *Science* **2006**, *313*, 940–943. [CrossRef]
4. Songwathana, K. The relationship between natural disaster and economic development: A panel data analysis. *Procedia Eng.* **2018**, *212*, 1068–1074. [CrossRef]
5. Wisner, B.; Blaikie, P.; Cannon, T.; Davis, I. *At Risk: Natural Hazards, People's Vulnerability and Disasters*; Routledge: London, UK, 2014.
6. Murray, A.T.; Grubesic, T. *Critical Infrastructure: Reliability and Vulnerability*; Springer Science & Business Media: Berlin, Germany, 2007.
7. Lupp, G.; Zingraff-Hamed, A. Nature-Based Solutions—Concept, Evaluation, and Governance. *Sustainability* **2021**, *13*, 3012. [CrossRef]
8. Hilly, G.; Vojinovic, Z.; Weesakul, S.; Sanchez, A.; Hoang, D.N.; Djordjevic, S.; Chen, A.S.; Evans, B. Methodological framework for analysing cascading effects from flood events: The case of Sukhumvit Area, Bangkok, Thailand. *Water* **2018**, *10*, 81. [CrossRef]
9. Espada, R.; Apan, A.; McDougall, K. Vulnerability assessment of urban community and critical infrastructures for integrated flood risk management and climate adaptation strategies. *Int. J. Disaster Resil. Built Environ.* **2017**, *8*, 375–411. [CrossRef]
10. Negi, I.S.; Kumar, K.; Kathait, A.; Prasad, P. Cost assessment of losses due to recent reactivation of Kaliasaur landslide on National Highway 58 in Garhwal Himalaya. *Nat. Hazards* **2013**, *68*, 901–914. [CrossRef]
11. Azevedo, J.; Guerreiro, L.; Bento, R.; Lopes, M.; Proença, J. Seismic vulnerability of lifelines in the greater Lisbon area. *Bull. Earthq. Eng.* **2010**, *8*, 157–180. [CrossRef]
12. Bruneau, P.; Brangbour, E.; Marchand-Maillet, S.; Hostache, R.; Chini, M.; Pelich, R.M.; Matgen, P.; Tamisier, T. Measuring the impact of natural hazards with citizen science: The case of flooded area estimation using Twitter. *Remote Sens.* **2021**, *13*, 1153. [CrossRef]
13. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [CrossRef]
14. De Albuquerque, J.P.; Herfort, B.; Brenning, A.; Zipf, A. A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 667–689. [CrossRef]
15. Dittrich, A.; Lucas, C. Is this Twitter event a disaster? In Proceedings of the AGILE'2014 International Conference on Geographic Information Science, Connecting a Digital Europe through Location and Place, Castellon, Spain, 13–16 June 2014. AGILE Digital Editions.
16. Florath, J.; Chanussot, J.; Keller, S. Rapid natural hazard extent estimation from twitter data: Investigation for hurricane impact areas. *Nat. Hazards* **2024**. [CrossRef]
17. Forati, A.M.; Ghose, R. Examining Community Vulnerabilities through multi-scale geospatial analysis of social media activity during Hurricane Irma. *Int. J. Disaster Risk Reduct.* **2022**, *68*, 102701. [CrossRef]
18. Agarwal, A.; Toshniwal, D. Face off: Travel habits, road conditions and traffic city characteristics bared using twitter. *IEEE Access* **2019**, *7*, 66536–66552. [CrossRef]
19. Utomo, M.N.Y.; Adji, T.B.; Ardiyanto, I. Geolocation prediction in social media data using text analysis: A review. In Proceedings of the 2018 International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 6–7 March 2018; pp. 84–89.
20. Martínez, N.J.F.; Periñán-Pascual, C. Knowledge-based rules for the extraction of complex, fine-grained locative references from tweets. *RAEL Rev. Electrón. Lingüística Apl.* **2020**, *19*, 136–163.
21. Izbicki, M.; Papalexakis, V.; Tsotras, V. Geolocating Tweets in any Language at any Location. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; pp. 89–98.
22. Hu, X.; Zhou, Z.; Sun, Y.; Kersten, J.; Klan, F.; Fan, H.; Wiegmann, M. GazPNE2: A general place name extractor for microblogs fusing gazetteers and pretrained transformer models. *IEEE Internet Things J.* **2022**, *9*, 16259–16271. [CrossRef]
23. Yu, L.; Li, D. Road-Related Information Mining from Social Media Data: A Joint Relation Extraction and Entity Recognition Approach. *Buildings* **2022**, *13*, 104. [CrossRef]
24. Panasyuk, A.; Yu, E.S.L.; Mehrotra, K.G. Improving geocoding for city-level locations. In Proceedings of the 13th International Conference on Semantic Computing (ICSC), Newport Beach, CA, USA, 30 January–1 Febuary 2019; pp. 416–421.
25. Zhang, W.; Gelernter, J. Geocoding location expressions in Twitter messages: A preference learning method. *J. Spat. Inf. Sci.* **2014**, *2014*, 37–70.
26. Gu, Y.; Qian, Z.S.; Chen, F. From Twitter to detector: Real-time traffic incident detection using social media data. *Transp. Res. Part Emerg. Technol.* **2016**, *67*, 321–342. [CrossRef]
27. Zhang, J.; DeLucia, A.; Dredze, M. Changes in Tweet Geolocation over Time: A Study with Carmen 2.0. In Proceedings of the Eighth Workshop on Noisy User-generated Text (W-NUT 2022), Gyeongju, Republic of Korea, 12–17 October 2022; pp. 1–14.
28. Milusheva, S.; Marty, R.; Bedoya, G.; Williams, S.; Resor, E.; Legovini, A. Applying machine learning and geolocation techniques to social media data (Twitter) to develop a resource for urban planning. *PLoS ONE* **2021**, *16*, e0244317. [CrossRef]

29. Zia, M.; Fürle, J.; Ludwig, C.; Lautenbach, S.; Gumbrich, S.; Zipf, A. SocialMedia2Traffic: Derivation of Traffic Information from Social Media Data. *Isprs Int. J. Geo-Inf.* **2022**, *11*, 482. [CrossRef]
30. Putra, P.K.; Mahendra, R.; Budi, I. Traffic and road conditions monitoring system using extracted information from Twitter. *J. Big Data* **2022**, *9*, 65. [CrossRef]
31. Sakaki, T.; Matsuo, Y.; Yanagihara, T.; Chandrasiri, N.P.; Nawa, K. Real-time event extraction for driving information from social sensors. In Proceedings of the 2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), Bangkok, Thailand, 27–31 May 2012; pp. 221–226.
32. Ünsal, A.D. Detection and Description of Traffic Events Using Floating Car and Social Media Data. Ph.D. Thesis, Middle East Technical University, Ankara, Turkey, 2022.
33. Vallejos, S.; Alonso, D.G.; Caimmi, B.; Berdun, L.; Armentano, M.G.; Soria, Á. Mining social networks to detect traffic incidents. *Inf. Syst. Front.* **2021**, *23*, 115–134. [CrossRef]
34. Gelernter, J.; Balaji, S. An algorithm for local geoparsing of microtext. *GeoInformatica* **2013**, *17*, 635–667. [CrossRef]
35. Knoop, V.L.; Snelder, M.; van Zuylen, H.J.; Hoogendoorn, S.P. Link-level vulnerability indicators for real-world networks. *Transp. Res. Part Policy Pract.* **2012**, *46*, 843–854. [CrossRef]
36. Guth, J. Enhancing OpenStreetMap for the Assessment of Critical Road Infrastructure in a Disaster Context. Ph.D. Thesis, Karlsruhe Institute of Technology, Karlsruhe, Germany, 2021.
37. Chen, A.; Yang, H.; Lo, H.K.; Tang, W.H. Capacity reliability of a road network: An assessment methodology and numerical results. *Transp. Res. Part Methodol.* **2002**, *36*, 225–252. [CrossRef]
38. Liu, Z.; Chen, H.; Liu, E.; Hu, W. Exploring the resilience assessment framework of urban road network for sustainable cities. *Phys. Stat. Mech. Its Appl.* **2022**, *586*, 126465. [CrossRef]
39. Ientile, S.; Bianne, G.; Chevalier, C.; Schmidt, F.; Rasol, M.; Orcesi, A.; Adelaide, L.; Nedjar, B. Measuring road network resilience by loss of serviceability index for critical road links. In *Proceedings of the Institution of Civil Engineers-Bridge Engineering*; Thomas Telford Ltd.: London, UK, 2022; Volume 175, pp. 160–171.
40. Redzuan, A.A.H.; Zakaria, R.; Anuar, A.N.; Aminudin, E.; Yusof, N.M. Road Network Vulnerability Based on Diversion Routes to Reconnect Disrupted Road Segments. *Sustainability* **2022**, *14*, 2244. [CrossRef]
41. Rebello, K.; Jaggi, K.; Costello, S.; Blake, D.; Oo, M.; Hughes, J.; Egbelakin, T. Testing a criticality framework for road networks in Auckland, New Zealand. *Int. J. Disaster Resil. Built Environ.* **2019**, *10*, 36–51. [CrossRef]
42. Gibbons, S.; Lyytikäinen, T.; Overman, H.G.; Sanchis-Guarner, R. New road infrastructure: The effects on firms. *J. Urban Econ.* **2019**, *110*, 35–50. [CrossRef]
43. Glover, J.D.; Tennant, S.K. *Remote Areas Statistical Geography in Australia: Notes on the Accessibility/Remoteness Index for Australia (ARIA+ version)*; Public Health Information Development Unit, The University of Adelaide: Adelaide, Australia, 2003.
44. Dong, S.; Gao, X.; Mostafavi, A.; Gao, J.; Gangwal, U. Characterizing resilience of flood-disrupted dynamic transportation network through the lens of link reliability and stability. *Reliab. Eng. Syst. Saf.* **2023**, *232*, 109071. [CrossRef]
45. Liao, Y. Feasibility of estimating travel demand using geolocations of social media data. *Transportation* **2021**, *49*, 137–161. [CrossRef]
46. Ghader, S. A copula-based continuous cross-nested logit model for tour scheduling in activity-based travel demand models. *Transp. Res. Part Methodol.* **2021**, *145*, 324–341. [CrossRef]
47. Marshall, S. Line structure representation for road network analysis. *J. Transp. Land Use* **2016**, *9*, 29–64. [CrossRef]
48. Oberoi, K.S.; Del Mondo, G.; Dupuis, Y.; Vasseur, P. Spatial modeling of urban road traffic using graph theory. In Proceedings of the Spatial Analysis and GEOmatics (SAGEO) 2017, Rouen, France, 27 November 2017; pp. 264–277.
49. Sarkar, T.; Sarkar, D.; Mondal, P. Road network accessibility analysis using graph theory and GIS technology: A study of the villages of English Bazar Block, India. *Spat. Inf. Res.* **2021**, *29*, 405–415. [CrossRef]
50. Kondo, R.; Shiomi, Y.; Uno, N. Network evaluation based on connectivity reliability and accessibility. In *Network Reliability in Practice: Selected Papers from the Fourth International Symposium on Transportation Network Reliability*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 131–149.
51. Daniel, C.B.; Saravanan, S.; Mathew, S. Gis based road connectivity evaluation using graph theory. In *Transportation Research: Proceedings of CTRG 2017*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 213–226.
52. Erath, A.; Löchl, M.; Axhausen, K.W. Graph-theoretical analysis of the Swiss road and railway networks over time. *Netw. Spat. Econ.* **2009**, *9*, 379–400. [CrossRef]
53. Park, K.; Yilmaz, A. A social network analysis approach to analyze road networks. In Proceedings of the ASPRS Annual Conference, San Diego, CA, USA, 26–30 April 2010; pp. 1–6.
54. Kanyou, C.; Kouokam, E.; Emvudu, Y. Structural Network Analysis: Correlation between Centrality Measures. 2022. Available online: https://hal.science/hal-03714191/ (accessed on 1 March 2024).
55. Antunes, A.; Seco, A.; Pinto, N. An accessibility–maximization approach to road network planning. *Comput.-Aided Civ. Infrastruct. Eng.* **2003**, *18*, 224–240. [CrossRef]
56. Santos, B.F.; Antunes, A.P.; Miller, E.J. Interurban road network planning model with accessibility and robustness objectives. *Transp. Plan. Technol.* **2010**, *33*, 297–313. [CrossRef]
57. Chen, A.; Yang, C.; Kongsomsaksakul, S.; Lee, M. Network-based accessibility measures for vulnerability analysis of degradable transportation networks. *Netw. Spat. Econ.* **2007**, *7*, 241–256. [CrossRef]

58. Dijkstra, E.W. A note on two problems in connexion with graphs. *Numer. Math.* **1959**, *1*, 269–271. [CrossRef]
59. Li, X.; Zhou, J.; Pedrycz, W. Linking granular computing, big data and decision making: A case study in urban path planning. *Soft Comput.* **2020**, *24*, 7435–7450. [CrossRef]
60. Bader, R.; Dees, J.; Geisberger, R.; Sanders, P. Alternative route graphs in road networks. In *Proceedings of the Theory and Practice of Algorithms in (Computer) Systems: First International ICST Conference, TAPAS 2011, Rome, Italy, 18–20 April 2011*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 21–32.
61. Li, L.; Cheema, M.A.; Lu, H.; Ali, M.E.; Toosi, A.N. Comparing alternative route planning techniques: A comparative user study on Melbourne, Dhaka and Copenhagen road networks. *IEEE Trans. Knowl. Data Eng.* **2021**, *34*, 5552–5557. [CrossRef]
62. Chondrogiannis, T.; Bouros, P.; Gamper, J.; Leser, U.; Blumenthal, D.B. Finding k-shortest paths with limited overlap. *Vldb J.* **2020**, *29*, 1023–1047. [CrossRef]
63. Bai, S.; Ji, X.; Dai, B.; Pu, Y.; Qin, W. An Integrated Model for the Geohazard Accident Duration on a Regional Mountain Road Network Using Text Data. *Sustainability* **2022**, *14*, 12429. [CrossRef]
64. Keller, S.; Atzl, A. Mapping natural hazard impacts on road infrastructure—The extreme precipitation in Baden-Württemberg, Germany, June 2013. *Int. J. Disaster Risk Sci.* **2014**, *5*, 227–241. [CrossRef]
65. Alivand, M.; Hochmair, H. Extracting scenic routes from VGI data sources. In *Proceedings of the ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information*; Association for Computing Machinery: New York, NY, USA, 2013; pp. 23–30.
66. Manandhar, P.; Marpu, P.R.; Aung, Z.; Melgani, F. Towards automatic extraction and updating of VGI-based road networks using deep learning. *Remote Sens.* **2019**, *11*, 1012. [CrossRef]
67. Shrestha, E. Data Validation and Quality Assessment of Voluntary Geographic Information Road Network of Castellon for Emergency Route Planning. Ph.D. Thesis, Universitat Jaume I, Castellon de la Plana, Spain, 2019.
68. Aremu, B.D. Assessment of Volunteered Geographic Information (VGI) for Network Analysis in Emergency Response within the Health Sector. Ph.D. Thesis, Department of Surveying and Geoinformatics, Faculty of Environmental Design, Obafemi Awolowo University, Ife, Nigeria, 2020.
69. Tzavella, K.; Fekete, A.; Fiedrich, F. Opportunities provided by geographic information systems and volunteered geographic information for a timely emergency response during flood events in Cologne, Germany. *Nat. Hazards* **2018**, *91*, 29–57. [CrossRef]
70. Schnebele, E.; Oxendine, C.; Cervone, G.; Ferreira, C.M.; Waters, N. Using non-authoritative sources during emergencies in urban areas. In *Computational Approaches for Urban Environments*; Springer: Cham, Switzerland, 2015; pp. 337–361.
71. Schnebele, E.; Cervone, G.; Waters, N. Road assessment after flood events using non-authoritative data. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 1007–1015. [CrossRef]
72. White, D.R.; Borgatti, S.P. Betweenness centrality measures for directed graphs. *Soc. Netw.* **1994**, *16*, 335–346. [CrossRef]
73. Nicoară, P.S.; Haidu, I. A GIS based network analysis for the identification of shortest route access to emergency medical facilities. *Geogr. Tech.* **2014**, *9*, 60–67.
74. Tansley, G.; Schuurman, N.; Amram, O.; Yanchar, N. Spatial access to emergency services in low-and middle-income countries: A GIS-based analysis. *PLoS ONE* **2015**, *10*, e0141113. [CrossRef]
75. Di Cara, N.H.; Maggio, V.; Davis, O.S.; Haworth, C.M. Methodologies for monitoring mental health on Twitter: Systematic review. *J. Med. Internet Res.* **2023**, *25*, e42734. [CrossRef]
76. Müller, K.; Schwarz, C. From hashtag to hate crime: Twitter and antiminority sentiment. *Am. Econ. J. Appl. Econ.* **2023**, *15*, 270–312. [CrossRef]
77. Florath, J.; Keller, S. Supervised Machine Learning Approaches on Multispectral Remote Sensing Data for a Combined Detection of Fire and Burned Area. *Remote Sens.* **2022**, *14*, 657. [CrossRef]
78. Florath, J.; Chanussot, J.; Keller, S. Utilizing volunteered geographic information for real-time analysis of fire hazards: Investigating the potential of Twitter data in assessing the impacted areas. *Fire* **2023**, *7*, 6. [CrossRef]
79. Senaratne, H.; Mobasheri, A.; Ali, A.L.; Capineri, C.; Haklay, M. A review of volunteered geographic information quality assessment methods. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 139–167. [CrossRef]
80. Trudeau, R.J. *Introduction to Graph Theory*, 2nd ed.; Dover Publications: Mineola, NY, USA, 1993.
81. Boeing, G. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Comput. Environ. Urban Syst.* **2017**, *65*, 126–139. [CrossRef]
82. Boeing, G. Street network models and measures for every US City, county, urbanized area, census tract, and zillow-defined neighborhood. *Urban Sci.* **2019**, *3*, 28. [CrossRef]
83. Freeman, L.C. A set of measures of centrality based on betweenness. *Sociometry* **1977**, *40*, 35–41. [CrossRef]
84. Brandes, U. A faster algorithm for betweenness centrality. *J. Math. Sociol.* **2001**, *25*, 163–177. [CrossRef]
85. Freeman, L.C. Centrality in social networks: Conceptual clarification. In *Social Network: Critical Concepts in Sociology*; Routledge: London, UK, 2002; Volume 1, pp. 238–263.
86. Li, D.; Tang, Y.; Chen, Q. Multi-Mode Traffic Demand Analysis Based on Multi-Source Transportation Data. *IEEE Access* **2020**, *8*, 65005–65019. [CrossRef]
87. Yu, W.; Zhang, Y.; Ai, T.; Guan, Q.; Chen, Z.; Li, H. Road network generalization considering traffic flow patterns. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 119–149. [CrossRef]

88. Guth, J.; Wursthorn, S.; Braun, A.C.; Keller, S. Development of a generic concept to analyze the accessibility of emergency facilities in critical road infrastructure for disaster scenarios: exemplary application for the 2017 wildfires in Chile and Portugal. *Nat. Hazards* **2019**, *97*, 979–999. [CrossRef]
89. Knoop, V.; van Zuylen, H.; Hoogendoorn, S. The influence of spillback modelling when assessing consequences of blockings in a road network. *Eur. J. Transp. Infrastruct. Res.* **2008**, *8*. [CrossRef]
90. Du, M.; Jiang, X.; Chen, A. Identifying critical links using network capacity-based indicator in multi-modal transportation networks. *Transp. Transp. Dyn.* **2022**, *10*, 1126–1150. [CrossRef]
91. Vespaziani, L. *Best Practices in Wildfire Mass Evacuation Planning*; University of British Columbia, Graduating Project; UBC Library: Vancouver, BC, Canada, 2019.
92. District of Squamish and Resort Municipality of Whistler. Sea to Sky Evacuation Plan. 2019. Available online: https://www.whistler.ca/services/emergency/evacuating-whistler/sea-sky-evacuation-plan/ (accessed on 7 September 2023).
93. Florath, J.; Chanussot, J.; Keller, S. Investigation of Volunteered Geographic Information Data for Near-Real-Time Spatial Analysis of Fire Hazards. In Proceedings of the 9th International Conference on Geographical Information Systems Theory, Applications and Management GISTAM, Prague, Czech Republic, 25–27 April 2023.