Spatial and temporal evolution of post-disaster data for damage assessment of civil infrastructure systems

Jorge Mario Lozano, Elliot Nichols, J. David Frost & Iris Tien

Published online: 25 Aug 2023.

Submit your article to this journal ⬈

View related articles ⬈

View Crossmark data ⬈

Taylor & Francis
Taylor & Francis Group

Spatial and temporal evolution of post-disaster data for damage assessment of civil infrastructure systems

Jorge Mario Lozano 🔟, Elliot Nichols, J. David Frost and Iris Tien 🔟

School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT

Assessing damage to civil infrastructure is a resource-intensive process that is critical during the response to a disaster. Various datasets facilitate this process but are often collected on an individual ad hoc basis by multiple separate entities. Consequently, there is a lack of a coordinated approach when collecting disaster data, which prevents effective data interoperability. Rather than viewing datasets individually, this paper provides a comprehensive analysis of post-disaster damage data to demonstrate the merits of a dynamic data collection process accounting for both spatial and temporal variations. Specifically, datasets from Hurricane Maria and the Indios Earthquake in Puerto Rico are used to illustrate the entities involved, resources used, and resulting datasets for this purpose. The paper analyzes the evolution of key metadata features as a function of time, including data availability, coverage, and resolution. The results show distinct stages of the data collection process and reveal challenges in collaboration between entities and a lack of data integration for disaster response. The findings also lead to recommendations about the essential metadata for increased shareability. With these outcomes, entities in the field can improve the quality of information extracted and facilitate interoperability and information integration across datasets for damage assessment.

1. Introduction

In the aftermath of a disaster, estimating physical damage to civil infrastructure systems is critical. Infrastructure systems are crucial for community recovery and, when functional, enable other aspects of society to operate effectively. Unfortunately, infrastructure systems are highly vulnerable to natural hazards, resulting in large and widespread impacts (Vamvatsikos et al. 2010; Hosseini Nourzad and Pradhan 2016; Willis et al. 2016). The socioeconomic status of many developing regions makes this problem even more critical (Nuti and Vanzi 2003; Zhou et al. 2010).

Recent technological advancements have increased the number of data tools available for use in post-disaster damage assessment. Data collection tools, such as drones and sensor networks are increasingly prevalent during the response to recent disasters (Akter and Wamba 2019). For example, Synthetic Aperture Radars (SAR) mounted on satellites started to be used by NASA for rapid damage assessment after the 2015 Gorkha earthquake (Yun et al. 2015). Additionally, Streetview camera systems are now being used for rapidly cataloging damaged and undamaged buildings after disasters, as was done after Hurricane Michael in 2018 (Roueche et al. 2018; Berman et al. 2020). These advances provide multiple benefits for emergency response operations, policy formulation, engineering research, and reconstruction purposes. However, the lack of coordination in data collection and publication practices brings multiple limitations to the process. Table 1 lists the main benefits and limitations of post-disaster damage data.

Knowing the location of infrastructure damage with high fidelity is paramount for activating appropriate emergency response protocols and prioritizing infrastructure recovery activities. For example, emergency managers use building damage information to determine which structures to evacuate (Liu et al. 2021). In the case of transportation networks, assessing which roads are closed improves the assignment of evacuation routes (Urbina and Wolshon 2003). Other critical tasks include identifying failures in pipelines and power networks to facilitate the restoration of essential services.

Post-disaster damage assessment, however, is extremely laborious and time-consuming, especially for local emergency managers who may work with limited budgets. As a result, damage data for large areas affected by a disaster is usually incomplete. When performing detailed damage assessments, these entities also have responsibilities associated with relocating affected communities and guaranteeing medical supplies for those in need, amongst others (McEntire 2007). This makes damage assessment one task among many for emergency responders. Since most regions work on restricted budgets, simply allocating more personnel or financial resources is not a viable solution for agencies to obtain a clearer and more detailed picture of post-disaster damage.

In addition to emergency managers, multiple entities participate during the disaster response period to catalog damage to buildings and lifelines. In the U.S., federal agencies, for instance, require this data to validate damage and distribute funds to local governments (FEMA 2021). Other international agencies include NGOs, such as the

Table 1. Benefits and limitations of post-disaster damage data.

Benefits	Limitations
- Paramount for activating emergency response protocols, such as rescue, recovery, and evacuations - Determines prioritization of assets for infrastructure recovery - Useful for researchers, construction companies, and public agencies to improve design-code practices and policy formulations - Enables affected communities and insurance companies to file and review insurance claims, respectively	- Detailed damage descriptions are not readily available for immediate emergency needs - Data is usually incomplete for large regions because it is time and resource expensive to collect - There are no practical ways to quantify uncertainty - There is limited coordination in data collection, creating inconsistent metadata and a lack of shareability - Data is usually studied in a single point of time, ignoring the dynamic nature of data collection

United Nations or the World Bank, which help regions that have poor access to resources (The European Commission et al. 2013). Researchers collect and analyze damage data for weeks after the event (Bray et al. 2018), to support activities, such as improving design codes and creating updated damage fragility functions (Shinozuka et al. 2000; Del Gaudio et al. 2017; Gautam 2018; Gautam et al. 2018; Rajapaksha and Siriwardana 2023). Detailed damage descriptions that may eventually become available as more data is collected, however, are not readily available in the immediate aftermath of a disaster to support emergency response, rescue, and recovery operations.

The diversity of parties involved in the damage assessment process results in data that vary in coverage, resolution, and format. For example, damage related to buildings and geotechnical structures can be retrieved from satellite imagery by federal agencies (Hong et al. 2006; Adams et al. 2013) or from on-the-ground field surveys performed by academic experts (O'Rourke and Toprak 1997; Bray et al. 2019; Galvis et al. 2020). The methods to process and catalog these data also differ depending on the region and user. For example, there are different guidelines for assessing building damage, with different scales, such as the U.S. Federal Emergency Management Agency (FEMA) damage assessment guideline (FEMA 2021) and the European Macroseismic Scale (EMS-98). These different methods make it difficult to have consistent comparisons across disasters because the same asset (building, road, pipeline) is not cataloged using the same process or with the same scale. In addition, incomplete information and the lack of consistency across damage datasets prevents information from being effectively integrated across datasets and reduce the ability to rigorously quantify uncertainty in damage estimates.

Different data users also require different levels of data quality. It is noted that data quality is related to the purpose that it is collected for and is not just a measure of whether a sensor is calibrated or not. For example, Landsat imagery can be of high quality, but it only captures data at 30 m resolution (USGS 2018). Depending on the intended utilization, this could be classified anywhere between high quality and low quality. For instance, most public entities measure damage with large-scale damage scales (such as the one from FEMA previously mentioned). This damage data, however, is not useful for supporting detailed assessments of damage, such as determining the root cause of damage to a building or specifics about the failure mechanism of infrastructure components. This detailed information is critical for researchers and practitioners interested in code design and policy formulation. As a result, the disparity in requirements complicates the possibility of having a standard for collecting damage data, especially when, in addition to collecting data of different quality, most entities collect data at different times. For example, depending on team coordination and the severity of the event, while emergency managers collect information hours after the event, academic teams can arrive to the field on the order of 1 week to 1 month after (see arrival of the Geotechnical Extreme Event Reconnaissance (GEER) team 6 days after the Chile earthquake in 2010, and 26 days after for the Western European floods in 2021 (Bray and Frost 2010; Lemnitzer 2022).

Moreover, the variety of data collection methods and purposes across entities creates an inconsistent publication of metadata given there is no coordination when

collecting and analyzing damage data. For example, essential information, such as collection time, location, and contact information, among others, is not always published with every dataset. This inconsistency results in a lack of shareability between entities, reducing opportunities to leverage the outcomes from one dataset in the collection of another and limiting the ability to integrate information across datasets.

Finally, another limitation to post-disaster data is that most studies have looked at these datasets in isolation (e.g. Youd et al. 2000; Bray and Frost 2010) rather than reviewing them together in an integrated manner to provide a comprehensive cumulative view of post-disaster damage. Post-disaster states are often assessed at a single point in time rather than longitudinally to provide a temporal assessment of the evolution of data availability over time. One reason for this state of practice is that most studies focusing on infrastructure resilience tackle the problem either before the damage assessment, i.e. optimizing mitigation strategies (Abruzzo et al. 2006; Costa et al. 2010; Yildirim and Demir 2021), or shortly thereafter, i.e. prioritizing recovery tasks (González et al. 2016; Rouhanizadeh et al. 2020; Ghannad et al. 2021). The former includes major uncertainty on the expected damage to structures and lifelines. The latter usually excludes the time it takes to collect the data and therefore assumes the damage assessment as a static process.

All these limitations in data collection and publication result in a limited understanding of infrastructure damage across a community after a disaster event. Thus, there is a need to fully recognize and analyze the damage assessment as a time-varying process—a series of tasks and systems with incomplete information and estimation that evolves during the response to a disaster.

To address these challenges, this paper illustrates the dynamic nature of the damage assessment process for infrastructure systems by studying the variety of datasets that are compiled after a disaster, including the motivations and needs of multiple entities for data throughout the post-disaster time period, and the resulting characteristics of varying datasets over time. Datasets from two disaster events in Puerto Rico are included in this study—Hurricane Maria and the Indios Earthquake. These events represent different types of disaster events, as well as different levels of severity, enabling comparison of post-disaster datasets by disaster type and impact level. A detailed metadata analysis for all publicly available datasets is conducted, where multiple data characteristics (including collection time, publication time, coverage, and resolution) are evaluated to determine the evolution of data characteristics in the response to a disaster. The analysis addresses several of the current limitations in post-disaster data. The data trends allow for the determination of critical gaps in data collection and the essential metadata features needed for damage assessment of infrastructure systems. The gaps and results found in this paper lead to recommendations to increase the shareability and integration of damage datasets.

While multiple systems exhibit these dynamic characteristics, this study will focus on physical damage to structures and infrastructure systems that impact a community, including buildings, lifelines (e.g. electric power, communications, water, transportation), and geo-structures (e.g. natural slopes, culverts, embankments). This paper, for the first time, treats post-disaster data collection as a time-varying process, assessing both spatial and temporal evolutions of post-disaster datasets with high

fidelity. Given the specificity used in the analysis, this work exposes key limitations in existing post-disaster data collection and publication practices, e.g. in the metadata (e.g. release date, resolution, and coverage) that aggravate the data ingestion and utilization processes.

While this study will focus on two major disaster events in Puerto Rico, the outcomes are generally applicable to other events and locations. Realistically, not every country has the same level of resources to produce the same amount of data. However, the results from this analysis provide insights into how to improve the data collection process and develop better practices to share datasets within communities across locations and events. Implementing these strategies is critical for overcoming current challenges in the assessment of damage across large regions after disasters and the difficulties of local governments in allocating sufficient resources for this purpose.

This paper is organized as follows. Section 2 describes the datasets for each disaster event in Puerto Rico, emphasizing key data features, such as resolution and coverage. Next, Section 3 analyzes the time evolution of datasets, including the variations in data publication times, and how the coverage and resolution change throughout the disaster response period. Quantitative measures are provided to enable comparisons across multiple datasets and hazard events. Then, Section 4 provides an overview of the different stages of post-disaster data collection divided into three stages identified with the metadata analysis. It also presents the timeline of the data collection process per stage, making emphasis on the availability of datasets and changing needs of several entities. Finally, Section 5 describes the main challenges of the damage assessment process, and Section 6 proposes a set of metadata features for improving data shareability and decreasing uncertainty in data collection for post-disaster damage assessment.

2. Post-disaster data availability and descriptions

Given how available data changes over time and for different purposes, it is of interest to conduct a systematic and comprehensive analysis of the evolution of available datasets. To conduct this dynamic analysis of post-disaster data and illustrate the spatio-temporal evolution of datasets, this section focuses on data from two major extreme events in Puerto Rico: Hurricane Maria and the Indios Earthquake. Puerto Rico's geographic location within the Caribbean, along with factors, such as the large number of households located in high-risk zones, high unemployment rates, and low per capita income, make the island highly vulnerable to natural hazards. In terms of hurricanes, Puerto Rico is located in an area that is prone to multiple dangerous storms per year, and the island is directly hit by a moderate or greater hurricane every 5 years on average (Boose et al. 2004). More severe storms as a result of global warming are expected in the future generating larger impacts on the island (Hall et al. 2020). Earthquakes are also present in the region. Located close to the northeastern corner of the Caribbean seismic zone, specifically near the Sombrero seismic zone, Puerto Rico has a long history of severe earthquakes (McCann 1985; Mueller et al. 2010; Meighan et al. 2013). Therefore, the occurrence of recent events, such as

the category 5 Hurricane Maria and the magnitude $6.4\,M_w$ Indios Earthquake presents an opportunity to investigate the state of publicly available data for damage assessment under varying disaster scenarios. As Puerto Rico is a U.S. territory, different emergency management entities are involved at different levels, e.g. the Puerto Rico Emergency Management Agency (PREMA) and the Federal Emergency Management Agency (FEMA). Having multiple institutions working in response to different disasters provides an opportunity to analyze available datasets and make comparisons between disasters. This section presents the public datasets used for damage assessment after Hurricane Maria in 2017, as well as the Indios Earthquake in 2020. While crowdsourced data specific to the purpose of damage assessment is included, general social media data (e.g. from Twitter or Facebook) was not considered given that there are no current systematic processes used in practice to catalog data from community sources. Notwithstanding this, there are current research efforts being implemented for this purpose, see Tien et al. (2016), Resch et al. (2018), Wu and Cui (2018), and Hao and Wang (2020).

The analysis presented herein treats the two events as independent events. In reality, there is likely to be some relationship as the result of the relatively close proximity of the events temporally. For example, a reconnaissance mission to Puerto Rico in 2022 after Hurricane Fiona showed that some of the temporary bridges and blue tarps were still being used five years after Maria (Morales et al. 2022). These relationships, however, do not represent major differences in the damage assessment, since each event affected different parts of the island, minimizing this effect. A metadata analysis was conducted on each dataset to investigate multiple data features (e.g. coverage, resolution) that enable evaluations of the evolution of the data and comparisons across datasets that are analyzed later in Section 3.

2.1. Hurricane Maria, 2017

During the hurricane season of 2017, Puerto Rico was struck by two major hurricanes: Hurricane Irma on 7 September and Hurricane Maria on 20 September, the latter considered a category 4 hurricane when passing through the island. The first hurricane did not cross through the mainland but did cause major power outages and water service interruptions, with sustained tropical storm force winds reaching sustained wind speeds of 89 km/h and wind gusts up to 179 km/h. Three indirect deaths occurred in Puerto Rico from Hurricane Irma (Cangialosi et al. 2018). Fifteen days after Irma, Hurricane Maria made landfall in Puerto Rico with the eye passing only 40 kilometers from the capital city of San Juan, resulting in devastating damages as a result of floods, sustained wind speeds of 103 km/h, and wind gusts up to 225 km/h. The NOAA estimate of damage in Puerto Rico and the U.S. Virgin Islands due to Maria is around $90 billion (Pasch et al. 2019). There is no reliable estimate of the death toll for this event; however, estimates in the range of several thousand have been presented. Maria knocked down 80% of the utility poles and essentially all transmission lines, resulting in power losses to practically all residents on the island (Kwasinski et al. 2019). From Table 2, the range of data types and characteristics available at different time periods can be seen. The publication time after the event is

Table 2. Dataset descriptions for Hurricane Maria.

Dataset	Entity	Description	Data type	Collection start date	First publication date	Publication time after the event	Spatial resolution	Coverage
Rainfall	NASA GPM	Interpolated map of estimated rain on the day of the event	Raster	9/20/17	9/21/17	12 h, 3 months	10 km	Complete
Wind speeds	ARA	Wind speed profile developed using meteorological station throughout Puerto Rico	Raster	9/20/17	9/21/17	1, 5, and 7 days	500 m	Complete
Cellphone coverage	FCC	Communications status reports of the percentage of cell towers per municipality that were not functional after the event	Table	9/21/17	09/21/17	1 day–6 months (daily update)	By county	Complete
Damage proxy map	NASA JPL	Damage estimation map created with pre- and post-disaster satellite images.	Raster	9/21/17	9/22/17	2, 9 days	15 m	Complete
Routes blocked	USAR	A snapshot of the state of the transportation network six days after the event	Point	9/21/17	9/26/17	6 days	By asset	Complete
Flood detection maps	FEMA/NLT	Spatio-temporal data integration approach to provide a single, comprehensive and accurate measure of hurricane-related flooding	Raster	9/21/17	10/4/17	2 weeks	15 m	Substantial
Building flood damage	FEMA	Geospatial damage assessment map with buildings that have clear evidence of flood water (either destroyed or affected)	Point	9/22/17	9/27/17	7 days	By asset	Substantial
Building wind damage	FEMA/Newbury	Geospatial damage assessments for certain impacted areas using NOAA Imagery and Modeled Preliminary Observation Depth Grids	Point	9/22/17	10/17/17	27 days	By asset	Completes over 1 month
Crowdsourced portal	Crowd-sourceHQ	Displays crowdsourced reports about closed roads, collapsed bridges, flooding, mudslides, areas where communication is unavailable, areas that don't have power, areas without drinking water, and areas with downed power lines	Point	9/24/17	9/27/17	7 days–1 year (daily update)	By asset	Moderate over 1 year

(continued)

Table 2. Continued.

Dataset	Entity	Description	Data type	Collection start date	First publication date	Publication time after the event	Spatial resolution	Coverage
Landslide density	USGS	Visual estimation of landslide concentration, with a combined study area of 8103 square km	Polygon	9/26/17	10/24/17	5 weeks	2 km	Complete
Slope failures	USGS	Detailed locations of slope failures across the island. Provided by USGS using imagery from NASA, FEMA, NOAA, and Digital Globe	Point	9/26/17	8/21/19	1 year 11 months	By asset	Complete
Nightlights	NASA	Combined nightlight data with Landsat and OpenStreetMap data to monitor electric power. Acquired with the Visible Infrared Imaging Radiometer Suite (VIIRS)	Raster	9/27/17	9/28/17	7–28 days (daily update)	12 m	Substantial
Bridge status	NAPSG	Map depicting the status of Puerto Rico's bridge network. Status include: open, closed, and unknown	Point	9/29/17	10/02/17	13 days–6 weeks	By asset	Substantial

given as a single time if published all at once. Otherwise, it is given as multiple times or with a range, to reflect updates to the dataset or the dataset being published in parts, respectively.

The coverage of each dataset is represented by four coverage levels: Minimal, which includes only a small portion of the affected area (e.g. a single neighborhood); Moderate, which focuses only on areas of high damage (e.g. the most affected county); Substantial, which covers most of the affected region but misses some areas due to the reconnaissance route or the tool coverage (e.g. satellite image that does not cover all the region); and Complete, which means that all of the impacted areas are included in the dataset (e.g. cellphone coverage for the whole island of Puerto Rico).

In terms of data collection for damage assessment, the most common datasets identified one day after the event were estimations of parameters, such as the strength of winds and broad estimations of damage. For Hurricane Maria, the National Hurricane Center released a hurricane path, detailing the location of the eye at different times (NOAA 2017). In terms of wind speed, Applied Research Associates, Inc. (ARA) generated a grid with the peak wind speeds across the island (ARA Inc. 2017). Descriptions of all datasets that were publicly available for this disaster are given in Table 2. This table shows the main attributes found for each dataset, including the dataset description, entity releasing the data, data types, temporal characteristics, such as the time of data collection and data release, and data characteristics of spatial resolution and coverage.

One week after the event, some datasets, such as the wind speeds, were refined based on new information or additional data processing. ARA released multiple versions of this dataset, decreasing the uncertainty by including more information from additional sensors and sources. Remote sensing data that requires more post-processing was also released during this period. For example, night-time light data from NASA captured light from buildings and streets to compare the levels of radiance with pre-event values (Zhao et al. 2020). This stage also included some field work to estimate detailed damage in the region. FEMA began damage assessment due to flooding and wind at this time for the metropolitan areas of Puerto Rico. In the same way, crowdsourced data became available online through CrowdSourceHQ, for the community to support the damage assessment by uploading information about the state of different lifelines and infrastructure.

In the subsequent weeks up to a month after the event, more detailed published data became available. For instance, USGS published a map depicting areas by landslide density (Bessette-Kirton et al. 2017). FEMA released a broader coverage of both the wind and flood damage. FEMA also released an estimation of flooding areas by combining satellite data with an estimation of flood depth. Interestingly, even years after the event, there are still some datasets being released that can be used toward improved damage assessment of the event. For instance, USGS updated the landslide map and published a map in 2019 with the locations of all slope failures in Puerto Rico (Hughes et al. 2019).

In addition to the varying data availability at different time periods after an event, datasets also varied by the spatial resolution and coverage they provided. Figure 1
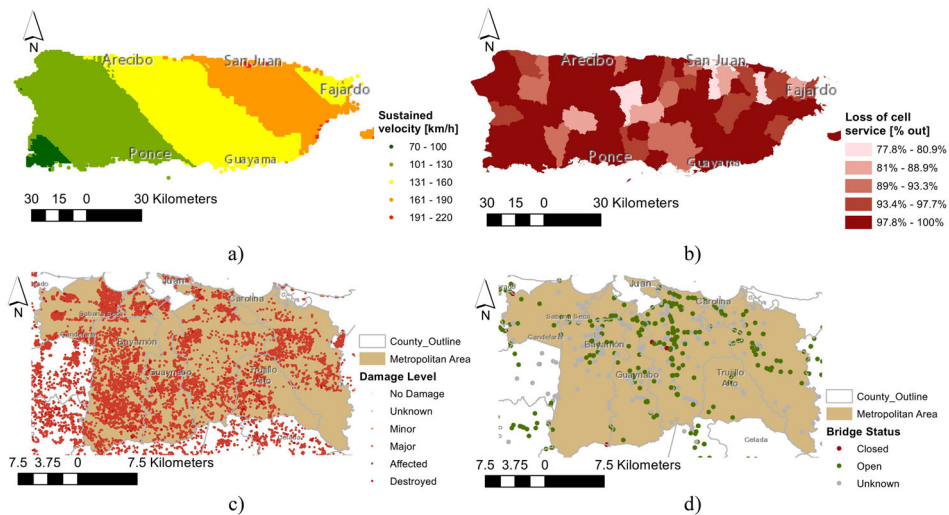
Figure 1. Different datasets collected after Hurricane Maria (2017) with varying spatial resolution and coverage: (a) wind profile from ARA, (b) cell towers out of power on 21 September 2017 (FCC), (c) FEMA wind damage in the San Juan metropolitan area, and (d) bridge status in the San Juan metropolitan area of Puerto Rico.

shows four example datasets from Hurricane Maria. Wind profile data is at 500 m resolution covering the whole island (Figure 1(a)), cell phone tower outage information is provided by the county (Figure 1(b)), and the two lower figures show detailed point data (Figures 1(c,d)) for the San Juan metropolitan area. The first point dataset represents the individual building wind damage, which was produced by FEMA using satellite images from NOAA. The second point dataset shows bridge locations along with their status information (i.e. closed, open, unknown), which were updated as new reports became available.

2.2. Indios Earthquake, 2020

The magnitude 6.4 M_w earthquake that struck due west of Ponce, Puerto Rico, on 7 January 2020, resulted in few casualties but an estimated $3.1 billion worth of damage (USGS 2018). The devastation from this event was due to the magnitude of the event itself, combined with the circumstances in which it occurred. Twenty-eight months before the earthquake, Puerto Rico had been impacted by Hurricanes Irma and Maria, as noted above, and the country was still recovering from these disasters at the time of the earthquake in 2020. The perishable data recovered from this event gives insights into the data collection process across hazard events in the same location. This section highlights the datasets collected after the earthquake and the similarities and differences in the types and characteristics of data compared to what was collected after Hurricane Maria. Table 3 shows the datasets available for damage assessment after the Indios Earthquake, including dataset characteristics.

In the day following the earthquake, intensity measures of the earthquake comprised the majority of the data, which is similar to early data from Hurricane Maria. The World Food Programme estimated the number of people per county exposed to

Table 3. Dataset descriptions for Indios Earthquake.

Dataset	Entity	Description	Data type	Collection start date	First publication date	Publication time after the event	Spatial resolution	Coverage
ShakeMap	USGS	USGS ShakeMap showing PGA, among multiple earthquake parameters	Raster	1/7/20	1/7/20	20 min–1 month (10 improved versions)	1000 m	Complete
Loss assessment	Hazus	Estimation of total economic impact on each barrio.	Polygon	1/7/20	1/7/20	3 h–8 days (1 improved version)	By barrio	Complete
ADAM population exposure	World Food Programme (WFP)	Number of people per county exposed using USGS ShakeMap	Table	1/7/20	1/7/20	3 h–8 days (1 improved version)	By county	Complete
Cellphone Coverage	FCC	Communication status reports of the percentage of cell towers per municipality that were not functional after the event.	Table	1/8/20	1/8/20	1–6 days (daily update)	By county	Complete
Nightlights	NASA	Combined nightlight data with OpenStreetmap data to monitor electric power. Acquired with the Visible Infrared Imaging Radiometer Suite (VIIRS).	Raster	1/8/20	1/9/20	1–17 days (daily update)	710 m	Substantial
Co-seismic landslides	NASA	Outlines of co-seismic landslides. Derived from imagery by Geo-Eye 1, WorldView-2, and WorldView-3 satellites	Point	1/8/20	1/15/20	8 days (partial online publication)	By asset level	Substantial
Damage Proxy Map	NASA ARIA JPL	Damage estimation map from pre- and post-radar imagery courtesy of Copernicus Sentinel-1.	Raster	1/9/20	1/10/20	3, 7, 19 days	30 m	Complete
Surface displacement	NASA ARIA JPL	Ground movement as a result of the event. Pre- and post-event radar imagery from Copernicus Sentinel-1	Raster	1/9/20	1/11/20	4, 7, 11, 19, 43 days	30 m	Complete
Building damage	FEMA	Individual assessments of building damage.	Point	1/11/20	1/14/20	7 days	By asset	Completes after 1 Week
Co-seismic landslides	USGS	Inventory of landslides from reconnaissance field trip and pre-event LIDAR data	Point	1/12/20	6/19/20	5 months	By asset	Complete

each shaking level intensity defined by the USGS ShakeMap, and the maximum peak ground acceleration (PGA) for each barrio was recorded. Using Hazus, preliminary estimations of economic loss for each barrio were generated. Following the initial publication of data, communication status updates *via* the Federal Communications Commission (FCC) were published for each county for the next week. Once the event struck, satellites began to position themselves to record data. Given the time it took to capture images and perform necessary post-processing, many more types of data were collected and then published in the days and the week following the earthquake.

One day after the event, NASA began the automatic process of publishing data that quantified the amount of light emitted from buildings in Puerto Rico daily. At the same time, ARIA JPL was using images from the Copernicus-Sentinel-1 satellite to locate areas of likely damage by comparing pre-and post-event imagery. The resulting Damage Proxy Map was updated twice after its initial publication on 9 January 2020, using similar imagery from 14 January and 26 January. Using the synthetic aperture radar (SAR) imagery from 9 January and 14 January, ARIA JPL also derived corresponding surface displacement maps for each date. FEMA used low-altitude airborne imagery as well as satellite imagery, collected over the first week after the event, to assign damage categories to buildings by inspecting the images manually, which resulted in a Damage Map that was updated every day starting one week after the event.

In the second week after the event, GIS specialists were able to derive the locations of co-seismic landslides in the mountainous regions of Puerto Rico using several satellites positioned over the island directly after the event. Beyond two weeks after the mainshock, only one dataset was published that pertained directly to the earthquake damage, as much of the focus turned to measuring humanitarian impact. An updated ARIA Damage Proxy Map was published 3 weeks after the event to measure how damaged areas had evolved, and similar to Hurricane Maria, the last published dataset was a description of landslide locations for the event, released five months after the earthquake, which marked the end of open-access and available maps and data.

Figure 2 shows example datasets from this event, including the initial Damage Proxy Map (Figure 2(a)), hazard assessment by PGA (Figure 2(b)), county-level cell phone outage information (Figure 2(c)), and point level building damage information (Figure 2(d)). As in the Hurricane Maria case, each dataset varies by several attributes, including the spatial resolution and coverage over the area that they provide.

3. Analysis of temporal and spatial evolution of datasets

The evolution of post-disaster data is studied in this paper with a detailed analysis of key metadata features that are highly dynamic during the response to a disaster, including publication time, spatial resolution, and coverage. The analysis shows how available data changes over time, and how the characteristics of post-disaster datasets that can be used to assess the extent of damage evolve in the post-disaster period. The rest of this section is organized into two sections: Section 3.1 describes the evolution of time features of data collection. In contrast, Section 3.2 analyzes the compound evolution of resolution and coverage over the response period.
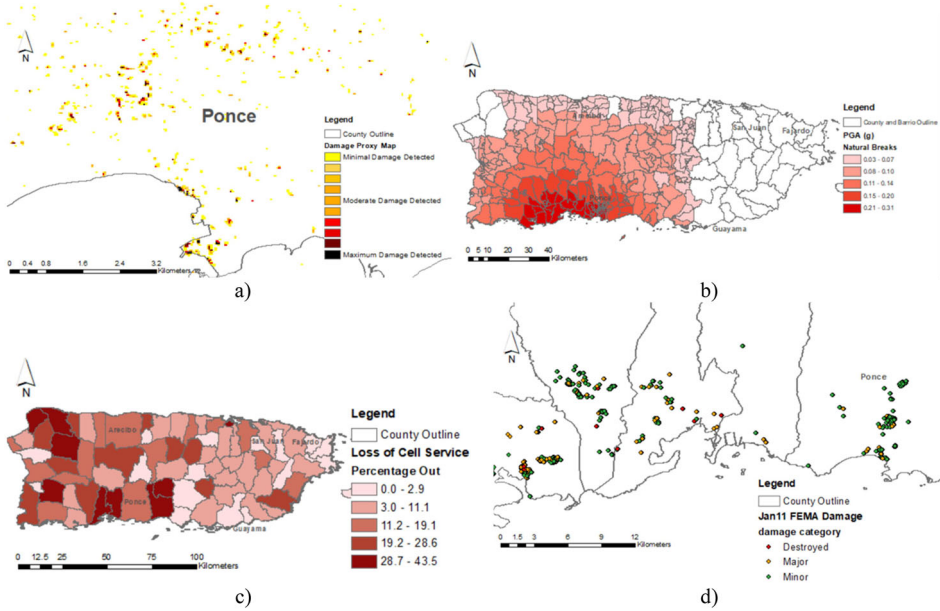
Figure 2. Different datasets collected after the Indios Earthquake (2020) with varying spatial resolution and coverage: (a) Damage Proxy Map from 9 January 2020 (ARIA JPL), (b) peak ground acceleration (USGS), (c) cellphone coverage on 9 January 2020 (FCC), and (d) FEMA building damage.

3.1. Time evolution of data

The temporal evolution of post-disaster data availability consists of two major elements: the time when data collection starts and the time when the data is first published (or made publicly available). Given the diversity of data tools and methodologies used in practice, the effort and time it takes to collect and publish a dataset can vary considerably. Quantifying the time element of dataset collection and publication is critical for its use in post-disaster damage assessment activities. It also supports recommendations for improving the data collection by leveraging and integrating multiple datasets in damage assessment.

Looking first at publication time, understanding the expected time range after the event occurs for a dataset to be published is key information for multiple entities. For instance, emergency managers are interested in the datasets released in the first couple of days to understand the damage in the whole region (Mukhopadhyay and Bhattacherjee 2015). On the other hand, researchers looking at infrastructure recovery are more interested in detailed descriptions of damaged facilities, information that can take months to be analyzed (Bray et al. 2018).

Considering the time when the data was collected is also important to fully understand the time-dependent elements of a dataset. For example, comparing the collection and publication times provides an understanding of the complexity of implementing a dataset. The longer it takes to publish a dataset from the collection start date, is likely reflective of either data collection or processing requirements. Moreover, during the response stage of a disaster, the performance of an
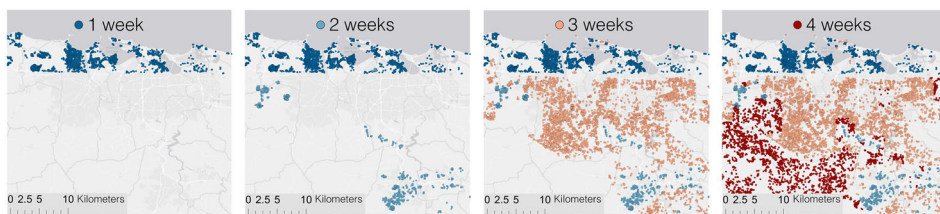
Figure 3. Data collection time evolution for building wind damage from Hurricane Maria. Points represent individual buildings colored by the week of collection time.

infrastructure system varies, where some regions continue to be affected by aftershocks in an earthquake, and some parts of the system are actively being restored. Considering the time of data collection is critical to creating a snapshot of the system's performance after the disaster, with the understanding that the condition is continually changing after a disaster. Power networks, for example, are known to be largely restored in days or weeks, compared to building infrastructure, which requires more time to become fully functional. The time-dependent element of datasets is shown in Figure 3, where data for building wind damage collected by FEMA from Hurricane Maria took almost a month to collect and process. This dataset was produced by visually inspecting satellite images by NOAA and assigning a damage level, as shown in Figure 1(c). Figure 3 displays distinctive regions for data collection per week and the increase in coverage as time progressed.

The evolution of some datasets can also be seen with refinements to previously published information. This relates to datasets that are released multiple times without changing their coverage or resolution. However, the base information of the dataset can change in some cases, such as the values of rainfall during Hurricane Maria released by NASA both 12 h after the event and 3 months after (Bell et al. 2018). The refined version of the later dataset typically has better estimations with less uncertainty. Figure 4 depicts two cases of updating datasets that do not change in format, coverage, or resolution, but the metric [i.e. wind speed or Peak Ground Acceleration (PGA)] changes due to increasing amounts of data and improvements in the estimation. This effect demonstrates the importance of utilizing the latest version of a dataset and acknowledging the uncertainty in some parameters in early data releases during the response stage.

The evolution of metadata time features can be visualized in a timeline of the response stage after the disaster. Figures 5 and 6 show the metadata features of all infrastructure damage datasets found for Hurricane Maria and the Indios Earthquake, respectively. Datasets are sorted by the data collection start time. These figures illustrate the evolution of the data collection start time (shown as green triangles) and the date when the dataset is first published (shown as blue dots). Also included are three types of data publication. For some datasets, partial data, shown as small blue circles, is published at periodic intervals indicating increasing data collection in terms of coverage or changes to the system over time (e.g. the increasing coverage of building wind damage shown in Figure 3). Next, is the date when data is updated (large blue circles), referring to information that does not change in coverage or resolution but is published repeatedly throughout the response to the disaster. One good example
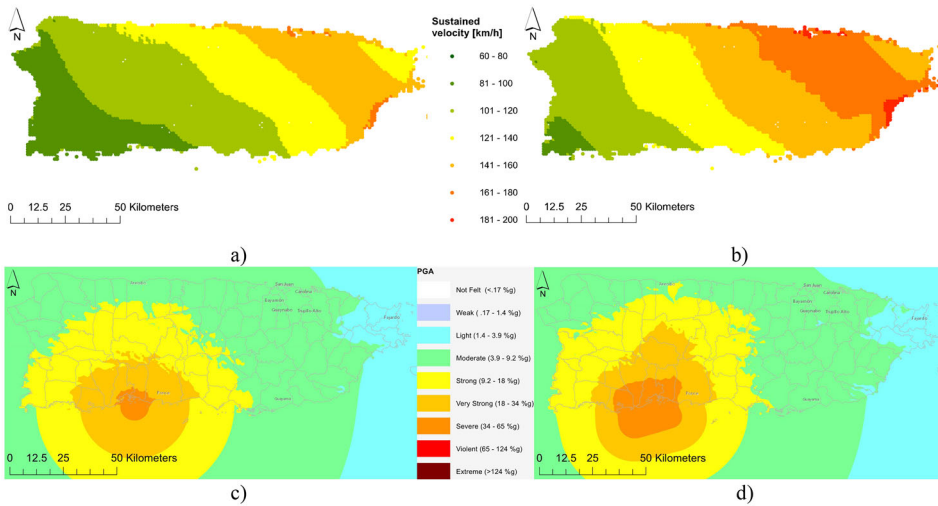
Figure 4. Datasets with refined features. Top: Hurricane Maria sustained wind speeds published (a) one day after event and (b) 7 days after event. Bottom: Indios Earthquake PGA published (c) 20 min after event and (d) 7 days after event.
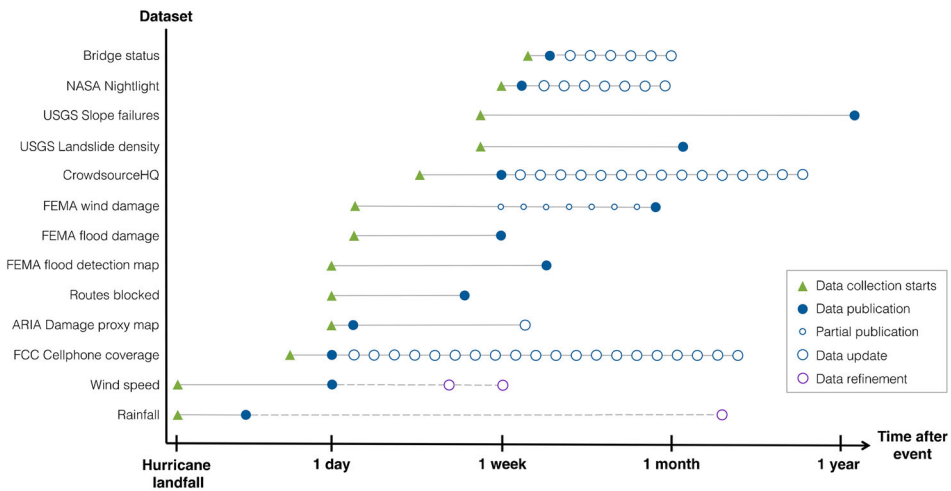

Figure 5. Time evolution of datasets after Hurricane Maria.

is NASA's nightlight data, which is published for both events. After the first image is available for this dataset, the process is repeated daily to account for changes in the power supply after the disaster. Finally, the date when data is refined (e.g. both datasets shown in Figure 4) is presented with purple circles.

These figures illustrate the various times when data is collected and published after an event, highlighting the dynamics of the damage assessment process. For example, for Hurricane Maria, collection times range from zero to nine days after the event, and publication times range from 12 h to almost two years after the event. It is
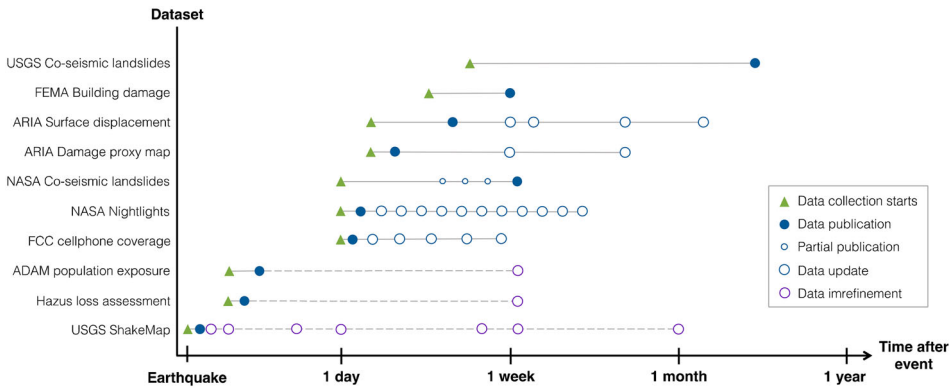
Figure 6. Time evolution of datasets after Indios Earthquake.

clear that the damage assessment process for infrastructure systems is a process that requires extensive time to complete, necessitating the analysis provided in this paper of the changes in types of datasets that are available over this long response period.

During the first day after both events, the data collection relies heavily on in-place sensors, given the difficulties of having personnel in the field during this time. For hurricane Maria, wind speed and rainfall data are collected using sensors that are distributed throughout the island. For the Indios earthquake, all three datasets released during the first day use data from the seismometer network of Puerto Rico (PRSN 2022). For instance, ground motion models are used to produce the ShakeMap, infrastructure loss models for Hazus, and population exposure for Automatic Disaster Analysis and Mapping (ADAM). This sequence demonstrates the value of having robust networks of sensors to rapidly collect exposure data after an event.

An analysis of the time of publication shows that different processing times are required, even for datasets that use the same input. For instance, both the damage proxy map (DPM) and FEMA wind damage for Hurricane Maria use satellite imagery as their input. However, their methodologies to estimate damage differ significantly. The DPM (published 2 days after the event) uses pre- and post-event satellite imagery to run precomputed algorithms that identify damage (Yun et al. 2015), while the FEMA wind damage data (completed 1 month after the event), requires the validation of damage through visual inspection of the images by FEMA personnel. The difference in interpretive approaches leads to different times required for publication. At the same time, however, the two approaches also carry different levels of uncertainty and damage reliability. The DPM, for example, is useful for identifying major damage regions, but its resolution is not sufficiently high to allow the identification of damage to individual buildings. On the contrary, FEMA wind damage data can be used to identify specific buildings needing repair or assistance by emergency managers. Thus, even though having data as soon as possible is preferable, some datasets require large post-processing to produce information at a high enough resolution to be useful for specific emergency purposes.

There is also a noted difference in publication time between the same dataset published for both types of disasters. In this case, the DPM is first released two days after Hurricane Maria and completes its coverage after nine days, while the one for the Indios Earthquake is periodically published from 3 to 19 days. This difference is a result of the pre-positioning of the satellites, which is possible given the increased time for preparedness for the hurricane compared to seismic events, leading to a shorter time to publication for hurricane DPM data. However, other datasets may take more time to be published for hurricanes. This is the case for NASA's night-time light data. This dataset requires clear skies for it to properly assess radiance from the region, which is usually not the case during or immediately after a hurricane. Therefore, night-time light data for Hurricane Maria started seven days after the storm, compared to one day after the Indios earthquake. These differences in publication time show that large coverage data require suitable weather for data collection, such as clear skies for satellite imaging, and no wind or rain for drone manipulation.

Not all datasets had different publication times for both disasters. Cellphone coverage reported by FCC is one dataset with similar publication times for both events. Both datasets had data available one day after the event and were updated daily for seven months after Hurricane Maria and for six days after the Indios earthquake. The difference in the number of publications is due to the impact each disaster had on the island's electric power network, where Hurricane Maria resulted in most of the Island being without power for weeks, with power outages in some areas lasting for ten months (Kwasinski et al. 2019).

The difference in the disaster impact area also results in varying data publications and post-processing times per dataset. More datasets were collected for Hurricane Maria compared to the Indios earthquake (13 as opposed to 10). This indicates the relationship between data collection and the type of disaster and its severity. Hurricane Maria was a category 5 hurricane (the highest on the scale), while the Indios earthquake was a moderate to high severity disaster. For the processing time per dataset (i.e. the time between data collection and publication), the average processing time for Hurricane Maria was 7.1 days, compared to 1.9 days for the Indios earthquake. These averages exclude both USGS landslide datasets, which are asset specific and can be considered outliers, with the one from Hurricane Maria taking 23 months to be published, and the one for the earthquake taking five months to be published. As a result, a disaster with larger impacts results in more datasets and quantities of data that require more time to be fully published.

The general evolution of data collection and publication is illustrated in Figure 7, which shows the cumulative percentage of datasets that were collected and published for both events over time. The figure enables a comparison of both collection time and publication time, as well as how collection and publication times of post-disaster datasets differ by type of hazard event. The cumulative collection times of both disasters are similar to each other, suggesting that the collection times for both events are somewhat event independent (with the earthquake having slightly earlier collection times). On the contrary, for publication times, there is a clear delay in the hurricane datasets, showing that for this event, it took longer to publish the datasets even when
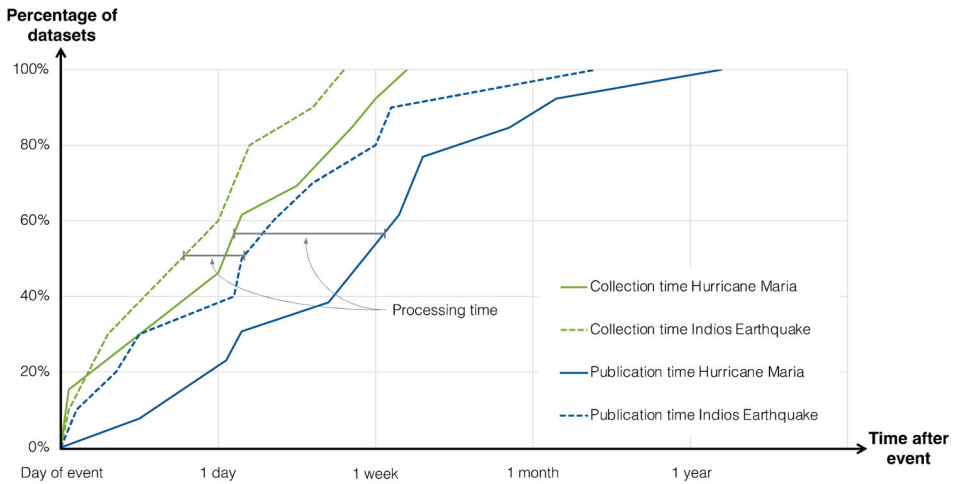
Figure 7. Percentage of total datasets collected and published for both disaster events.

they were collected at similar times after the event compared to the earthquake. Taken in combination with Figures 5 and 6, Figure 7 shows how data availability varies by the type and severity of a disaster. This difference is due to the increased processing time for the hurricane, which includes the time of collection, computation, and refinement, among others.

The evolution of time features in Figure 7 can also be interpreted as a trend in data collection and publication. Note that the time axis is not linear. However, these trends demonstrate how the data collection process can be improved. Mainly, the publication trend could be shorter for it to be more efficient. This can be accomplished by making two changes to the trends. First, data collection can be started earlier through increased data integration of early datasets and expansion of collaboration between entities to make the data collection more efficient. These changes can result in earlier publication times. The second change to the trend is to speed up processing times through automated procedures, distributed processing, or other strategies, which will make the publication trend and thus timing closer to the collection trend.

3.2. Resolution and coverage evolution

In addition to the time evolution of data availability, two critical parameters that characterize the utility of post-disaster datasets are resolution and coverage. The greater spatial resolution the data has, the more informed the decision-making can be in determining where to conduct field reconnaissance. In a similar fashion, a larger coverage area provides a more comprehensive picture of the damage extent. The most useful data, in this sense, is high spatial resolution data that covers the greatest area of impact and beyond. Thus, the spatial resolution and coverage of the datasets are equally important to several entities as the release and updating of data occurs. Figure 8 summarizes the times at which all 23 datasets were released following the two Puerto Rico disasters, along with their spatial resolution and coverage level.
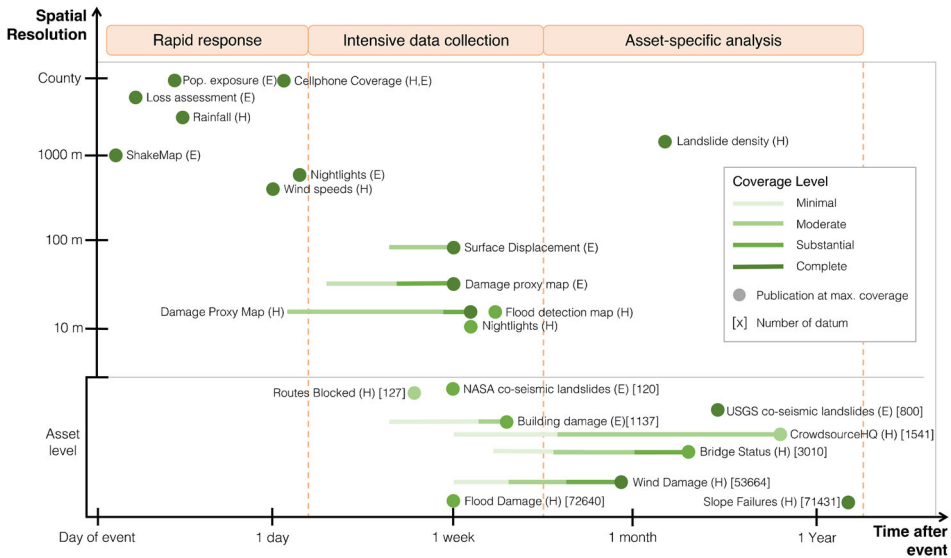
Figure 8. Spatial resolution and coverage *vs.* data release date for Puerto Rico disaster datasets.

Each dataset in Figure 8 is illustrated as a point located at the time when the publication of maximum coverage was released with its corresponding resolution. Thus, datasets that release multiple data versions or updates are only shown for the first time they were published (e.g. night-time light data, which does not change in resolution or coverage, but it is published daily, as shown in Figures 5 and 6). Regarding spatial resolution, the range varies from low resolution at the county level to high resolution at the asset level (e.g. point data per damaged bridge or line data per blocked route). Asset level is the highest resolution considered in this category because it provides information about a single infrastructure asset all data at the sub-component level (e.g. piers of a bridge or columns of a building) is considered at the asset-level as well. Within this category, the number of unique data observations is included to reflect the level of complexity and of the dataset. For example, in the Indios Earthquake, NASA's co-seismic landslide dataset includes 120 failures, whereas USGS's dataset includes 800 failures.

The dataset coverage is represented with a color scheme that follows the 4-level scale of coverage described in the previous section (i.e. Minimal, Moderate, Substantial, and Complete). Some datasets increase the coverage level as more data is collected. For example, data points on building wind damage from Hurricane Maria started being collected a week after the event, covering a minimal region of the island. Then, after a month, the coverage reaches a complete level after the island is fully studied. This evolution is indicated by a bar ranging from light green to dark green with time.

In analyzing the results shown in Figure 8, three distinct stages of data collection and processing are identified: an initial rapid response stage within the first day of the event, a period of intensive data collection from one day to a few weeks after the event, and then asset-specific analyses that take longer to complete with results that

are released from a few weeks to after one year following the event. These identified stages are described in further detail, including how the data collected during these varying stages are used by different entities, in the following section.

With respect to the evolution of data features, Figure 8 demonstrates a clear increase in the spatial resolution of datasets throughout the response stage for the Puerto Rico disasters. During the two days after the event, all eight published datasets had resolutions exceeding 500 m. Then, close to a week after the event, seven datasets were released with higher resolution, ranging from 90 m (earthquake surface displacement) down to the asset level (building flood damage). Following this phase, all datasets have a resolution at the asset level, except for the landslide density, which has a resolution of 2000 m. Still, it was then superseded 23 months later with the asset-level slope failures dataset. This evolution shows that the earliest time when asset-level datasets are published is about a week after the event.

The evolution of coverage is different from the spatial resolution. While resolution tends to increase with time after an event, the coverage area tends to decrease. All the datasets are released within two days after the events have complete coverage. Then, during the first week, remote sensing datasets with lower coverage start being collected (e.g. damage proxy map, surface displacement). Most of these datasets end with complete coverage after a week of the event since multiple images are taken to complete the coverage. Later in the data collection process, asset-level data is published at multiple coverage levels. For example, routes blocked from Hurricane Maria have moderate coverage, while building flood damage and NASA's co-seismic landslides have substantial coverage. This shows that the high-resolution datasets are often published at the expense of coverage. At the same time, the evolving coverage of some asset-level datasets, such as building wind damage or bridge status, shows that it is possible to eventually achieve high coverage in asset-level datasets. These datasets, however, require significant collection and processing times in the current practice.

Overall, the difference in coverage and resolution across the datasets arises from the nature of various user needs, tool capabilities, and available resources (personnel, tools, time). In this sense, early datasets prioritize high coverage because of rapid emergency needs and to provide broad situational awareness across an impacted community. This is shown in Figure 8, where all datasets released within the first two days after the event have complete coverage but low spatial resolution (exceeding 500 m). Then, datasets at a higher resolution are collected for specific regions or infrastructure systems without needing to cover all of the damage extent regions. For instance, NASA's co-seismic landslides and FEMA's building damage only cover the region close to the earthquake epicenter, where the most severe damage is anticipated. Finally, datasets with high resolution and coverage are expected to be published a long time after the event. This is the case of the hurricane wind damage (released one month after), USGS's co-seismic landslides (released five months after), and the hurricane slope failures (released 23 months after). The evolution of coverage and resolution of post-disaster data shows that the level of precision is heavily linked to the needs of the user and the limitations of the tool used for collection and the time required for data processing.

4. The three stages of post-disaster data collection

During the aftermath of a disaster, multiple entities use data at different times for different purposes. For instance, emergency managers prioritize evacuations and transportation of goods using preliminary assessments. Data about possible building collapse and open routes is important for this purpose. Some entities require more specific data (e.g. economic impact per neighborhood), which requires the use of different data that is usually not available at the early stages of the event response. The data needs and use of the data change as the post-disaster response and recovery process evolves. From Figure 8, three distinct stages of post-disaster data collection are identified; (1) a rapid response stage, where resources are being devoted to rescue and relief efforts; (2) an intensive data collection phase, where high spatial resolution and high coverage datasets are published for public use; and (3) an asset-specific stage where data on individual assets are being collected for the purposes of case studies and long-term studies including reconstruction. This section describes the differences in data availability for each of the three identified stages of post-disaster data collection. In addition, it analyzes how the available data matches with the range of entity need to support post-disaster activities, including how the data is used by entities with different backgrounds, purposes, and priorities in each stage. The information depicted in this section is supported not only by the specificity of the results shown in section 3 but also by interviews with emergency managers, disaster reconnaissance academics, and the authors' personal experiences conducting infrastructure damage reconnaissance.

4.1. Rapid response

The first stage of the data collection process comprises the post-disaster data used for rapid emergency tasks and situational awareness. The length of this stage typically ranges from 1 to 3 days, depending on weather conditions and available resources. In some cases, satellites need to be moved farther than in other cases to reach the needed positioning for data collection over the affected area, or information from sensors takes time to be collected and processed for publication. This stage is characterized by data of low resolution and high coverage area that often requires subsequent refinement due to its low precision.

Regarding data types, there are two metrics obtained from collected data that allow entities to estimate the damage in a region immediately after a disaster: the magnitude of the event and the impacted area. In the case of an earthquake, the magnitude is usually measured with the moment magnitude, a measure of the energy released in the event. The impacted area can be estimated with the interpolation of multiple seismometers. USGS's ShakeMap, published for the Indios earthquake, includes a version of both metrics as early as 5 min after the earthquake (Wald et al. 2005). On the other hand, the metrics for a hurricane differ in time and scale. Since hurricane forecasting is more predictive in nature than earthquake forecasting and detection, the likely intensity of the hurricane can be better estimated than the intensity of a future earthquake. The National Hurricane Center releases an estimated path of the event days before it makes landfall, along with periodic updates. This greatly improves the

preparedness and response to the event compared to an earthquake. However, esti-mating the impacted area can be more difficult than in an earthquake, given that the damage can be more spatially distributed compared to the centralized damage of an earthquake. Further, the characteristics of a hurricane change significantly once it makes landfall compared to when it passes over an open ocean. By coupling these event-specific datasets with the vulnerability metrics of a region, FEMA releases an early estimate of economic loss using the Hazus platform (Kircher et al. 2006; Scawthorn et al. 2006), which is updated as new versions of the ShakeMap and field surveys are released.

Local emergency managers, who conduct damage assessments within the first days of the event, use these intensity measures to gain situational awareness and estimate the resources needed for response. Also, since satellite images only provide an aerial view of the region, they are not capable of finding all buildings with high damage or in need of evacuation, only those that have collapsed and where roads are blocked by debris. As a result, local emergency managers with sufficient resources often conduct fly-overs in rotary or fixed-wing aircraft hours after an event to get a closer view of the affected areas. This enables qualitative assessment of the damage extent before more quantitative-based measures can be deployed in later stages. A problem with fly-over data and its role in post-disaster data collection is that it is rarely recorded but plays a role in early decision-making.

The collection procedure of some datasets during this stage makes them ideal for being consistently published after every major disaster. For instance, FCC's cellphone coverage data reports were identical for Hurricane Maria and Indios earthquake (i.e. same coverage, data structure, and format) because Puerto Rico counts with network outage data provided by the Disaster Information Reporting System (DIRS), facilitat-ing the workflow of post-disaster data. Another example of earthquake events is USGS's ShakeMap, which publishes a ShakeMap for every major earthquake around the world. Having datasets with reliable publication times makes them useful for implementing damage models that require consistent inputs after every disaster. In addition, entities in need of hazard characterization or damage data can expect with high confidence the fast publication of these datasets after a disaster.

The lack of high-resolution data, however, demonstrates that there is still signifi-cant room for improvement in the rapid collection and publication of perishable data. Instead, datasets published in this stage provide low-resolution, hazard-specific information (as shown in Figure 8), such as the ShakeMap, loss estimates, and rain-fall. While these datasets provide useful intensity measures, they do not provide enough detailed information to first responders about on-the-ground conditions. As a result, most of these datasets need to be refined for improved precision in the later stages of the disaster response. Note that in Figures 5 and 6, all the datasets collected within the first day of the event include a refinement component later in the timeline of the event. This stage provides an opportunity for implementing emerging technologies, such as machine learning, to accelerate data processing and improve damage estimations by training models with impacts from previous disas-ters. Some current efforts for this purpose can be found in Yeum et al. (2019) and Miura et al. (2020).

4.2. Intensive data collection

The second stage of the post-disaster data collection stage usually starts half a week after the event (3–4 days), when most rescue and recovery missions are completed and finishes around 1–2 weeks after when most of the critical perishable information has been collected. During this stage, determining what type of data to collect becomes more challenging after the initial rescue and recovery phase, given the diverse set of assets that need to be assessed. Information on critical infrastructures, such as hospitals, schools, and fire stations is prioritized not only during the initial phases of data collection but also throughout later stages as a function of local conditions and needs. Thus, one entity's needs often intersect or are adjacent to another's, such as a scientific reconnaissance team studying the effect of flooding on bridges and a local government planning bridge infrastructure recovery/reconstruction to support the transportation of goods to a community. If there is communication between data producers and local entities that need data, it is often initiated at this stage. However, this process is typically done in an ad-hoc manner or based on prior standing relationships and has yet to be optimized to meet the needs of as many entities as possible.

Regarding the needs of local governments, emergency managers are interested in cataloging damage in greater detail than was possible in the immediate aftermath of the event. In the U.S., for example, local governments must begin the process of completing the Initial Damage Assessment (IDA), which the state or territorial government helps coordinate. The goal of the IDA is to document and categorize all relevant damage impacts and information, which they then convey to FEMA; this is the basis by which all additional external support and resources are distributed. As extensive as the IDA may be at classifying and identifying damage, this data is not made public and is only used internally for financial decision-making purposes. Nonetheless, local governments use publicly available data from this data collection stage to support reconnaissance missions and validate damage.

Once the IDA is complete and if the damage is severe enough, FEMA will send representatives to the field with local and state workers to begin a Preliminary Damage Assessment (PDA). The goal of the PDA is not to discover and classify more damage but only to validate the damage originally presented to FEMA in the IDA. The PDA is then used to determine the amount of financial support that is given to the area and to inform the potential decision of a Presidential disaster declaration (FEMA 2021). As a result of the PDA, multiple types of infrastructure damage data are released during this stage. From the Puerto Rico disasters studied in this paper, FEMA datasets in this stage include the flood detection map and building damage datasets at the asset level (i.e. flood damage, wind damage, and seismic damage).

Compared to datasets from the previous stage, datasets of the intensive data collection stage are characterized by their increased spatial resolution. This is possible due to the availability of more specialized tools and personnel on the field. One of the most common data types seen during this stage is remotely sensed images. Among these images, there are multiple datasets collected by different entities, such as NASA or NOAA as well as private satellite operators. Some of these images use a special type of satellite sensor called Synthetic Aperture Radar (SAR), which allows the

collection of more information compared to conventional optical or electromagnetic spectrum imagery. NASA releases a product to estimate damage called the Damage Proxy Map (DPM), in which they compare pre- and post-event images through computer vision algorithms to determine regions of damage (Yun et al. 2015). The resolution of these images is 15–30 m, a higher resolution compared to maps released during the rapid response stage, which range between 500 and 10,000 m. This dataset provides more information about the distribution of infrastructure damage. However, as of now, the DPM has not been used to determine specific building damages at the asset level, given the high uncertainty associated with this dataset compared to, for example, on-the-ground observations. In addition to the DPM, the use of computer vision on high-resolution images is also observed in datasets, such as NASA's surface displacement, night-time light data, and co-seismic landslides. FEMA's use of computer vision includes the flood detection map and building flood damage.

Even when most of the data at this stage has higher resolution, not all of it has complete coverage. Some datasets use collection procedures that require complex computations or large efforts to complete, and thus, data producers prioritize data collection in areas of high expected damage. For example, FEMA's flood detection maps from Hurricane Maria only cover the coastal area of Puerto Rico that surrounds the two biggest cities on the island (San Juan and Ponce). In a similar manner, NASA's co-seismic landslides use high-resolution satellite imagery to identify landslides only in the area close to the Indios epicenter. Another reason for variable coverage is the time it requires to collect damage data at high resolution. For instance, this stage of data collection is the first to collect asset-level data, which requires cataloging a damage level per asset (e.g. one damage point per bridge or building). This procedure requires time to complete, especially after events of high impact, such as Hurricane Maria. The coverage evolution of asset-level data can be seen on the building wind damage dataset, which starts being collected a week after the event in small regions of the island at a minimal coverage level.

The variety of data published during this stage shows that multiple entities are interested in obtaining high-resolution data over the highest coverage possible. In that sense, having multiple data producers working at the same time opens an opportunity to increase data integration and entity collaboration. While conversations between public, private, and academic institutions do happen during this stage and there are current platforms, such as DesignSafe that advocate for improved data publication (Rathje et al. 2017), there is no systematic procedure to collect and share post-disaster data. Thus, from this work, it is recommended that the intensive data collection stage be the time when entities collaborate the most to produce data of better quality in less time.

4.3. Asset-specific analysis

In the weeks following an event, publicly available data begins to shift from the broad characteristics of the event impact to characterizing the built and natural environment's specific responses and impacts from the event. Therefore, this data is more useful for long-term academic studies and community recovery strategies. For

instance, asset-level data increases in coverage to support building-by-building recovery strategies, crowdsourcing initiatives are undertaken for reaching smaller or more remote affected regions, and more field reconnaissance teams with a range of sensing systems are mobilized to study the impact of the disaster. It is at this stage in the data collection process that data unrelated to emergency tasks is processed to produce more detailed scientific artifacts, such as detailed landslide maps for understanding complex causative mechanisms and bridge 3D scans for studying earthquake damages (Chen et al. 2012; Hughes et al. 2019).

Given the specific needs and procedures of academic institutions and the intricacies of scientific research, the time frame for datasets produced during this stage varies considerably, starting 2–3 weeks after the event, after most rescue missions are completed, and finishing one or multiple years after the event, after all data is analyzed and published.

These datasets are typically not used by local or state governments, as municipalities may still be working on the IDA and PDA at this point in the process rather than on quantifying damage. The IDA and PDA are more concerned with discovering and validating damage rather than quantifying it. The limiting factor for local governments in a post-disaster data collection process is often human labor. Local governments must work quickly in the aftermath of a disaster to complete the IDA, which is entirely done by in-person reconnaissance efforts. As a result, local governments do not prioritize the collection of remote sensing data at this stage, even though it may be used to support their IDA findings or be of interest to the research community or others.

Once the IDA and PDA are completed, local entities are no longer heavily involved in the data collection process. At this point, data publication begins to slow down and become increasingly niche-focused, mainly to inform specific academic research or practice-oriented questions. Thus, at the expense of increasing the time of publication, most datasets have a high resolution at the asset level and high levels of spatial coverage (see Figure 8). The post-IDA and -PDA environment is also the time when holistic reconnaissance reports are published, including those from the Geotechnical Extreme Events Reconnaissance (GEER) and Structural Extreme Events Reconnaissance (STEER), which document engineering effects after disasters to improve scientific practice (Bray et al. 2018; Kijewski-Correa et al. 2021). These reports and research papers usually mark the end of the damage assessment process that includes perishable data. The cumulative dataset publication time depends on the severity of the event and the procedure used to analyze the data. For example, in terms of data collection, for the Indios earthquake, STEER deployed a small team of four people for 3 days and published a report <2 months after the earthquake (Miranda et al. 2020). On the contrary, for Hurricane Maria, there was a large GEER team of 14 people who cataloged infrastructure damages for 12 days and published a report 9 months after the hurricane (Silva-Tulla 2018).

5. Improving post-disaster infrastructure damage assessment

This section focuses on key challenges and gaps in the collection and publication of post-disaster data identified from analyzing the evolution of spatial and temporal data

features. This is important for entities interested in improving methodologies to assess damage in infrastructure systems, both by acquiring better data and by sharing it in a way that is helpful for others in the post-disaster damage assessment process. Doing so benefits entities interested in rapid damage assessment, since being able to plan for a dataset to be released at a certain point in time, with a certain spatial resolution, and a certain coverage area, allows for more in-depth and secure planning on the part of response operations. The analysis conducted in Section 3 also revealed that datasets are generally not well documented, with the authors needing to undertake significant efforts to infer key data features. In fact, most existing datasets lack all the proper metadata to understand how they were collected and the features of the resulting datasets. An improvement of this does not necessarily require a standard for all datasets, which would be impractical given the diversity of features and users. However, understanding the challenges and opportunities described in this and the following section can help future researchers and emergency managers improve the process of both collection and publication of post-disaster data, and identify where the greatest gains might be made in improving post-disaster damage assessment.

5.1. Improving entity collaboration and integration across datasets

The intensive data collection stage from section 4.2 (i.e. the second stage identified during the post-disaster data collection process) proved to be the time when most datasets are published, and at the same time, the period when more entities work simultaneously collecting data. However, the inconsistency in metadata published during this time and the lack of data integration show that there is an opportunity for improving entity collaboration. For instance, some of the data is not yet being used at its full potential during this stage. FEMA's Hazus estimates loss impacts after disasters by coupling hazard data (e.g. PGA of the earthquake) with structural components of the building environment. However, the results of this methodology are highly uncertain given the low precision of the structural parameters. Thus, this dataset can be improved by validating damage with partial field data or high-resolution imagery that is collected during the second stage of data collection. By improving data integration, the data collection can be more efficient by speeding up the coordination of data collection, resulting in earlier publication times. In addition, data integration increases data precision by joining multiple sources of damage data, decreasing uncertainty and increasing the reliability of the estimations (Lee and Tien 2018).

During this second stage of post-disaster data collection, it is important to note that some of the field data are not recorded with the purpose of improving the response or recovery for that specific event. Most academic reconnaissance teams arrive at the field a couple of weeks after the event, so they do not interfere with rescue or recovery missions. However, some of the data collected by researchers can be useful to emergency managers during the first stage of the damage assessment. Thus, when possible, researchers should be in contact with local governments to provide support in the damage assessment process. This collaboration is crucial in cases where there is a lack of resources from the government, particularly on-the-ground

resources, which limits the ability of emergency managers to fully and precisely assess damage in their community. In these cases, additional data collected by other entities, such as researchers, can augment the amount of information available to support community disaster response.

Finally, given the importance of Initial Damage Assessments (IDAs) and Preliminary Damage Assessments (PDAs) in the ability for communities to receive resources and financial support for disaster response and recovery, other post-disaster datasets can be used to support these efforts. Instead of relying entirely on in-person reconnaissance efforts to complete the IDA or PDA, datasets, such as satellite and drone imagery can be used to estimate building damage. Moreover, these datasets can be coupled with partial in-person data to increase their reliability. Including in-person datasets in this process would require both in-person and remote datasets to have consistent availability and accessibility, as the in-person reconnaissance data is most often private.

5.2. Increasing rapid high-resolution data

Remote sensing is the most common method of obtaining large-extent datasets in a post-disaster environment. From the analysis of the resolution evolution of the previous section, we demonstrated that all remote sensing data from the two first days after a disaster are low resolution (exceeding 500 m). As a result, during the first stage of data collection, available data can only be used for rapid emergency tasks and situational awareness. The lack of tools and rapid automated processes during this stage makes post-disaster data unable to be used for more decision-support tools, such as reliable building evacuations and speeding up the federal assistance process. Thus, high-resolution data from the second stage should be published earlier by enhancing current methodologies used to process post-disaster data. For example, building damage assessments can be conducted *via* crowdsourcing of aerial photographs (Barrington et al. 2011) or automated processes (Eguchi et al. 2009), and slope failures can be derived from satellite imagery (Scaioni et al. 2014); both examples result in valuable high-resolution data that can be on the sub-meter scale.

During the rescue and recovery phases, local emergency managers can be resistant to developing a data-driven system to inform their decision-making (Campion 2020; Delaney and Kitchin 2023). The lack of early high-resolution data on buildings and high-risk infrastructure, such as dams, levees, and slopes contributes to the need to implement new data collection tools and increase the use of data analysis techniques in the damage assessment process. From the results of this study, the recommendation is to leverage and increase the use of new technologies to support needed early high-resolution data collection. The requirement for in-person preliminary damage classifications of buildings is becoming inefficient with the emergence of UAVs, 360° StreetView cameras, and high-resolution satellite imagery, all of which can be used to classify building damage at the level needed by emergency response teams. Data acquisition using these technologies is much quicker than in person-classifications; the post-processing of images is the step that requires the most time.

In addition to data collection, some high-resolution data can take a long time to be published. It is recommended that data be published as it is collected and updated with greater coverage as more data is collected in the days following rather than waiting for publication as one complete dataset. Following this methodology of data publication allows local entities more information on which to base decision-making from the outset. Making this change is especially significant for data that requires considerable manual post-processing. If datasets that contain a large amount of information, including high-resolution information, are published months or years after the event, the novelty and value of the information is diminished. However, if data is systematically processed as a function of location, the data can be published as the post-processing occurs, releasing different geographic regions of information as the post-processing continues. In the case of Hurricane Maria, for example, a dense dataset of 71,431 slope failures was published over one year after the event. In contrast, a dataset of wind damage to buildings during Hurricane Maria was first published one week after the event, which only represented a small geographic region. In the month following, geographic regions and the corresponding building damages were added incrementally. Early data with a defined, albeit incomplete, coverage combined with consistent updates, is more valuable than a lack of data for a significant duration followed by one large publication. Thus, from the results, we identified that earlier data availability, although incomplete, enables increased use and interoperability of datasets in the critical earlier time window for post-disaster damage assessment.

5.3. Developing strategies for data shareability

One of the main challenges with sharing post-disaster data is that it is often collected to serve current needs. It has been shown that these needs change not only by entity but by time as well. Therefore, having robust information about a dataset allows a user to maintain a high value of the data over a longer time, increasing the possibilities of future users being able to extract more information even when they come from different backgrounds or are using it years after the data is collected. In addition, improved data-sharing practices are not only helpful for future users. Even within a public or private entity, having a clearer structure of the data and more precise information about a dataset improves the understanding of its capabilities. In the case of researchers, a better understanding of the data can lead to more rigorous use of these datasets in research. Further, having detailed information about datasets supports improved data integration and interoperability. For example, if it is desired to overlay multiple datasets, information including temporal characteristics and spatial resolution specific to each dataset is critical to produce accurate information overlays and analyses.

The lack of shareability applies to multiple aspects of the metadata. For instance, recognizing the limitations of damage maps constructed with computer vision algorithms gives emergency managers the necessary information to locate damaged regions, but not enough precision to locate damaged buildings, even when the maps are constructed at high resolution. This requires an understanding of what each parameter means for the building damage and what is the likelihood of the damage

based on the imagery information. Another example is failing to define the spatial resolution of point, line, and polygon data in a geographic information system (GIS); when examining an individual point and line datum, the larger context and dataset to which each datum belongs should not be lost. In this study, seven data sources failed to mention the spatial resolution of the tool used to collect point, line, or polygon data, which required a distinct area of the graph that could contain geospatial data of unspecified resolution.

Recent tools have proven to increase data shareability in the context of natural hazards. One of them is the DesignSafe cyberinfrastructure (Rathje et al. 2017). This online platform integrates diverse datasets so that researchers can easily share and find experimental or reconnaissance data. However, most of the metadata for post-disaster needs is very user specific, making the data curation process hard to standardize for damage data in other infrastructure sectors [see the StEER guidelines for building infrastructure (Roueche 2020)]. In addition, current metadata features only include the last updating date for the dataset, meaning that there is no information about the time when the data was collected or when it was first available. Thus, while some aspects of data shareability have improved with DesignSafe (e.g. including DOIs for datasets and having a single platform for infrastructure data), some features are still missing, making it hard to obtain all the information needed for a rigorous damage assessment. To overcome this, the next section proposes a generalized set of metadata features to fully capture the characteristics of an infrastructure damage dataset.

6. Metadata for improved data shareability

A detailed analysis of post-disaster infrastructure damage data requires detailed cataloging of all metadata to enable comparison and integration of information across datasets. For the datasets included in this study, the resulting collected metadata are shown in Tables 2 and 3. However, the process of discovering and reporting the metadata for various infrastructure datasets proved to be challenging. Most datasets are inconsistent when presenting the metadata, and some of them publish little to no metadata at all. For instance, USGS's ShakeMap only publishes the time of publication of the last version of the map on the website. In this case, it was necessary to contact an expert from USGS to know that there is a public server catalogue that can be used to download all other metadata. Another example is the crowdsourced data for Hurricane Maria, which was inconsistent in terms of metadata about both geolocation and time. Specifically, some datapoints included location coordinates for damages and others did not, and some datapoints did not include a timestamp. In addition, information is often not easily accessible or downloadable. For example, the authors had to use web scraping to read pdf reports from the FCC to geolocate information on the percentage of cellphone coverage. This process highlighted the lack of coordination and consistency in publishing metadata for infrastructure data. Metadata should be as easily accessible as the data itself. The additional activities required to access and infer metadata create barriers for users to extract more information from collected datasets.

Table 4. Essential metadata for post-disaster datasets related to damage assessment.

Feature	Value	Example	Datasets with feature (%)
Collection date	Gives a snapshot of the state of the damage depending on when the data was collected	Collected: 01-01-21	91
Publication date	Gives a proxy of the time it took to both collect and process the data before publication	Published: 01-03-21	74
Updating date	Demonstrates that more information was gathered to decrease the uncertainty in the past data or to increase the total coverage	Updated: 01-10-21	88
Geo-location	Allows to intersect different formats and datasets to the same infrastructure asset	Coordinate pairs [latitude, longitude]	91
Resolution	Provides a look at the precision of the data and its reliability	1 m grid, point data per building	39
Coverage	Gives boundaries to the analysis and establish regions with no data	10 × 10 km, County level	28
Tool	Speaks to the uncertainty in the data depending on the technology of the tool	WorldView-2 Satellite Sensor	58
Format	Informs the type of software needed to process the data (e.g. word processor, GIS tool, programming language) and its complexity	.txt .HDF .csv	45
Contact info	Enables more information and interoperability of the data with the connection of people who collected and curated the data	Email, phone number	74

Given the need for improved shareability in post-disaster datasets, this work recommends that there be a minimum set of metadata features for all post-disaster datasets to solve some common issues in using the datasets to relate the full story of a disaster. This set of features defines components of post-disaster metadata that are vital to emergency managers, scientific reconnaissance teams, and future long-term studies alike. The set is composed of nine main features, described in Table 4.

Of the proposed data features, the first three capture the temporal evolution of datasets. Detailing the time stamps across elements of the data collection and publication process reduces temporal uncertainty, specifying when the point of interest was in the recorded state. The criticality of assigning a time stamp applies to photographic and imagery data, as well as any published dataset from a post-disaster environment. Reconnaissance data often has a collection date, publication date, date when updated, and when more data was added, improved, or otherwise modified after further investigation. These three distinct dates—collection data, publication data, and date when updated—are all equally important to distinguish in the metadata. Without these dates, potential future studies of the timeline of an event or investigations of cascading failures that occur in the time period after an event, to name a few, are made impossible. Thus, the specification of data timelines facilitates emerging areas of resilience research.

Geo-locating every photo, video, or survey taken at the disaster site is vital. The latitude and longitude, similar to the time stamp, reduces spatial uncertainty,

specifying where a point of interest from the event is located. In addition to a specific location where the data is collected, emerging data collection systems are now capturing orientation at the time of data collection too. This information is critical to understanding the subsequent data features of resolution, coverage, data collection tool or equipment, and format of the unprocessed information, which give a description of the geographical details of the data and facilitates cross-dataset comparisons. Finally, having contact information creates a social network of persons and entities responsible for the dataset in the study. Including these nine metadata features will improve data use and shareability for post-disaster datasets.

Table 4 also provides examples of each metadata feature and the percentage of datasets from the Puerto Rico disasters that included each feature. This analysis helps to understand the current state of metadata for damage assessment compared to a state where datasets have all features published. Of the nine metadata features, only three of them are included in 75% or more of the datasets. These three features are publication date, updating date, and geo-location. While this demonstrates the implementation of GIS tools and timeline features in the data, there are still important features that are not being recorded in the metadata. There is no feature with 100% implementation, which means that none of the features were recorded in all datasets.

These features provide sufficient information to understand how useful a dataset is within disaster response or recovery efforts and for later research purposes. Moreover, it positions post-disaster data to be in line with the FAIR guiding principles proposed for data management and stewardship in the scientific community (Wilkinson et al. 2016). This guideline includes four main principles: Findability, Accessibility, Interoperability, and Reusability. By implementing the small set of proposed features on any post-disaster dataset as described here, multiple communities will benefit from the increased information and interoperability of the datasets beyond what currently exists.

7. Conclusions

With the advent of new technological tools and data globalization, the publication of higher fidelity post-disaster datasets for damage assessment of infrastructure systems has increased in recent years. However, most publicly available data is not envisaged for or prepared to be used by entities with multiple purposes. The work in this study demonstrates the temporal and spatial evolution of post-disaster datasets and recommends strategies to improve data availability, shareability, and use of the data to decrease uncertainty in the data for damage assessment. Given the correlations between the needs and motivations of an entity and the availability of each dataset, this study provides an in-depth analysis of the data based on temporal features (data collection, publication, and updating times), as well as spatial resolution and coverage. We accomplish this by analyzing critical metadata features at an unprecedented specificity, gathering and examining infrastructure damage data from two disaster events: 2017s Hurricane Maria and 2020s Indios Earthquake in Puerto Rico.

The results of the metadata analysis show how the temporal features of post-disaster datasets evolve during the response to a disaster. In Hurricane Maria, for example,

collection times range from zero to nine days while publication times range from 12 h to almost two years after the event, highlighting the complexity of the damage assessment process, which can require multiple months to be completed.

A clear difference was shown between the data collection process for a hurricane compared to an earthquake event. For instance, the magnitude of the hazard is better predicted before the impact of a hurricane event, giving time for scientists and data collection entities to prepare for the disaster. However, the reliability and robustness of the instrumentation for earthquakes make it possible to have accurate measures of event intensities in just a couple of minutes after the event, even when current technologies cannot fully predict the location and magnitude of an impending earthquake. Publication of hurricane-related damage datasets are also found to take longer compared to those for earthquake damage, particularly for severe storm event with widespread impacts. The average processing time for Hurricane Maria was 7.1 days, compared to 1.9 days for the Indios earthquake.

The detailed analysis conducted in this study of the evolution of coverage and spatial resolution of the post-disaster datasets reveals a clear increase in spatial resolution throughout the post-disaster response period for the Puerto Rico disasters. First, one day after the event, the data collection relies heavily on remote sensing tools given the difficulties of having personnel in the field, offering information with broad coverage but low resolution. Then, the use of more specialized tools and field missions increases the accuracy and resolution of damage estimations at later stages of the response. At the same time, coverage levels tend to decrease with time with the use of specialized tools that cover less area and the implementation of complex data collection and processing procedures that require significant time to cover large areas.

As a result of the metadata analysis, three distinct stages of post-disaster data collection are identified: (i) a rapid response stage, where resources are devoted to rescue and relief efforts and data is characterized by high coverage and low resolution; (ii) an intensive data collection stage, where high spatial resolution is incorporated, and most datasets are collected and published; and (iii) an asset-specific stage, where data on individual assets are being collected for the purposes of case studies and long-term studies.

In addition to the metadata analysis, three main challenges and opportunities for improving the post-disaster damage assessment process were identified: (i) improve data integration and entity collaboration to make damage assessment more systematic and potentially decrease uncertainty in the collected data; (ii) increase the publication of rapid high-resolution data to support emergency tasks during the first days after a disaster, particularly with the implementation of new technologies combined with current practice; and (iii) increase data shareability to enable combining of information from multiple datasets and facilitate extracting more information from the data even when it comes from different backgrounds and entities or is used years after it was originally collected.

Through rigorous analysis of post-disaster datasets and to meet the need for improving data shareability, several vital pieces of metadata have been identified: collection date, publication date, updating date, geo-location, resolution, coverage, tool/equipment, format, and contact information, with the recommendation to include each of these pieces of metadata in dataset publication. The purpose of this recommendation is not

to make the publication process tedious, but to move towards a standard approach that facilitates dataset interoperability with the aim is to minimize the uncertainty not only in data collection but also in the use of the datasets in future studies of the disaster. Post-disaster data collection may serve a singular purpose during the time of collection, but it could well inform myriad future studies of significant value.

The representation of post-disaster data in this study as a dynamic rather than static process lends itself to many possible future studies of the evolution of data temporally and spatially. Work can be done to quantify the uncertainty of data and how the uncertainty of datasets contributes to damage quantification over time. This work was outside the scope of this study and cannot be implemented with the currently accessible information about the datasets. Future work could also include quantifying how the inclusion of a metadata standard affects the uncertainty of post-disaster data and determining to what degree other possible metadata or auxiliary information can help minimize uncertainty in damage estimation. Such studies building off the findings presented in this paper would help facilitate not only reconnaissance work but also early relief and long-term recovery efforts.

ORCID

Jorge Mario Lozano http://orcid.org/0000-0002-9219-3879
Iris Tien http://orcid.org/0000-0002-1410-632X

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Abruzzo J, Matta A, Panariello G. 2006. Study of mitigation strategies for progressive collapse of a reinforced concrete commercial building. J Perform Constr Facil. 20(4):384–390. doi: 10.1061/(ASCE)0887-3828(2006)20:4(384).

Adams SM, Levitan ML, Friedland CJ. 2013. High resolution imagery collection utilizing unmanned aerial vehicles (UAVs) for post-disaster studies. In: Advances in hurricane engineering: learning from our past; p. 777–793. doi: 10.1061/9780784412626.067

Akter S, Wamba SF. 2019. Big data and disaster management: a systematic review and agenda for future research. Ann Oper Res. 283(1–2):939–959. doi: 10.1007/s10479-017-2584-2.

ARA Inc. 2017. Hurricane Maria wind data. https://disasters.geoplatform.gov/publicdata/NationalDisasters/2017/HurricaneMaria/Data/Wind/ARA/.

Barrington L, Ghosh S, Greene M, Har-Noy S, Berger J, Gill S, Lin AY-M, Huyck C. 2011. Crowdsourcing earthquake damage assessment using remote sensing imagery. Ann Geophys. 54(6):680–687.

Bell J, Schultz LA, Molthan A, Kirschbaum D, Roman M, Yun S-H, Meyer FJ, Hogenson K, Gens R, Goodman H. 2018. NASA earth science disasters program response activities during Hurricanes Harvey, Irma, and Maria in 2017. American Meteorological Society Annual Meeting 2018: IN31F-04.

Berman JW, Wartman J, Olsen M, Irish JL, Miles SB, Tanner T, Gurley K, Lowes L, Bostrom A, Dafni J, et al. 2020. Natural hazards reconnaissance with the NHERI RAPID facility. Front Built Environ. 6:573067. doi: 10.3389/fbuil.2020.573067.

Bessette-Kirton EK, Coe JA, Kean JW, Godt JW, Rengers FK, Schulz W, Baum RL, Jones ES, Staley DM. 2017. Map data showing concentration of landslides caused by Hurricane Maria in Puerto Rico; [accessed 2021 Jun 8]. https://www.sciencebase.gov/catalog/item/59de6459e4b05fe04ccd39d8.

Boose ER, Serrano MI, Foster DR. 2004. Landscape and regional impacts of hurricanes in Puerto Rico. Ecol Monogr. 74(2):335–352. doi: 10.1890/02-4057.

Bray JD, Frost JD. 2010. Geo-engineering reconnaissance of the 2010 Maule, Chile Earthquake. http://www.geerassociation.org.

Bray JD, Frost JD, Rathje EM, Garcia FE. 2018. Turning disaster into knowledge in geotechnical earthquake engineering. In: Geotechnical Earthquake Engineering and Soil Dynamics V. Austin, TX: American Society of Civil Engineers; p. 186–200. [accessed 2021 Apr 12]. doi: 10.1061/9780784481462.018.

Bray JD, Frost JD, Rathje EM, Garcia FE. 2019. Recent advances in geotechnical post-earthquake reconnaissance. Front Built Environ. 5(5). doi: 10.3389/fbuil.2019.00005.

Campion P. 2020. Building a data-driven culture in emergency management. https://www.domesticpreparedness.com/preparedness/building-a-data-driven-culture-in-emergency-management/.

Cangialosi JP, Latto AS, Berg R. 2018. Hurricane Irma tropical cyclone report. National Hurricane Center (NHC) Report No.: AL112017. https://www.nhc.noaa.gov/data/tcr/AL112017_Irma.pdf.

Chen S-E, Liu W, Bian H, Smith B. 2012. 3D LiDAR scans for bridge damage evaluations. In: Forensic Engineering 2012. San Francisco, CA: American Society of Civil Engineers; p. 487–495. [accessed 2020 Sep 8]. doi: 10.1061/9780784412640.052.

Costa AC, Sousa ML, Carvalho A, Coelho E. 2010. Evaluation of seismic risk and mitigation strategies for the existing building stock: application of LNECloss to the metropolitan area of Lisbon. Bull Earthq Eng. 8(1):119–134. doi: 10.1007/s10518-009-9160-3.

Del Gaudio C, De Martino G, Di Ludovico M, Manfredi G, Prota A, Ricci P, Verderame GM. 2017. Empirical fragility curves from damage data on RC buildings after the 2009 L'Aquila earthquake. Bull Earthq Eng. 15(4):1425–1450. doi: 10.1007/s10518-016-0026-1.

Delaney A, Kitchin R. 2023. Progress and prospects for data-driven coordinated management and emergency response: the case of Ireland. Territ Polit Gov. 11(1):174–189. doi: 10.1080/21622671.2020.1805355.

Eguchi RT, Huyck CK, Ghosh S, Adams BJ, McMillan A. 2009. Utilizing new technologies in managing hazards and disasters. In: Showalter PS, Lu Y, editors. Geospatial techniques in urban hazard and disaster analysis. Dordrecht: Springer Netherlands; p. 295–323. [accessed 2023 Jan 6]. doi: 10.1007/978-90-481-2238-7_15.

FEMA. 2021. Preliminary damage assessment guide. FEMA. https://www.fema.gov/disaster/how-declared/preliminary-damage-assessments/guide.

Galvis FA, Miranda E, Heresi P, Dávalos H, Ruiz-García J. 2020. Overview of collapsed buildings in Mexico City after the 19 September 2017 (M_w 7.1) earthquake. Earthq Spectra. 36(2_suppl):83–109. doi: 10.1177/8755293020936694.

Gautam D. 2018. Observational fragility functions for residential stone masonry buildings in Nepal. Bull Earthq Eng. 16(10):4661–4673. doi: 10.1007/s10518-018-0372-2.

Gautam D, Fabbrocino G, Santucci de Magistris F. 2018. Derive empirical fragility functions for Nepali residential buildings. Eng Struct. 171:617–628. doi: 10.1016/j.engstruct.2018.06.018.

Ghannad P, Lee Y-C, Choi JO. 2021. Prioritizing post-disaster recovery of transportation infrastructure systems using multiagent reinforcement learning. J Manage Eng. 37(1):04020100. doi: 10.1061/(ASCE)ME.1943-5479.0000868.

González AD, Dueñas-Osorio L, Sánchez-Silva M, Medaglia AL. 2016. The interdependent network design problem for optimal infrastructure system restoration: the interdependent network design problem for infrastructure restoration. Comput-Aided Civ Infrastruct Eng. 31(5):334–350. doi: 10.1111/mice.12171.

Hall J, Muscarella R, Quebbeman A, Arellano G, Thompson J, Zimmerman JK, Uriarte M. 2020. Hurricane-induced rainfall is a stronger predictor of tropical forest damage in Puerto Rico than maximum wind speeds. Sci Rep. 10(1):4318. doi: 10.1038/s41598-020-61164-2.

Hao H, Wang Y. 2020. Leveraging multimodal social media data for rapid disaster damage assessment. Int J Disaster Risk Reduct. 51:101760. doi: 10.1016/j.ijdrr.2020.101760.

Hong Y, Adler R, Huffman G. 2006. Evaluation of the potential of NASA multi-satellite precipitation analysis in global landslide hazard assessment. Geophys Res Lett. 33(22):L22402. doi: 10.1029/2006GL028010.

Hosseini Nourzad SH, Pradhan A. 2016. Vulnerability of infrastructure systems: macroscopic analysis of critical disruptions on road networks. J Infrastruct Syst. 22(1):04015014. doi: 10.1061/(ASCE)IS.1943-555X.0000266.

Hughes KS, Schulz WH, Baum RL. 2019. Map of slope-failure locations in Puerto Rico after Hurricane Maria; [accessed 2021 Jun 8]. https://www.sciencebase.gov/catalog/item/5d4c8b26e4b01d82ce8dfeb0.

Kijewski-Correa T, Roueche DB, Mosalam KM, Prevatt DO, Robertson I. 2021. StEER: a community-centered approach to assessing the performance of the built environment after natural hazard events. Front Built Environ. 7:636197. doi: 10.3389/fbuil.2021.636197.

Kircher CA, Whitman RV, Holmes WT. 2006. HAZUS earthquake loss estimation methods. Nat Hazards Rev. 7(2):45–59. doi: 10.1061/(ASCE)1527-6988(2006)7:2(45).

Kwasinski A, Andrade F, Castro-Sitiriche MJ, O'Neill-Carrillo E. 2019. Hurricane Maria effects on Puerto Rico electric power infrastructure. IEEE Power Energy Technol Syst J. 6(1):85–94. doi: 10.1109/JPETS.2019.2900293.

Lee C, Tien I. 2018. Probabilistic framework for integrating multiple data sources to estimate disaster and failure events and increase situational awareness. ASCE-ASME J Risk Uncertainty Eng Syst Part A Civ Eng. 4(4):04018042. doi: 10.1061/AJRUA6.0000995.

Lemnitzer A. 2022. Geotechnical reconnaissance of the 2021 Western European floods; [accessed 2023 May 19]. http://www.geerassociation.org/administrator/components/com_geer_reports/geerfiles/Blank.pdf.

Liu J, Chen Y, Chen Y. 2021. Emergency and disaster management-crowd evacuation research. J Ind Inf Integr. 21:100191. doi: 10.1016/j.jii.2020.100191.

McCann WR. 1985. On the earthquake hazards of Puerto Rico and the Virgin Islands. Bull Seismol Soc Am. 75(1):251–262. doi: 10.1785/BSSA0750010251.

McEntire DA. 2007. Local emergency management organizations. In: Rodríguez H, Quarantelli EL, Dynes RR, editors. Handbook of disaster research. New York, NY: Springer New York; p. 168–182. doi: 10.1007/978-0-387-32353-4_10.

Meighan HE, Pulliam J, ten Brink U, López-Venegas AM. 2013. Seismic evidence for a slab tear at the Puerto Rico Trench: Puerto Rico trench slab tear. J Geophys Res Solid Earth. 118(6):2915–2923. doi: 10.1002/jgrb.50227.

Miranda E, Archbold J, Heresi P, Messina A, Rosa I, Robertson I, Mosalam K, Kijewski-Correa T, Prevatt D, Roueche D. 2020. StEER – Puerto Rico earthquake sequence December

2019 to January 2020: Field Assessment Structural Team (FAST) Early Access Reconnaissance Report (EARR); [accessed 2022 Nov 30]. https://www.designsafe-ci.org/data/browser/public/designsafe.storage.published/PRJ-2712.

Miura H, Aridome T, Matsuoka M. 2020. Deep learning-based identification of collapsed, non-collapsed and blue tarp-covered buildings from post-disaster aerial images. Rem Sens. 12(12):1924. doi: 10.3390/rs12121924.

Morales A, Hughes S, Lang K, Rivera-Hernandez F, Vargas P, Lozano JM, Karantanellis E, Kassem M, Gomberg D, Plescher R. 2022. Geotechnical impacts of Hurricane Fiona in Puerto Rico. GEER Association (Report-81); [accessed 2023 May 10]. https://geerassociation.org/components/com_geer_reports/geerfiles/GEER_HurricaneFiona_report.pdf.

Mueller C, Frankel A, Petersen M, Leyendecker E. 2010. New seismic hazard maps for Puerto Rico and the U.S. Virgin Islands. Earthq Spectra. 26(1):169–185. doi: 10.1193/1.3277667.

Mukhopadhyay B, Bhattacherjee B. 2015. Use of information technology in emergency and disaster management. Am J Environ Prot. 4(2):101–104. doi: 10.11648/j.ajep.20150402.15.

NOAA. 2017. Hurricane Maria advisory archive. https://www.nhc.noaa.gov/archive/2017/MARIA.shtml.

Nuti C, Vanzi I. 2003. To retrofit or not to retrofit? Eng Struct. 25(6):701–711. doi: 10.1016/S0141-0296(02)00190-6.

O'Rourke T, Toprak S. 1997. GIS assessment of water supply damage from the Northridge earthquake. In: Post-earthquake reconstruction strategies: NCEER-INCEDE center-to-center project; p. 109–122.

Pasch RJ, Penny AB, Berg R. 2019. Hurricane Maria tropical cyclone report. National Hurricane Center (NHC) Report No.: AL152017. https://www.nhc.noaa.gov/data/tcr/AL152017_Maria.pdf.

PRSN. 2022. Puerto Rico seismic network. https://redsismica.uprm.edu/english/our_work/instrumentation.php.

Rajapaksha RADV, Siriwardana CSA. 2023. A systematic review on different approaches used in the development of fragility curves for buildings. In: Dissanayake R, Mendis P, Weerasekera K, De Silva S, Fernando S, Konthesingha C, editors. 12th International Conference on Structural Engineering and Construction Management; Vol. 266 (Lecture Notes in Civil Engineering). Singapore: Springer Nature Singapore; p. 407–426. [accessed 2023 May 12]. doi: 10.1007/978-981-19-2886-4_29.

Rathje EM, Dawson C, Padgett JE, Pinelli J-P, Stanzione D, Adair A, Arduino P, Brandenberg SJ, Cockerill T, Dey C, et al. 2017. DesignSafe: new cyberinfrastructure for natural hazards engineering. Nat Hazards Rev. 18(3):06017001. doi: 10.1061/(ASCE)NH.1527-6996.0000246.

Resch B, Usländer F, Havas C. 2018. Combining machine-learning topic models and spatio-temporal analysis of social media data for disaster footprint and damage assessment. Cartogr Geogr Inf Sci. 45(4):362–376. doi: 10.1080/15230406.2017.1356242.

Roueche D. 2020. Virtual Assessment Structural Team (VAST) handbook: data enrichment and quality control (DE/QC) for StEER earthquake rapid evaluation form. National Science Foundation (NSF), Structural Extreme Events Reconnaissance (StEER). https://www.steer.network/challenge-page/cc17aae5-6fae-4a55-8ad9-0c55300071bf

Roueche D, Cleary J, Gurley K, Marshall J, Pinelli J-P, Prevatt D, Smith D, Alipour A, Angeles K, Davis B, et al. 2018. StEER – Hurricane Michael: Field Assessment Team 1 (FAT-1) Early Access Reconnaissance Report (EARR); [accessed 2021 May 10]. https://www.design-safe-ci.org/data/browser/public/designsafe.storage.published/PRJ-2111.

Rouhanizadeh B, Kermanshachi S, Dhamangaonkar VS. 2020. Reconstruction of critical and interdependent infrastructure due to catastrophic natural disasters: lessons learned. In: Construction Research Congress 2020. Tempe, AZ: American Society of Civil Engineers; p. 895–904. [accessed 2023 May 15]. doi: 10.1061/9780784482858.096.

Scaioni M, Longoni L, Melillo V, Papini M. 2014. Remote sensing for landslide investigations: an overview of recent achievements and perspectives. Rem Sens. 6(10):9600–9652. doi: 10.3390/rs6109600.

Scawthorn C, Blais N, Seligson H, Tate E, Mifflin E, Thomas W, Murphy J, Jones C. 2006. HAZUS-MH flood loss estimation methodology. I: overview and flood hazard characterization. Nat Hazards Rev. 7(2):60–71. doi: 10.1061/(ASCE)1527-6988(2006)7:2(60).

Shinozuka M, Feng MQ, Lee J, Naganuma T. 2000. Statistical analysis of fragility curves. J Eng Mech. 126(12):1224–1231. doi: 10.1061/(ASCE)0733-9399(2000)126:12(1224).

Silva-Tulla F. 2018. Geotechnical impacts of Hurricane Maria in Puerto Rico; [accessed 2022 Nov 30]. http://geerassociation.org/administrator/components/com_geer_reports/geerfiles/180629_GEER_PR_Report_No_GEER-057.pdf.

The European Commission, United Nations Development Group, World Bank. 2013. Post-disaster needs assessments volume A. https://www.gfdrr.org/en/publication/post-disaster-needs-assessments-guidelines-volume-2013.

Tien I, Musaev A, Benas D, Ghadi A, Goodman S, Pu C. 2016. Detection of damage and failure events of critical public infrastructure using social sensor big data. In: Proceedings of the International Conference on Internet of Things and Big Data. Rome: SCITEPRESS-Science and Technology Publications; p. 435–440. [accessed 2023 May 19]. doi: 10.5220/0005932104350440.

Urbina E, Wolshon B. 2003. National review of hurricane evacuation plans and policies: a comparison and contrast of state practices. Transp Res Part Policy Pract. 37(3):257–275. doi: 10.1016/S0965-8564(02)00015-0.

USGS. 2018. U.S. Landsat Analysis Ready Data. Reston, VA Fact Sheet Report No.: 2018–3053. http://pubs.er.usgs.gov/publication/fs20183053.

Vamvatsikos D, Kouris L, Panagopoulos G, Kappos A, Nigro E, Rossetto T, Lloyd T, Stathopoulos T. 2010. Structural vulnerability assessment under natural hazards: a review. p. 17–18.

Wald DJ, Worden BC, Quitoriano V, Pankow KL. 2005. ShakeMap manual: technical manual, user's guide, and software guide. Version 1.0. Techniques and Methods Report No.: 12-A1. http://pubs.er.usgs.gov/publication/tm12A1.

Wilkinson MD, Dumontier M, Aalbersberg I, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, et al. 2016. The FAIR guiding principles for scientific data management and stewardship. Sci Data. 3(1):160018. doi: 10.1038/sdata.2016.18.

Willis HH, Narayanan A, Fischbach JR, Molina-Perez E, Stelzner C, Loa K, Kendrick L. 2016. Current and future exposure of infrastructure in the United States to natural hazards. Santa Monica (CA): Rand Corporation (Research Report).

Wu D, Cui Y. 2018. Disaster early warning and damage assessment analysis using social media data and geo-location information. Decis Support Syst. 111:48–59. doi: 10.1016/j.dss.2018.04.005.

Yeum CM, Dyke SJ, Benes B, Hacker T, Ramirez J, Lund A, Pujol S. 2019. Post-event reconnaissance image documentation using automated classification. J Perform Constr Facil. 33(1):04018103. doi: 10.1061/(ASCE)CF.1943-5509.0001253.

Yildirim E, Demir I. 2021. An integrated flood risk assessment and mitigation framework: a case study for Middle Cedar River Basin, Iowa, US. Int J Disaster Risk Reduct. 56:102113. doi: 10.1016/j.ijdrr.2021.102113.

Youd T, Bardet J-P, Bray J. 2000. Kocaeli, Turkey, earthquake of August 17, 1999 reconnaissance report. EERI No. 2000–03.

Yun S-H, Hudnut K, Owen S, Webb F, Simons M, Sacco P, Gurrola E, Manipon G, Liang C, Fielding E, et al. 2015. Rapid damage mapping for the 2015 M_w 7.8 Gorkha earthquake using synthetic aperture radar data from COSMO–SkyMed and ALOS-2 satellites. Seismol Res Lett. 86(6):1549–1556. doi: 10.1785/0220150152.

Zhao N, Liu Y, Hsu F-C, Samson EL, Letu H, Liang D, Cao G. 2020. Time series analysis of VIIRS-DNB nighttime lights imagery for change detection in urban areas: a case study of devastation in Puerto Rico from hurricanes Irma and Maria. Appl Geogr. 120:102222. doi: 10.1016/j.apgeog.2020.102222.

Zhou Y, Banerjee S, Shinozuka M. 2010. Socio-economic effect of seismic retrofit of bridges for highway transportation networks: a pilot study. Struct Infrastruct Eng. 6(1–2):145–157. doi: 10.1080/15732470802663862.