Chapter 14
Using Non-authoritative Sources During Emergencies in Urban Areas

Emily Schnebele, Christopher Oxendine, Guido Cervone, Celso M. Ferreira, and Nigel Waters

Abstract During emergencies in urban areas, it is paramount to assess damage to people, property, and environment in order to coordinate relief operations and evacuations. Remote sensing has become the de facto standard for observing the Earth and its environment through the use of air-, space-, and ground-based sensors. These sensors collect massive amounts of dynamic and geographically distributed spatiotemporal data daily and are often used for disaster assessment, relief, and mitigation. However, despite the quantity of big data available, gaps are often present due to the specific limitations of the instruments or their carrier platforms. This chapter presents a novel approach to filling these gaps by using

E. Schnebele (✉)
GeoInformatics and Earth Observation Laboratory (GEO), Department of Geography and Institute for CyberScience, The Pennsylvania State University, University Park, PA, USA
e-mail: eschnebe@gmu.edu

C. Oxendine
Department of Geography & Environmental Engineering, United States Military Academy, West Point, NY, USA
e-mail: christopher.oxendine@usma.edu

G. Cervone
GeoInformatics and Earth Observation Laboratory (GEO), Department of Geography and Institute for CyberScience, The Pennsylvania State University, University Park, PA, USA

Research Application Laboratory, National Center for Atmospheric Research, Boulder, CO, USA
e-mail: cervone@ucar.edu

C.M. Ferreira
Department of Civil & Environmental and Infrastructure Engineering, George Mason University, Fairfax, VA, USA
e-mail: cferrei3@gmu.edu

N. Waters
GeoInformatics and Earth Observation Laboratory (GEO), Department of Geography and Institute for CyberScience, The Pennsylvania State University, University Park, PA, USA

The University of Calgary, Calgary, Canada
e-mail: nwaters@gmu.edu

M. Helbich et al. (eds.), *Computational Approaches for Urban Environments*, Geotechnologies and the Environment 13, DOI 10.1007/978-3-319-11469-9_14

non-authoritative data including social media, news, tweets, and mobile phone data. Specifically, two applications are presented for transportation infrastructure assessment and emergency evacuation.

Keywords Infrastructure assessment • Evacuation • Remote sensing • Inundation modeling • Social media • Geospatial analysis • Big data

14.1 Introduction

Never in the history of humankind have we known so much about our planet. Never in the history of humankind have we had such easy access to data. Never in the history of humankind has our civilization been so much at risk.

Hazards pose a constant threat to the development and sustainment of our infrastructure and our society. Hazards can be natural, anthropogenic, or technological. They are, respectively, events that naturally occur, events resulting from human activities or accidents, or the catastrophic collapse of infrastructure, such as roads, communication networks, or power grids, which are needed for our society to function.

A single catastrophic event can claim thousands of lives; cause billions of dollars of damage; trigger an economic depression that might directly or indirectly affect the entire world; destroy natural landmarks; cause tsunamis, floods, and landslides; render a large territory uninhabitable; and destabilize the military and political balance in a region (Cutter 1993; Alexander 2002; Wisner et al. 2004). Such potential catastrophic consequences are due to the emergence of megacities and the proliferation of nuclear power plants and nuclear waste storage facilities, high dams, and other facilities whose destruction poses an unacceptable risk of global reach (Freudenburg et al. 2008; Casti 2012). Thus, the study of natural hazards and of the processes that govern their occurrence has become a fundamental challenge for the survival of our civilization.

Advances in our ability to observe the Earth and its environment through the use of air-, space-, and ground-based sensors has led to the collection of massive amounts of dynamic and geographically distributed spatiotemporal data. Numerical models are initialized with these high-resolution observations to forecast the future or to simulate the past, generating simulations that can be several orders of magnitude larger than the initial observations. Remote sensing data from air- and space-borne platforms have also become the de facto standard for providing high-resolution information for the assessment, relief, and mitigation of damaged areas during and after emergencies caused by natural, human-made, and technological disasters (Jensen and Cowen 1999; Voigt et al. 2007). However, due to limitations in orbital revisit time, sensor characteristics, and the presence of clouds, there may be gaps in these remote sensing data.

This chapter presents applications for data collected from non-authoritative sources to fill the gaps in remote sensing data during disasters and emergencies. Non-authoritative sources include data volunteered by citizens (also known as volunteered geographic information or VGI (Goodchild 2007; Sui and Goodchild

2011; Sui et al. 2013)) or collected for purposes other than disaster assessment, such as traffic cameras or mobile phone locations. This general class of data, often voluntarily contributed and made available, can consist of pictures, videos, sounds, text messages, etc. Due to the spread of the Internet to mobile devices, an unprecedented and massive amount of data have become available, often geolocated and often in real time. These sources provide a large, rapidly changing, dynamic dataset that not only complements remote sensing observations but also adds an additional, subjective view of how people perceive and react to hazards.

Although non-authoritative data are often published without scientific intent, and usually carry little scientific merit, it is still possible to mine mission critical information. For example, volunteered photos and videos about natural hazards have emerged as a data source during crises and hazardous events to derive local meteorological information, capture and record the physical features of an event, and identify and document flood height (De Longueville et al. 2009; Hyvärinen and Saltikoff 2010; Poser and Dransch 2010). During Hurricane Sandy, geolocated pictures and videos searchable through Google provided early emergency response with ground view information.

Mining these massive amounts of "big data," it is possible to reconstruct a spatiotemporal human terrain that provides knowledge when remote sensing data are unavailable or incomplete. Additionally, non-authoritative data may provide unique knowledge that is not possible to acquire solely from remote sensing instruments.

This chapter discusses the fusion of remote sensing and non-authoritative sources to assess road infrastructure and plan evacuations in an urban environment during emergencies. Two specific applications are discussed:

1. An assessment of New York City transportation infrastructure during and after Hurricane Sandy using crowdsourced remote sensing imagery, numerical models, social media, and ground observations
2. Identification of evacuation routes during emergencies in New York City using traffic information and mobile phone data

14.2 Transportation Infrastructure Assessment

The first application presented in this chapter is a damage assessment of roads during and after Hurricane Sandy in New York City. Multiple sources of data are combined including aerial images contributed by the Civil Air Patrol (CAP), numerical inundation models, VGI harvested from social media, and ground observations.

The utilization of data from multiple sources can help provide a more complete description of a phenomenon. For example, data fusion is often employed with remote sensing data to combine information of varying spatial, temporal, and spectral resolutions as well as to reduce uncertainties associated from using a single source (Zhang 2010). The fused data then provides new or better information than would be available from a single source (Pohl and Van Genderen 1998).

The incorporation of multiple data sources or methods for improved performance or increased accuracy is not limited to the field of remote sensing. Boosting, a common machine learning technique, has been shown to be an effective method for generating accurate prediction rules by combining rough, or less than accurate, algorithms together (Freund et al. 1999). While the individual algorithms may be singularly weak, their combination can result in a strong learner. Furthermore, redundancies in observations provide an increase in the confidence of observations or estimates, while data from multiple sources can provide information when they might not do so if used in isolation.

Figure 14.1 shows the stacked layer approach used to fuse heterogeneous data at different spatial, temporal, and radiometric resolutions. The data are first processed in a GIS environment by resampling them at the highest resolution available using

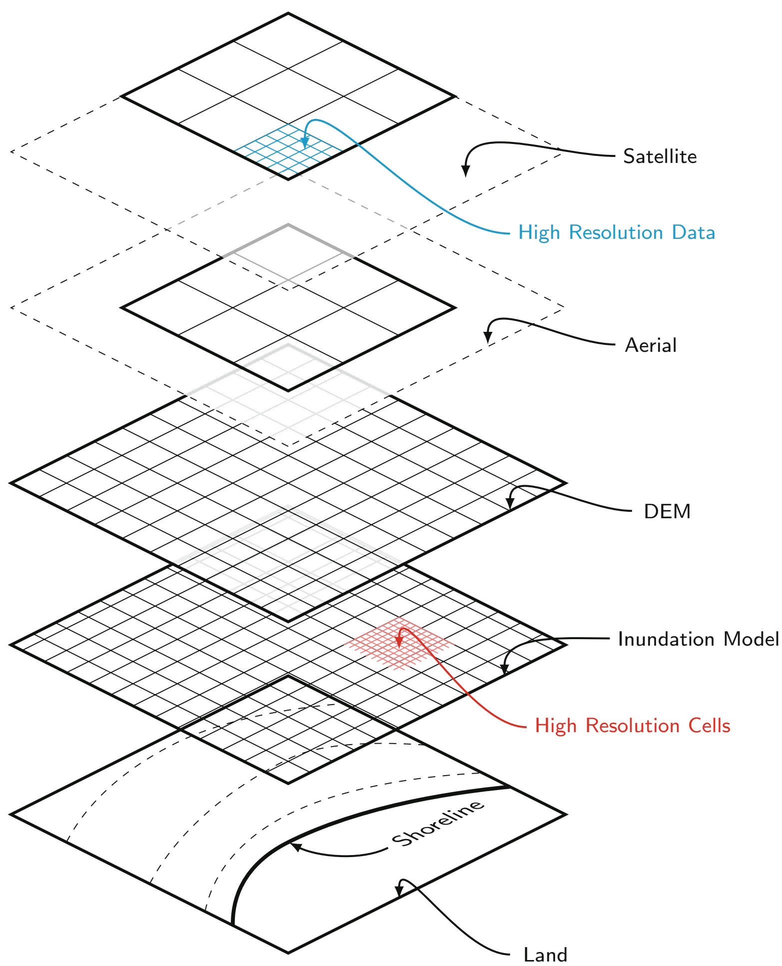

Fig. 14.1 Stacked layer approach used in the methodology. Data comes in different formats and with different resolutions

spatial statistical algorithms (Zhang et al. 2014). For example, points which identify flooding (e.g., photos) are plotted, georeferenced, and then smoothed using a kernel interpolation to create a GIS layer of estimated flood extent. This is a task performed for each data source, resulting in multiple, individual flood extent layers. The analysis is then performed on the fused layers by applying statistical or machine learning algorithms to classify the data and identify anomalies.

Fusing data from multiple sources leads to an improved damage estimation and an increased understanding of the sequence of events that leads to transportation infrastructure failure. In this example, non-authoritative data are used in two different scenarios:

1. Damage assessment during an event
2. Damage assessment after the event

It is assumed that ground truth data might not be available. The novelty of this study is the development of a methodology that takes advantage of "citizens as sensors" (Goodchild 2007) and of various other data, including remote sensing and numerical models, not necessarily designed to be used during emergencies to improve damage assessment. These non-authoritative or nontraditional sources are used to create additional layers which augment traditional sources when they may be lacking or incomplete. The result is shown in the bottom layer, where a flood hazard map is generated. The resulting flood hazard map is then paired with a high-resolution road network to create a road damage map.

14.2.1 Data Sources

14.2.1.1 Remote Sensing Data

High-resolution remote sensing data are routinely used to assess damage during and after hazards, in both urban and rural areas. Two or more images are acquired for an area showing the differences before and after the hazard.

The Civil Air Patrol, the civilian branch of the US Air Force, was tasked with collecting aerial photos of the US East Coast following the impact of Hurricane Sandy in October 2012. Within days of the storm making landfall, hundreds of missions were flown by volunteers from Cape Cod, MA, to Cape May, NJ. From these missions, thousands of aerial photos of the coastline were generated, including those documenting heavily flooded areas.

14.2.1.2 Numerical Surge Model

Recent improvements in understanding the physics of storm surge combined with rapid increases in High Performance Computing (HPC) power have led to the development of physics-based, high-resolution numerical models capable of predicting and simulating hurricane storm surge with reasonable accuracy in coastal areas.

The Sea, Lake, and Overland Surge from Hurricane model (SLOSH) (Jelesnianski et al. 1992) developed by the National Weather Service (NWS) has been extensively used by decision makers to predict storm surge inundation for planning and emergency management, and it is currently the NWS official operational forecast model for storm surge. Several other numerical models have been developed over the years to calculate water levels and currents resulting from hurricane storm surges along the continental shelves and coasts.

Among others are the Advanced Circulation (ADCIRC) model developed by Luettich and Westerink (2004), the fully nonlinear Finite Volume Coastal Ocean Model (FVCOM) developed by Chen et al. (2003), and the Semi-implicit Eulerian-Lagrangian Finite Element (SELFE) model developed by Zhang and Baptista (2008). Recently, the Coastal and Ocean Modeling Testbed (COMT) compared the models' prediction skills (Kerr et al. 2013) and concluded that they all, except SLOSH, generated similar predictions for Hurricane Ike in 2008 and Hurricane Rita in 2005, thus demonstrating the maturity level of storm surge model development.

For this study, a lower-resolution/faster computational time numerical mesh was used to simulate the Hurricane Sandy storm surge in order to maintain similarity to models used in operational forecasts (e.g., Advanced Flooding Guidance System [ASGS]) to represent information that would be available to decision makers before a hurricane landfall. The coupled version of the two-dimensional depth integrated version of the Advanced Circulation (ADCIRC) model and the SWAN wave model (Dietrich et al. 2011) was used to simulate hurricane storm surge along the coast. The ADCIRC model (Luettich and Westerink 2004) is a finite element, shallow water model that solves for water levels and currents at a range of scales and is widely used for storm surge modeling (e.g., Ferreira et al. 2014). This version of the program solves the Generalized Wave Continuity Equation (GWCE) and the vertically integrated momentum equations. SWAN is a third-generation spectral wave model that computes random, short crested wind-generated waves and wave transformation in the near shore and inland waters (Booij et al. 1999). For storm surge simulation, ADCIRC is forced by the wind and pressure fields and the wave radiation stress resulting from the wave model. Tides and river inflow can also be added as a boundary.

The East Coast Mesh (ECM2001) presented by Mukai et al. (2002) was utilized with approximately 250,000 nodes and a resolution of approximately 1.2 km along the study area. ADCIRC allows for the use of an unstructured finite element mesh with variable resolution along the model domain. The hurricane surge model was forced by wind and pressure fields developed by a parametric asymmetric wind model (Mattocks and Forbes 2008) that computes wind stress, average wind speed, and direction inside the Planetary Boundary Layer (PBL) based on the National

Hurricane Center (NHC) Advisory Archive for Hurricane Sandy (NOAA 2013a) track data and meteorological conditions (e.g., central pressure, forward speed, and radius to maximum wind).

To simulate the Hurricane Sandy storm surge, a simulation was run for October 18th until the 28th including tides (tidal potential components M2, S2, N2, K2, K1, O1, and Q1) but neglecting river inflows. The simulations were performed under the HPC environment provided by the Extreme Science and Engineering Discovery Environment (XSEDE) supported by the National Science Foundation (NSF). Results were recorded every 15 min around the study region at every model node and at NOAA Tides and Currents stations (NOAA 2013b). The model results generally overestimate the measured water levels most likely due to the differences between the hypothetical asymmetrical wind and pressure fields and the actual storm conditions.

The spatial flood levels were calculated using a Digital Elevation Model (DEM) with a 1 arc-second resolution from the National Elevation Dataset (NED) for the study region (USGS 2013). The maximum water levels for each model node were extracted for the 29th and the 30th of November and converted to the NAVD88 vertical datum. A spline interpolation with tension was applied to create a maximum water level surface for the study region according to the methodology suggested by Berenbrok et al. (2009). Finally, the water levels were subtracted from the DEM to calculate the spatial flooded extent (Fig. 14.2).

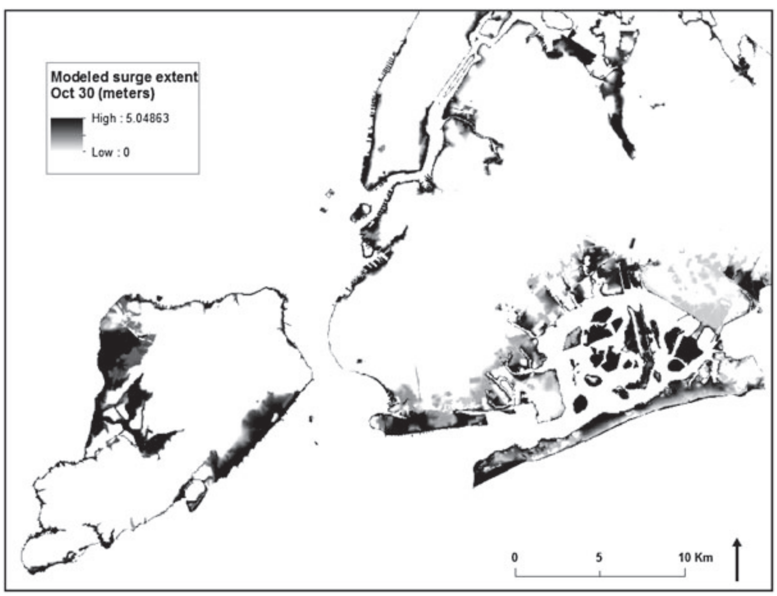

Fig. 14.2 Modeled surge extent for October 30 at 1:00pm

14.2.1.3 Non-authoritative Data

Non-authoritative data are not produced or distributed from necessarily trusted sources and often lack any assertion of verification or accuracy. However, regardless of varying levels of certainty or trustworthiness, non-authoritative sources can provide valuable, real-time, on-the-ground information during disasters when traditional sources are unavailable, lacking, or slow to respond. For example, following the Fukushima nuclear disaster in 2011, the Japanese public supplemented authoritative government sensors with user-generated content. Individuals throughout the country bought personal Geiger counters and contributed to a crowdsourced Geigermap.[1]

Crowdsourced Damage Assessment

Remote sensing data (photos) acquired by the Civil Air Patrol were assessed for damage by thousands of people across the world. The photos were placed on a Hurricane Sandy Google Crisis Map website (Fig. 14.3) for the public to assess visible damage through a crowdsourcing portal supported by MapMill. This yielded a large damage assessment dataset generated from crowdsourced, non-authoritative, nontraditional sources. The photos were also made available online through a Federal Emergency Management Agency (FEMA) website for residents to search by street address to see what, if any, damage their homes may have sustained.

The crowdsourced damage assessments of photos captured between October 31 and November 11, 2012 for the area from 33N to 26N latitude and 90W to 84W longitude were downloaded directly from MapMill. Because of the large volume of photos and the scale of the domain, the photos were aggregated into a 500 m grid structure. The value for each grid point is a function of the number of images present in each grid and their average crowdsourced damage assessment. As a result, each grid has a value from 1 to 10, with 1 representing no damage and 10 severe damage/flooding.

Volunteered Geographic Information (VGI)

VGI was ascertained from YouTube videos which documented flooding and damage in New York City following Hurricane Sandy. The data were collected from a Hurricane Sandy Google Earth website where YouTube videos were supplied by Storyful. YouTube, a video-sharing website, is utilized by millions of people for the sharing of videos covering a wide range of topics and experiences. Through this site the public voluntarily shares information, often documenting damage resulting from natural hazards. The videos were provided with geolocated information and were visually assessed by the authors. The small number of videos ($n = 15$) did

[1] http://japan.failedrobot.com/

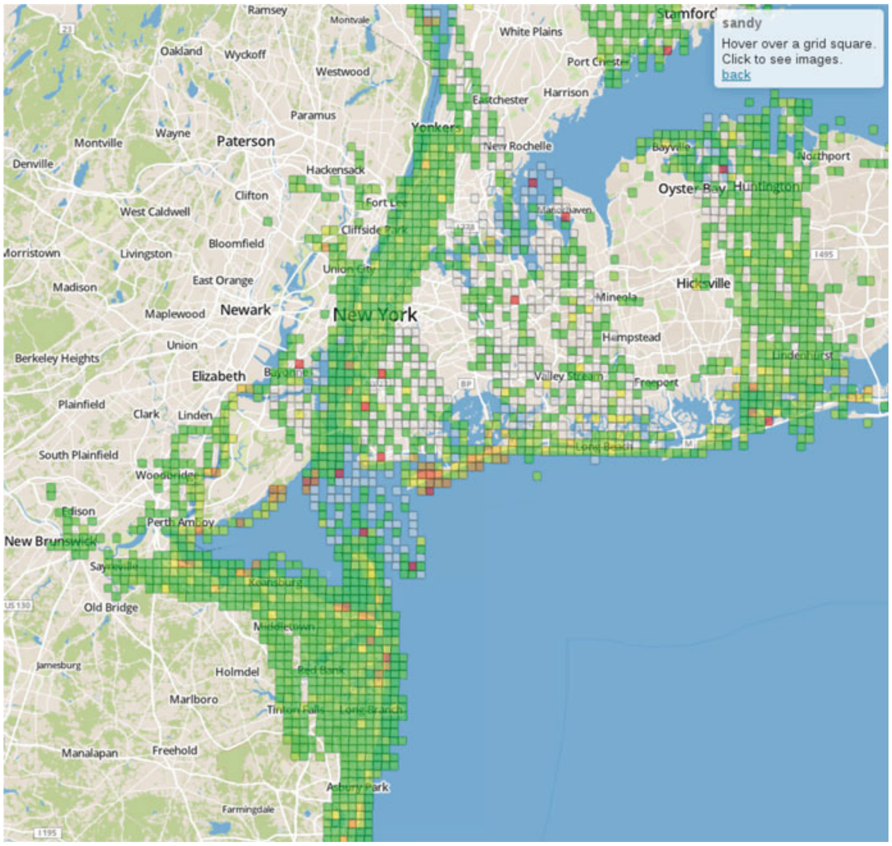

Fig. 14.3 Crowdsourced assessments for the Civil Air Patrol data. Damage assessment: *red* = high, *yellow* = medium, *green* = none

not require any crowdsourcing or automated assessment. Furthermore, it is shown in Schnebele and Cervone (2013) that even a small number of properly located VGI data can help improve flood assessment. Each location corresponding to a video point was assigned a value of 10 (severe damage/flooding).

Photos ($n = 25$) which documented flooding within the study domain were downloaded using the Google search engine and were also visually assessed by the authors. The point locations were georeferenced to create a GIS layer of flooded locations. Each point was assigned a value of 10 (severe damage/flooding).

Twitter, a popular social networking site, is often utilized by the public to share information about their daily lives through micro-blogging. Arizona State University's TweetTracker provided Twitter data for this project (Kumar et al. 2011). Tweets generated in the New York City area extending from 40.92N to 40.54N latitude and 73.75W to 74.13W longitude from October 26 to November 3, 2012 containing the word "flood" were used to provide a temporal framework.

14.2.1.4 Authoritative Data

Authoritative data are collected, produced, and managed by professional cartographers, geographers, and/or government agencies. Information which comes from these official, authoritative persons or agencies carries a certain level of trust which affords them credibility (Flanagin and Metzger 2008; Goodchild and Glennon 2010). Examples of authoritative data may include remote sensing imagery collected and calibrated by NASA or stream flow information collected from USGS river gauges. These are our traditional sources of data and information during disasters and emergencies.

Federal Emergency Management Agency (FEMA)

The FEMA Modeling Task Force (MOTF) consists of experts in hazard assessment and the modeling of hazard losses. Following Hurricane Sandy, FEMA MOTF used field-verified high water marks and storm surge sensor data to create storm surge maps for the US East Coast. For this work, a FEMA MOTF storm surge shapefile for New York City was downloaded from FEMA's GeoPlatform website. The surge map was the finalized version (dated February 14, 2013) with a 1 m horizontal resolution and a New York State Plane coordinate system (Fig. 14.4 (right)).

Water depth data were also collected at inundated New York City public schools by FEMA MOTF. The water depth at schools was ascertained from water marks taken from on-site structures (Fig. 14.4 (left)). A GIS layer was created from georeferenced point locations of the schools with measured water depths.

United States Geological Survey (USGS)

Water height collected by the USGS storm-tide monitoring provides an additional source of authoritative ground data. These official measurements were taken at

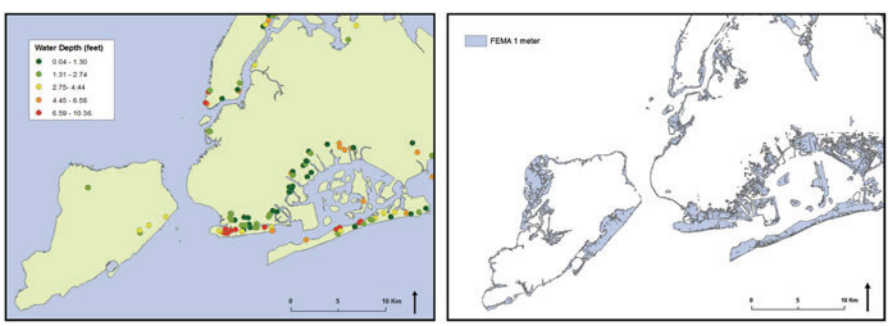

Fig. 14.4 FEMA water depth measured at public schools (*left*) and official flood inundation map (*right*)

different locations throughout the domain. Water height modeled for each point is interpolated using a spline function to create a water height surface. A DEM with a 1 arc-second resolution from the National Elevation Dataset (NED) is subtracted from the water height surface to create a water depth layer (USGS 2013).

United States Census Bureau

A 2012 TIGER/line® shapefile of road networks for the New York City area was downloaded from the US Census Bureau and was georeferenced to New York State Plane coordinates.

14.2.2 Damage Assessment During Emergencies

After individual data layers are generated from available remote sensing and authoritative and non-authoritative data, they are integrated together using an artificial neural network machine learning algorithm. Artificial neural networks are nonlinear data modeling tools for discovering patterns in data from a series of inputs (Atkinson and Tatnall 1997). The network consists of interconnected nodes comprising an input layer, a hidden layer, and an output layer (Fig. 14.5). In this research, the nodes of the input layer consist of the flood identification layers created during preprocessing, and the output layer is a flood assessment surface. The hidden layer nodes, or neurons, are the computational units of the network. The neuron receives the inputs and produces responses. Benediktsson et al. (1990) defines the simplest formal model of the neuron, where the output value is approximated by the function

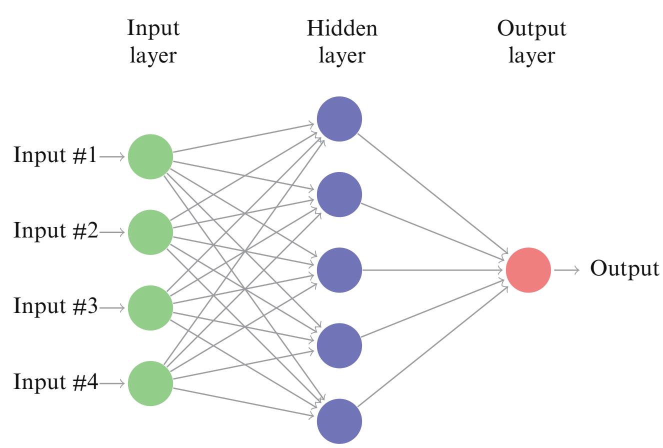

Fig. 14.5 Depiction of an artificial neural network

$$o = K\phi\left(\sum_{j=1}^{n} w_j x_j - \theta\right) \tag{14.1}$$

where K is a constant, ϕ is a nonlinear function, w_j are the weights assigned by the network, and θ is a threshold. The network takes inputs x and produces a response o_i from the output units i. The outputs are either $o_i = 1$ if the neuron i is active for the input x or $o_i = 0$ if it is inactive. The network learns the weights through iterative training and will converge when there is no change from one iteration to the next.

The trained network can then be used for the classification of a new dataset. A feedforward artificial neural network was implemented for this work using the **R** statistical package (Venables and Ripley 2002).

The goal is to classify each pixel as being flooded or not flooded. The neural network classifier is trained using the data layers from October 29 and tested on the October 30 layers. Figure 14.6 illustrates the training of the neural network classifier (blue lines) using available sources of remote sensing and authoritative and non-authoritative data from October 29. The data are first preprocessed (e.g., georeferenced and interpolated) to create individual flood extent estimations which are fed into the neural network to create a classifier. The operational step uses this classifier along with data collected from the subsequent day, October 30. These data (green lines) are preprocessed and then passed through the trained classifier to create a flood extent map.

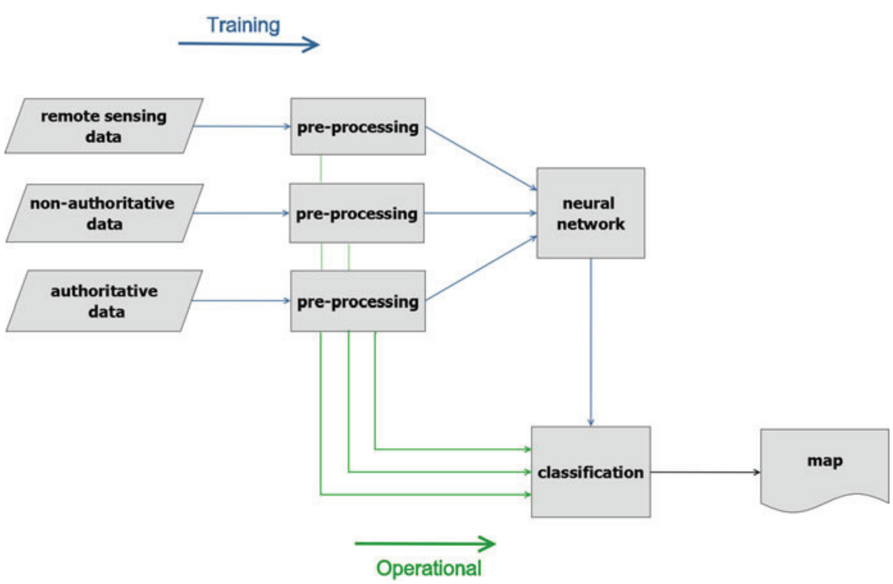

Fig. 14.6 Illustration of the application of a neural network classifier. The classifier is created from training data (*blue lines*) and is then used to create a flood extent map by passing data from a subsequent day (*green lines*) through the classifier

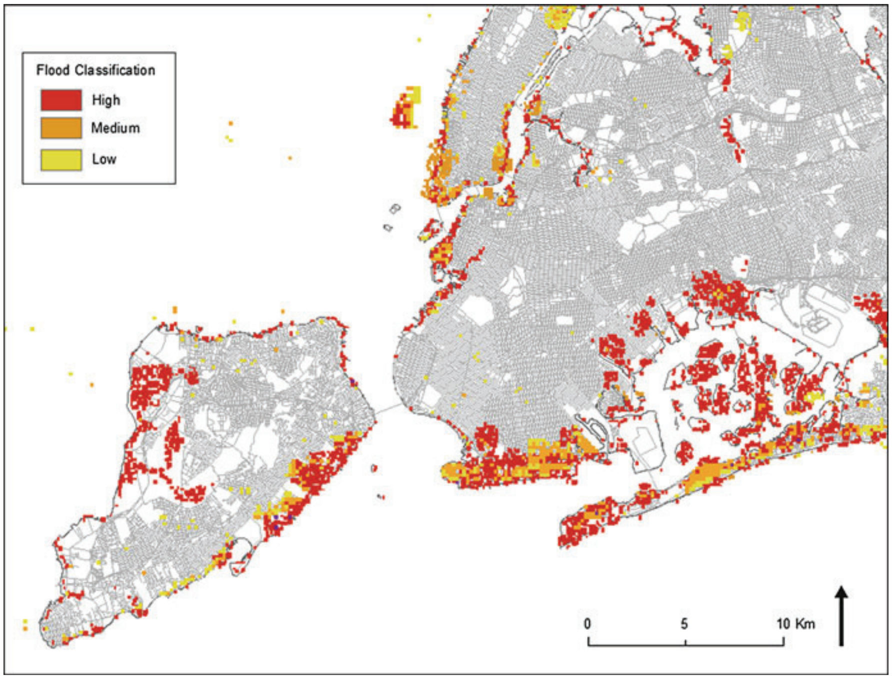

Fig. 14.7 Classification of flooding (high, medium, low flood severity/damage) in New York using an artificial neural network

Because the inundated schools, USGS, and Civil Air Patrol data represented maximum flood extent, it was possible to generate only one layer from each dataset; therefore, these data were used for both days. The initial training and testing datasets produced results indicating flooding along the coastlines of New York City with the greatest damage identified in lower Manhattan and southern edges of Brooklyn and Queens (Fig. 14.7).

14.2.3 Damage Assessment After Emergencies

After an emergency, remote sensing and volunteered data can be employed to provide a damage assessment. In this particular work, the official FEMA flood map is color coded to show not only which areas have been flooded but also which areas have been most affected. In addition, the damage assessment surface is then used to identify roads which may be compromised or may require site inspections (Schnebele et al. 2013).

Crowdsourced data (CAP photos) and VGI (YouTube videos), which are illustrated in Fig. 14.8b, are fused together using a kriging interpolation. Kriging allows for spatial correlation between values (i.e., locations/severity of flooding)

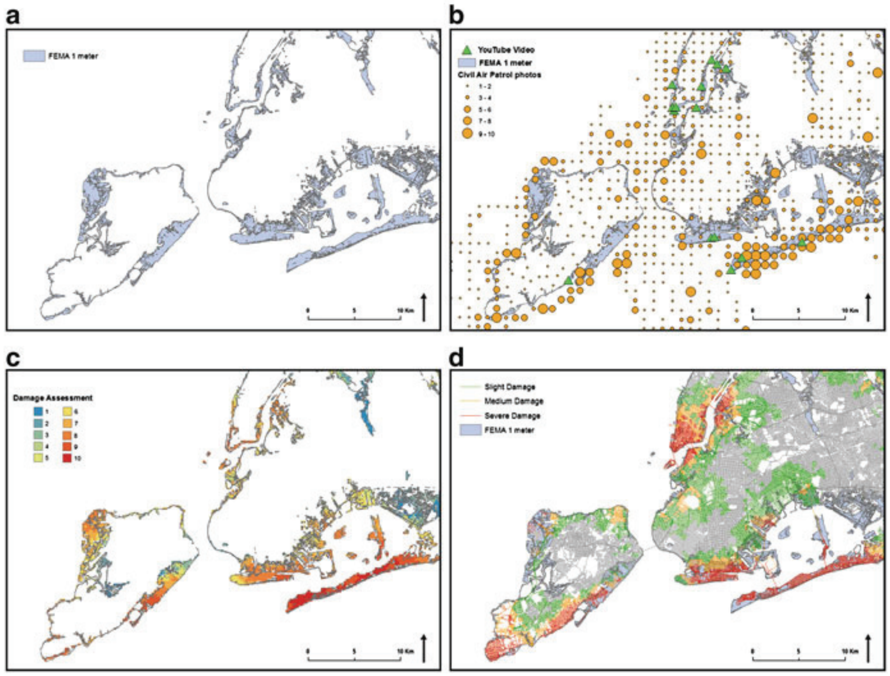

Fig. 14.8 Storm surge extent generated by FEMA and the locations of Civil Air Patrol photos and geolocated videos (**a** and **b**). Flood damage assessment generated from non-authoritative data and the subsequent classification of potential road damages (**c** and **d**)

to be considered and is often used with Earth science data (Oliver and Webster 1990; Olea and Olea 1999; Waters 2009). Ordinary kriging generated a strong interpolation model. Cross-validation statistics yielded a standardized mean prediction error of 0.0008 and a standardized root-mean-squared prediction error of 0.9967. Figure 14.8c illustrates the damage assessment, with values ranging from 1 (no damage) to 10 (severe damage), created from the interpolated surface which is clipped to the boundaries of the FEMA surge extent (Fig. 14.8a) demonstrating how non-authoritative sources can be used to add value to the FEMA map.

Ground information in the form of geolocated videos (Fig. 14.9) enhances the non-authoritative dataset by providing flood information not conveyed in the CAP photos. As illustrated in Fig. 14.8b, the locations of the videos (green triangles) did not coincide with the locations of photos rated as medium/severe damage (larger orange circles, values 7–10). Reasons for this disparity may include flooding captured on video had receded before the Civil Air Patrol flights or were captured at night or flooding may have occurred in areas which were not in a flight path or were unable to be seen from aerial platforms (i.e., flooding in tunnels, under overpasses). By using multiple data sources, flood or damage details not captured by one source can be provided by another.

Fig. 14.9 Example of YouTube video documenting flooding

Overall, there is a very good agreement between the flood extent from FEMA and the assessment generated with the proposed methodology. Figure 14.10 provides examples of agreement between photos identifying flooding/damage and the FEMA-generated flood extent, while Fig. 14.11 includes examples where the locations of flooding or damage did not agree between the Civil Air Patrol and the FEMA data. These areas were located along coastal edges, and therefore a lack of spatial precision in the data is most likely the cause of the discrepancies.

Sources of error in non-authoritative data, such as incorrect information (false positive/negative) or improper geolocation, needed to be considered. Incorrect information can be mitigated by including visually verified photos/videos and the application of multiple sources. Crowdsourcing, in particular, can increase accuracy and enhance information reliability compared to single-source observations (Giles 2005). Geolocation errors can be reduced with automation.

Sparse data or data skewed in favor of densely populated or landmark areas makes the use of non-authoritative data sources especially challenging. Increasing data volume and integrating authoritative data into the methodology can yield increased confidence and include underrepresented areas. Table 14.1 compares and summarizes some features of each type of data. Although non-authoritative data can provide timely, local information often in large volume, they are often viewed with uncertainty. Conversely, the verification and authentication of authoritative data yield trusted results at the cost of time.

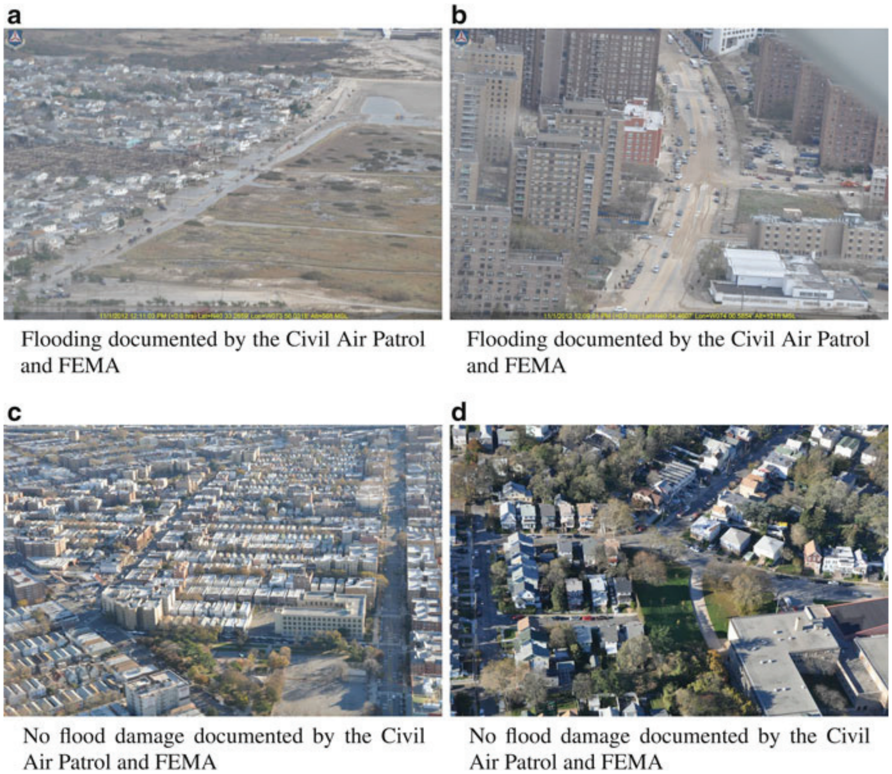

Fig. 14.10 Agreement between Civil Air Patrol photos and FEMA evaluation for flooded (**a** and **b**) not flooded (**c** and **d**)

14.2.3.1 Road Damage Map

In Fig. 14.8c, the damage assessment is limited to the FEMA-generated surge extent for the sake of comparison. For the classification of road damage, the non-authoritative assessment is not limited by the FEMA boundary. The fusion of the non-authoritative data predicted flooding and damage outside the FEMA surge boundary, so the full damage assessment was utilized for the road classification. A road network from the TIGER/line® shapefile was layered over the damage assessment surface. Road damage was then classified based on the underlying damage assessment (Fig. 14.8d).

By using the damage assessment surface along with a high-resolution road network layer, roads which may have severe damage can be identified at the street level. This allows authorities to prioritize site inspections, task additional aerial data collection, or identify routes which may be compromised. The identification of potential damage to transportation infrastructure is also crucial to the planning of evacuation routes during and after emergencies.

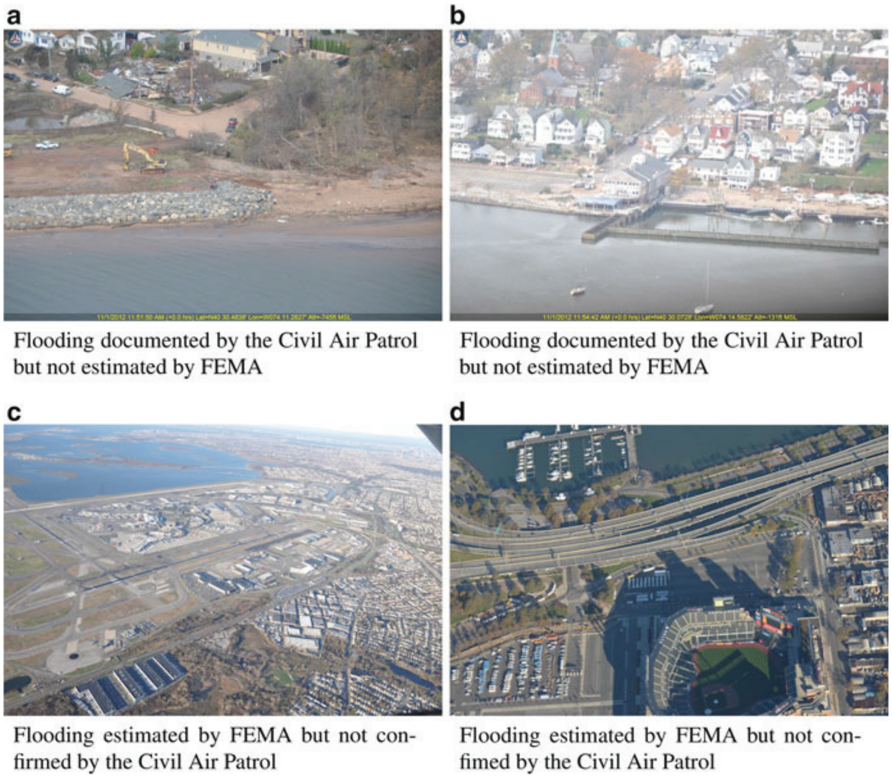

Fig. 14.11 Disagreement between Civil Air Patrol photos and FEMA evaluation for flooded (**a** and **b**) not flooded (**c** and **d**)

Table 14.1 Comparison between non-authoritative and authoritative data

	Non-authoritative data	Authoritative data
Benefits	Volume	Reliable
	Real time	Verified
	Citizens as sensors	Authenticated
Challenges	Sampling bias	Slow
	Unconfirmed	Unavailable

14.2.3.2 Temporal Assessment

For this study, Twitter data were used to provide a temporal rather than spatial assessment. Although tweets were geolocated using TweetTracker, uncertainty in their location did not allow for a study at a street resolution. However, they provide precise temporal information that can be used to understand the progression of the surge extent over time. To understand the temporal progression is crucial during and after flood events and is very hard to understand using remote sensing instruments, due to their inherent carrier limitations. Twitter data can effectively be used to

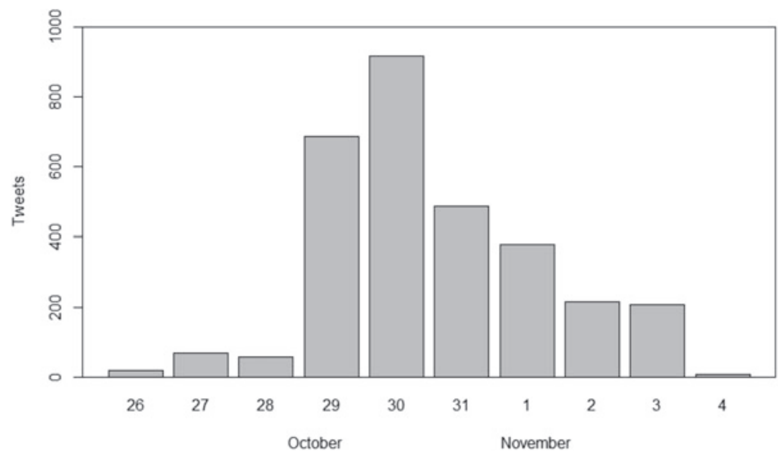

Fig. 14.12 Progression of tweets mentioning the word "flood" in the New York City area

overcome this limitation because of its high temporal resolution. For example, the peak in the number of tweets containing the word "flood" occurs on October 29 and 30, 2012 immediately before and during landfall of Hurricane Sandy in New York City the night of October 29th (Fig. 14.12). Having an indication of event timing can be a very effective aid for emergency managers and response initiatives.

14.3 Evacuations During Emergencies

Emergency evacuations during natural and anthropogenic disasters are time sensitive and require detailed spatiotemporal information for emergency planners to mitigate evacuee risk. One of the primary goals of emergency personnel is to reduce the likelihood of injury and death for citizens within an evacuation zone. First responders are entrusted with providing accurate evacuation information to the public, especially to special populations (e.g., children, elderly, and disabled people are at increased risk during evacuations). Fusing non-authoritative and authoritative information provides increased situational awareness for crisis response personnel. This information enables them to better understand where incidents are occurring, how many people are at risk, and where to allocate resources. The ability to minimize response time is critical to minimizing loss of life and limb.

This section discusses the importance and challenges of evacuations, mobile phone data and its collection, issues of privacy with mobile phone data, and how the use of mobile phone data in evacuations can enhance situational awareness of emergency response planners.

Emergency response personnel can use non-authoritative data, such as near real-time mobile location data, to increase their situational awareness and to redirect resources as needed within their jurisdiction. Figure 14.13 shows a subset of

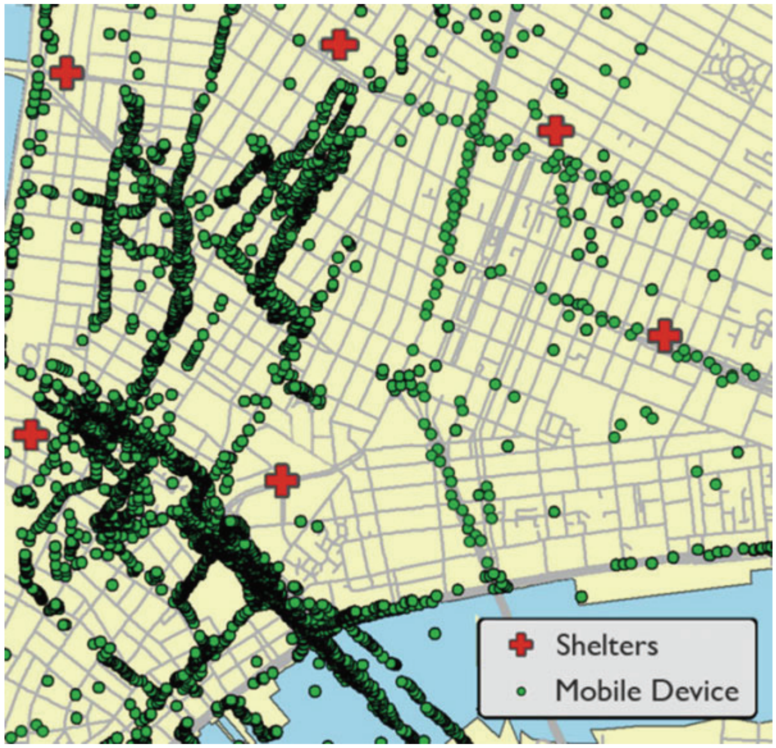

Fig. 14.13 Emergency shelters and example of mobile phones geolocated in New York City. Emergency response personnel can utilize mobile phone data to answer unknown questions. For example, where are citizens located within the evacuation zone, or are there areas where additional police forces should be positioned to reduce congestion?

mobile phone data (over 15,000 phones) collected from OpenCellID in 2011 for NYC as well as the locations of five emergency shelters (OpenCellID 2011). In this example, mobile phone locations provide enhanced situational awareness of spatial and temporal population fluctuations, unlike census data which is a static representation of data collected during the last census. Using mobile phone data, as shown in Fig. 14.13, emergency response personnel can determine evacuation routes to the closest shelters, determine where to position police officers to reroute traffic to reduce congestion, or estimate how many citizens are at risk in an evacuation zone.

14.3.1 Importance of Evacuations

Emergency evacuations often occur with limited or no planning. Notifying the correct individuals based on their location and the type of risk is key to minimizing loss of life and limb. Emergency response personnel manage and execute evacuations. They collect information from various sources and share this information with first responders and other emergency response personnel through a common operational picture.

Much of the information collected and required during a crisis have a spatial component. Analyzing the area of impact for a crisis event provides emergency responders with a spatial footprint of the affected area. Further spatial analysis may provide estimated damage costs and an estimation of the number of people affected and their location. Emergency planners use such information to evaluate where to allocate resources in order to minimize further loss of life, limb, or property.

14.3.2 Challenges of Evacuations

At the onset of an evacuation, emergency planners have many questions that are inherently spatial: How many people were affected by the incident, and where are they located? Where are the available emergency shelters? What is the status of the transportation network including public transportation? What is the status and location of available emergency personnel and equipment? Much of this information is spatial in nature; however, it may be found in disparate databases or systems. Quickly finding answers to these questions can save lives; however, limited knowledge can lead to increased risk to evacuees and response personnel.

One key challenge in evacuations is determining where the people are located. Determining the location of individuals and the risks they may be exposed to is of the utmost importance for planning an emergency response. Populations fluctuate during the day as people travel to work, school, recreational activities, sporting/cultural events, etc. Their location is constantly changing and not always predictable. Currently, the allocation of resources is based on the expertise of first responders, information that has been collected about the situation, and situational awareness of emergency planners.

14.3.3 Mobile Phone Data

During emergencies, traditional mobile phones provide voice and SMS services; however, smartphones also provide enhanced services through mobile applications. Some of these applications include social media networks, notification services, mapping, navigation, e-mail, Internet access, photo and video capture, and crowdsourcing applications. Individuals use these to improve their situational awareness and share information with others who may be affected by the crisis.

Mobile phone companies collect information from individual mobile phones in the United States and internationally. Policies for how long the data is stored vary by mobile provider. Mobile phone data that are collected may include date, time, latitude, longitude, identification number of the cell phone, and signal strength. After these data are collected, they may be aggregated for additional analysis, cleansed of personally identifiable information, and sold for use in other applications, such as navigation, traffic services, or business development.

For example, mobile phone data are often used to improve navigation by monitoring congestion along interstates and highways. The average speed of mobile phones is compared with the speed limit along a section of highway to determine which routes are flowing normally and which are congested. Older models of mobile phones determine their location by triangulating between mobile phone towers, whereas most modern phones now use GPS to determine their location. This information provides individuals with the ability to find the shortest or quickest route to their destination using mobile phone data. Further analysis of mobile phones that are traveling along highways provides a method to estimate the number of vehicles traveling along a route at different time periods. This information is useful for transportation departments for traffic planning, for businesses for determining locations for new franchises, and for numerous other applications.

14.3.4 Issues of Privacy with Mobile Location Data

Although the collection of near real-time mobile location data to support emergency response greatly increases situational awareness of response personnel and planners; citizens are often concerned with the government or individuals using such information to invade their privacy, track their movement, or in an investigation following an attack.

Personally identifiable information (PII) includes information such as name, social security number, phone number, home address, etc. PII is not needed when collecting real-time mobile location data to support emergency response. Protective measures such as stripping PII from data before mobile phone companies share it with government officials during emergencies provide a measure of privacy. Another measure for protecting individual privacy is through the aggregation of data. In metropolitan areas, mobile phone location data could be aggregated to the nearest road intersections with an estimate of the number of people near that intersection. These data would reduce privacy concerns while also providing invaluable data to emergency response personnel. In suburban or rural areas, data could be assigned by census block or tract, or the spatial and temporal accuracy of the data could be reduced to reduce privacy concerns while still providing increased awareness to response personnel.

With the increasing use of mobile phone data in the private sector to support business development, government policies must be developed to address the application of mobile phone data in large-scale emergencies. Policies on the application of mobile phone location data in emergency response are limited to nonexistent. Although citizens are concerned with privacy, they are also concerned with an accurate and efficient response by emergency response personnel.

14.3.5 Application of Mobile Phone Data in Evacuations

Fusing mobile phone data with existing data sources provides increased situational awareness and fills existing gaps in other data. During large-scale dynamic events, especially in high population density areas, the common operational picture is further supplemented with mobile phone data.

One major challenge for emergency response personnel and planners is to determine where people are located at various times of the day. This challenge becomes even more complicated with limited communications or a complete loss of communications. During emergencies with limited to no communications, it is possible for emergency personnel to use archived data to estimate the density of populations. Previously collected mobile data that accounts for diurnal population change and population change during major events, such as sporting events, parades, festivals, etc., provide response personnel with prediction models to use in their planning. This information can prove very important in international response to disasters where population data are often unreliable.

Providing assistance at the right location is often a challenge for emergency response personnel. Often limited crowdsourced or social media data are available in developing countries; however, mobile phones are much more prevalent and provide a source of data on movements of people within the country. Information on the movement of individuals within a country enables the government and nongovernment organizations to prioritize their response based on spatial and temporal data.

As smartphones become more prominent in developing counties, it is necessary to devote increased efforts to increase awareness of mobile applications and their contributions to citizens. These tools provide citizens and emergency personnel with increased information and damage estimates on their response areas.

Finally, mobile phones provide an opportunity for evacuees to provide data to emergency personnel through text messaging, social media, or phone calls. Although there may be a loss of power, mobile phones usually work on battery power for several hours after an incident. This provides an alternate means of communication and enhances situational awareness when other communication methods are limited or nonexistent.

Mobile phones provide emergency personnel with enhanced situational awareness; however, new policies for protecting the privacy of individuals with respect to mobile phone data are necessary to limit concerns for PII. Standards for collection of data and who can handle data are necessary to protect citizens. Additional research is needed to evaluate how mobile phone data is represented in Geographic Information Systems (GIS). This research should address and refine database storage methods and analysis of mobile phone data.

14.4 Conclusions

During disasters and emergencies in urban areas, timely and accurate information of road networks, infrastructure conditions, and locations of citizens is crucial. But this information can be limited, incomplete, lengthy to acquire, or out of date. Although not necessarily created or posted with the intent of being used for scientific research, non-authoritative data can be harvested during disasters and emergencies to provide timely, on-the-ground information. Although often viewed with uncertainty because of concerns related to, for example, producer anonymity and lack of authoritative verification, these data often provide relevant information that may be difficult for authorities to collect. Non-authoritative sources also provide an additional layer of subjective information which can indicate the severity of potentially dangerous events, as well as how citizens are reacting to the developing danger or coping with damage resulting from a disaster. By presenting varied applications, such as damage assessments and emergency evacuations, the practicality of non-authoritative sources and how they can add value to official data is demonstrated. It is hoped that one day this research will help save lives.

Acknowledgements Work performed under this project has been partially supported by the Office of the Assistant Secretary for Research and Technology, US Department of Transportation award # RITARS-12-H-GMU (GMU #202717) and also partially funded by the Office of Naval Research (ONR) award #N00014-14-1-0208 (PSU #171570) and the Office of Naval Research (ONR) award #N00014-14-1-0208 (PSU #171570). DISCLAIMER: The views, opinions, findings and conclusions reflected in this presentation are the responsibility of the authors only and do not represent the official policy or position of the USDOT/OST-R, or any State or other entity.

References

Alexander DE (2002) Principles of emergency planning and management. Oxford University Press, Oxford/New York

Atkinson PM, Tatnall A (1997) Introduction neural networks in remote sensing. Int J Remote Sens 18(4):699–709

Benediktsson JA, Swain PH, Ersoy OK (1990) Neural network approaches versus statistical methods in classification of multisource remote sensing data. IEEE Trans Geosci Remote Sens 28(4):540–552

Berenbrok C, Mason RR, Blanchard SF (2009) Mapping Hurricane Rita inland storm tide. J Food Risk Manag 2:76–82

Booij N, Ris RC, Holthuijsen LH (1999) A third generation wave model for coastal regions: model description and validation. J Geophys Res 104:7649–7666

Casti JL (2012) X-events: the collapse of everything. HarperCollins, New York

Chen C, Liu H, Beardsley R (2003) An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries. J Atmos Ocean Technol 20:159–186

Cutter SL (1993) Living with risk: the geography of technological hazards. Edward Arnold, London

De Longueville B, Smith R, Luraschi G (2009) OMG, from here, I can see the flames!: a use case of mining location based social networks to acquire spatio-temporal data on forest fires. In: Proceedings of the 2009 international workshop on location based social networks, Seattle. ACM, pp 73–80

Dietrich J, Zijlema M, Westerink J, Holthuijsen L, Dawson C, Luettich R, Jensen R, Smith J, Stelling G, Stone G (2011) Modeling hurricane waves and storm surge using integrally-coupled scalable computations. Coast Eng 58:45–65

Ferreira CM, Irish J, Olivera F (2014) Uncertainty in hurricane surge simulation due to land cover specification. J Geophys Res-Oceans 119(3):1812–1827

Flanagin A, Metzger M (2008) The credibility of volunteered geographic information. GeoJournal 72(3):137–148

Freudenburg WR, Gramling R, Laska S, Erikson KT (2008) Organizing hazards, engineering disasters? Improving the recognition of political-economic factors in the creation of disasters. Soc Forces 87(2):1015–1038

Freund Y, Schapire R, Abe N (1999) A short introduction to boosting. Jpn Soc Artif Intell 14(771–780):1612

Giles J (2005) Internet encyclopaedias go head to head. Nature 438(7070):900–901

Goodchild M (2007) Citizens as sensors: the world of volunteered geography. GeoJournal 69(4):211–221

Goodchild MF, Glennon JA (2010) Crowdsourcing geographic information for disaster response: a research frontier. Int J Digit Earth 3(3):231–241

Hyvärinen O, Saltikoff E (2010) Social media as a source of meteorological observations. Mon Weather Rev 138(8):3175–3184

Jelesnianski CP, Chen J, Shaffer WA (1992) SLOSH: sea, lake, and overland surges from hurricanes. National Oceanic and Atmospheric Administration, Technical report NWS 48:1–77

Jensen JR, Cowen DC (1999) Remote sensing of urban/suburban infrastructure and socio-economic attributes. Photogramm Eng Remote Sens 65:611–622

Kerr PC et al (2013) U.S. IOOS coastal and ocean modeling testbed: inter-model evaluation of tides, waves, and hurricane surge in the Gulf of Mexico. J Geophys Res Oceans 118:5129–5172

Kumar S, Barbier G, Abbasi MA, Liu H (2011) TweetTracker: an Analysis tool for humanitarian and disaster relief. In: Fifth international AAAI conference on weblogs and social media (ICWSM), Barcelona

Luettich R, Westerink J (2004) Formulation and numerical implementation of a 2D/3D ADCIRC finite element model version 4. www.adcirc.org

Mattocks C, Forbes C (2008) A real-time event-triggered storm surge forecasting system for the state of North Carolina. Ocean Model 25:95–119

Mukai AY, Luetich JWR, Mark D (2002) East coast 2001, a tidal constituent database for Western North Atlantic, Gulf of Mexico and Caribbean Sea. Coastal inlets research program report, coastal and hydraulics laboratory ERDC/CHL TR-02-24. US Army Corps of Engineers. Engineer Research and Development Center, Vicksburg

NOAA (2013a) National Oceanic and Atmospheric Administration, Atlantic basin hurricane database (HURDAT). http://www.aoml.noaa.gov/hrd/hurdat/. Accessed July 2013

NOAA (2013b) National Oceanic and Atmospheric Administration tides and currents. http://tidesandcurrents.noaa.gov/. Accessed 15 Mar 2013

Olea RA, Olea RA (1999) Geostatistics for engineers and earth scientists. Kluwer Academic, Boston

Oliver MA, Webster R (1990) Kriging: a method of interpolation for geographical information systems. Int J Geogr Inf Syst 4(3):313–332

OpenCellID (2011) OpenCellID. http://opencellid.org/. Accessed Feb 2012

Pohl C, Van Genderen J (1998) Review article multisensor image fusion in remote sensing: concepts, methods and applications. Int J Remote Sens 19(5):823–854

Poser K, Dransch D (2010) Volunteered geographic information for disaster management with application to rapid flood damage estimation. Geomatica 64(1):89–98

Schnebele E, Cervone G (2013) Improving remote sensing flood assessment using volunteered geographical data. Nat Hazards Earth Syst Sci 13:669–677

Schnebele E, Cervone G, Waters N (2013) Road assessment after flood events using non-authoritative data. Nat Hazards Earth Syst Sci 14:1007–1015

Sui D, Goodchild M (2011) The convergence of GIS and social media: challenges for GIScience. Int J Geograph Inf Sci 25:1737–1748

Sui D, Elwood S, Goodchild M (2013) Crowdsourcing geographic knowledge: volunteered geographic information (VGI) in theory and practice. Springer, Dordrecht/New York

USGS (2013) U.S. Geological survey national elevation dataset. http://ned.usgs.gov/. Accessed 04 Aug 2013

Venables WN, Ripley BD (2002) Modern applied statistics with S, 4th edn. Springer, New York. http://www.stats.ox.ac.uk/pub/MASS4, ISBN 0-387-95457-0

Voigt S, Kemper T, Riedlinger T, Kiefl R, Scholte K, Mehl H (2007) Satellite image analysis for disaster and crisis-management support. IEEE Trans Geosci Remote Sens 45(6):1520–1528

Waters N (2009) Representing surfaces in the natural environment: implications for research and geographical education, Ch 3. In: Mount NJ, Harvey GL, Aplin P, Priestnall G (eds) Representing, modeling & visualizing the natural environment: innovations in GIS 13. CRC Press, Boca Raton/London/New York, pp 21–39

Wisner B, Blaikie, P, Cannon T, Davis I (2004) At Risk: natural hazards, people's vulnerability and disasters, 2nd edn. Routledge, New York

Zhang J (2010) Multi-source remote sensing data fusion: status and trends. Int J Image Data Fusion 1(1):5–24

Zhang Y, Baptista AM (2008) A semi-implicit Eulerian-Lagrangian finite-element model for cross-scale ocean circulation. Ocean Model 21:71–96

Zhang J, Atkinson P, Goodchild M (2014) Scale in spatial information and analysis. CRC, Boca Raton