

Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study

Guido Cervone, Elena Sava, Qunying Huang, Emily Schnebele, Jeff Harrison & Nigel Waters

To cite this article: Guido Cervone, Elena Sava, Qunying Huang, Emily Schnebele, Jeff Harrison & Nigel Waters (2016) Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study, International Journal of Remote Sensing, 37:1, 100-124, DOI: [10.1080/01431161.2015.1117684](https://doi.org/10.1080/01431161.2015.1117684)

To link to this article: <http://dx.doi.org/10.1080/01431161.2015.1117684>



Published online: 13 Dec 2015.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study

Guido Cervone^a, Elena Sava^a, Qunying Huang^b, Emily Schnebele^a, Jeff Harrison^c and Nigel Waters^a

^aGeoinformatics and Earth Observation Laboratory, Department of Geography and Institute for CyberScience, The Pennsylvania State University, University Park, PA, USA; ^bDepartment of Geography, University of Wisconsin, Madison, WI, USA; ^cThe Carbon Project, Arlington, VA, USA

ABSTRACT

A new methodology is introduced that leverages data harvested from social media for tasking the collection of remote-sensing imagery during disasters or emergencies. The images are then fused with multiple sources of contributed data for the damage assessment of transportation infrastructure. The capability is valuable in situations where environmental hazards such as hurricanes or severe weather affect very large areas. During these types of disasters it is paramount to 'cue' the collection of remote-sensing images to assess the impact of fast-moving and potentially life-threatening events. The methodology consists of two steps. First, real-time data from Twitter are monitored to prioritize the collection of remote-sensing images for evolving disasters. Commercial satellites are then tasked to collect high-resolution images of these areas. Second, a damage assessment of transportation infrastructure is carried out by fusing the tasked images with contributed data harvested from social media such as Flickr and Twitter, and any additional available data. To demonstrate its feasibility, the proposed methodology is applied and tested on the 2013 Colorado floods with a special emphasis in Boulder County and the cities of Boulder and Longmont.

ARTICLE HISTORY

Received 26 May 2015

Accepted 28 October 2015

1. Introduction

Environmental hazards pose a significant threat to the development and sustainment of our society. Rapid population growth, the emergence of megacities, and high-risk facilities such as high dams and nuclear power plants have increased the risk posed by natural hazards at unprecedented levels (Tate and Frazier 2013). A single catastrophic event can claim thousands of lives, cause billions of dollars of damage, trigger a global economic depression, destroy natural landmarks, render a large territory uninhabitable, and destabilize the military and political balance in a region (Keilis-Borok 2002).

1.1. Remote sensing

Remote-sensing data are paramount during disasters and have become the de facto standard for providing high-resolution imagery for damage assessment and the coordination of disaster relief operations (Cutter 2003; Joyce et al. 2009). Organizations such as the International Charter for Space and Disasters (<http://www.disasterscharter.org/>) provide high-resolution images from commercial and research air- and space-borne instruments within hours of major events, frequently including 'before' and 'after' scenes of the affected areas (Stryker and Jones 2009; Duda and Jones 2011; Cervone and Manca 2011). These 'before' and 'after' images are quickly disseminated through official government portals and news channels to inform the public of the magnitude of the event, and often serve to sensitize citizens about the unfolding tragedies. In addition, first responders rely heavily on remotely sensed images to coordinate relief and response efforts as well as to prioritize resource allocations (Cutter 2003).

Determining the location and severity of damage to transportation infrastructure is particularly critical for establishing evacuation and supply routes as well as repair and maintenance agendas (Oxendine, Sonwalkar, and Waters 2012). Following the Colorado floods of September 2013, over 1000 bridges required inspection and approximately 200 miles of highway and 50 bridges were destroyed. The assessment of transportation infrastructure over such a large area could have been accelerated through the use of high-resolution imagery and geospatial analysis (Uddin 2011).

Despite the wide availability of large remote-sensing data sets from numerous sensors, specific data might not be collected in the time and space most urgently required. Geo-temporal gaps result due to satellite revisit time limitations, atmospheric opacity, or other obstructions. Tasking instruments on board satellites and other aerial platforms for data collection is thus crucial for the timely delivery of data for damage assessment and disaster relief. However, satellite tasking is usually limited by orbital restrictions and the locations of data-receiving stations. It is usually predefined and based on the statistical likelihood that data for an area are needed. A small number of satellite instruments can be oriented to collect data at an oblique angle with respect to the satellite path. For this class of instruments, a correct tasking can greatly increase the data coverage during emergencies.

Furthermore, aerial platforms, especially Unmanned Aerial Vehicles (UAVs), can be quickly deployed to collect data over specific regions. UAVs are capable of providing high-resolution, near-real-time images often with less expense than manned aerial- or space-borne platforms. Their quick response time, high manoeuvrability, and resolution make them important tools for disaster assessment (Tatham 2009). Tasking data collection for specific regions affected by a hazard is particularly difficult during major events. For extremely large events such as Hurricane Sandy in 2012 or the Colorado floods of 2013, rapid and systematic evaluations were difficult because the area affected was so extensive. This difficulty is further enhanced when events quickly unfold and are mitigated by local infrastructure and relief efforts.

Therefore, it is apparent that during emergencies the tasking of data collection from remote-sensing platforms must be constantly assessed and refined based on the needs of emergency responders and the constraints dictated by the number and type of instruments available. This assessment has historically been based on official

measurements and established plans, and does not account for the availability of real-time, on-the-ground data freely contributed by citizens.

1.2. Social media and volunteered geographical information

Novel information streams, such as social media-contributed videos, photographs, and text as well as other open sources, are redefining situation awareness during emergencies. When these contributed data contain spatial and temporal information, they can provide valuable Volunteered Geographical Information (VGI), harnessing the power of 'citizens as sensors' to provide a multitude of on-the-ground data, often in real time (Goodchild 2007). There are several opportunities and challenges associated with the use of VGI. Elwood (2008); Elwood, Goodchild, and Sui (2012) examine the content and characteristics of VGI, the technical and social processes through which it is produced, appropriate methods for synthesizing and the use of these data in research, and emerging social and political concerns related to this new form of information.

Although these volunteered data are often published without scientific intent, and usually carry little scientific merit, it is still possible to mine mission critical information (Huang and Xiao 2015). For example, during hurricane Katrina, geolocated pictures and videos searchable through Google provided early emergency responders with ground-view information. These data have been used during major events, with the capture, in near-real-time, of the evolution and impact of major hazards (De Longueville, Smith, and Luraschi 2009; Pultar et al. 2009; Heverin and Zach 2010; Vieweg et al. 2010; Acar and Muraki 2011; Verma et al. 2011; Earle, Bowden, and Guy 2012; Tyshchuk et al. 2012). Specifically, VGI based on Twitter and other non-authoritative data have been shown to contain valuable data that can be used for improving flood estimation in near-real-time (Poser and Dransch 2010; McDougall and Temple-Watts 2012; Triglav-Čekada and Radovan 2013; Schnebele and Cervone 2013; Schnebele et al. 2014a, 2014b).

Volunteered data can be employed to provide timely damage assessment, help in rescue and relief operations, as well as for the optimization of engineering reconnaissance (Laituri and Kodrich 2008; Dashti et al. 2014). While the quantity and real-time availability of VGI make it a valuable resource for disaster management applications, data volume, as well as its unstructured, heterogeneous nature, makes the effective use of VGI challenging. Volunteered data can be diverse, complex, and overwhelming in volume, velocity, and in the variety of viewpoints they offer (Huang and Xu 2014). Negotiating these overwhelming streams is beyond the capacity of human analysts. Current research offers some novel capabilities to utilize these streams in new, ground-breaking ways, leveraging, fusing, and filtering this new generation of air-, space-, and ground-based sensor-generated data (Oxendine et al. 2014).

This research presents a novel approach to prioritizing the collection of remote-sensing data from satellites, airplanes, and UAVs during hazard events by utilizing VGI as a filtering tool. In addition, it proposes the use of VGI for disaster assessment to fill in the gaps when remote-sensing data are lacking or incomplete. In order to use social media efficiently and effectively and VGI to 'cue' or augment satellite observations, it is necessary to filter the data for content and to geolocate them using a variety of text-mining and network analysis algorithms. Filtering yields a rapid and direct identification

of affected areas, which can aid authorities to prioritize site visits and response initiatives as well as to task additional data collection.

1.3. 2013 Boulder CO floods

We present an application of this new methodology to the 2013 floods that occurred around the City of Boulder, CO. The September 2013 flooding in Colorado was the worst hydrological disaster in the state's history. The severity of the event was the result of extreme precipitation, pre-existing drought, and recent wildfires that hardened the ground and reduced the vegetation layer. With precipitation exceeding 50 cm in parts of Boulder County, the area experienced catastrophic flooding, property destruction totalling over \$2 billion in damage, the evacuation of over 180,000 people, and a tragic loss of eight lives. Damage to transportation infrastructure was especially severe and affected large areas of the state.

Boulder is located near the foothills of the Rocky Mountains approximately 50 km northwest from Denver. This unique location has made Boulder the number one flash flood community in the state of Colorado as major flash floods have developed in the past and still continue to occur (Elaine 2014). Development in Boulder started in 1910 when prominent planner Frederick Law Olmstead Jr. recommended Boulder Creek be lined with a park and preserved as open space to limit structural damage in case of a flood (Crumpacker 1985). More recently the city's rapidly growing population raised concerns about the loss of open space. In the last 15 years alone, Boulder has spent more than \$45 million dollars on sustainable projects based on the blueprint of the community (Elaine 2014). However, the 2013 flood was not the large flash flood officials had been planning for. It was an unprecedented event where Colorado received the amount of its average annual precipitation in only eight days.

The 2013 Colorado flood propagated quickly and with little warning. From 11 to 15 September, a catastrophic storm system formed over the Rocky Mountain Front Range impacting 18 counties and resulting in severe flooding in many parts of Colorado. Record-breaking precipitation occurred along the Colorado Front Range from Larimer and Weld Counties southward to El Paso County. Boulder County was the worst hit with 25 cm of rain on 12 September 2013 and over 43 cm by 15 September 2013. The effects of the storm were not only felt in-state, but flooding and heavy precipitation also extended into several neighbouring states such as Nebraska, New Mexico, and Wyoming (Gochis et al. 2014).

1.4. Objectives

The present article describes a methodology and its implementation to achieve two main objectives.

- (1) Using Twitter data to task the collection of remote-sensing data during emergencies. A program called CarbonScanner is described to automatically scan Twitter data, and to identify hotspots and keywords for an unfolding emergency. While the current research focuses on commercial high-resolution satellite platforms, it

can also be applied for the collection of aerial images using manned or unmanned platforms.

- (2) Fusing the tasked remote-sensing images with additional contributed data, including tweets and VGI, for damage assessment. The current research specifically discusses the assessment of transportation infrastructure; however, it can be extended to other situations.

2. Data

Multiple sources of contributed, remote-sensing, and open source geospatial data were collected and used during this research. A summary of the sources and collection dates of the contributed and remote-sensing data are available in [Table 1](#).

2.1. Twitter

Twitter is one of the largest social networking sites, and it is widely used to share information through micro-blogging (Waters 2014). These micro-blogs, or ‘tweets’, are limited to 140 characters, so abbreviations and colloquial phrasing are common, making the automation of filtering by content challenging. Twitter is very popular during emergencies and disasters, and it is being used by both official government agencies and the public to disseminate information.

Central to the operation of Twitter is the use of hashtags, words or unspaced phrases prefixed with the sign #. They are identifiers unique to Twitter and are frequently used to search and filter information. The creation and use of a hashtag can be established by any user who wants to create a concept category to share specific information about a subject. For example, during the 2013 Boulder floods, the hashtag #boulderflood was used by users and agencies to share information about this particular event.

Twitter data can be queried for specific hashtags or text present in the tweets, and for spatial and temporal constraints. There are several web-based tools and an API for the automatic querying, filtering, and displaying of tweets. For this study, tweets are harvested using the CarbonScanner application to identify ‘hotspots’ and task satellite data collection. CarbonScanner scans tweets, identifies relevant keywords and hashtags, and georectifies the data (see [Sections 3.1.2](#) and [3.1.1](#)).

Table 1. Quantity of data collected for each source from 11 September 2013 to 21 September 2013. The ‘many’ for the Falcon UAV indicates that a stream of images was collected.

| Data Source | September | | | | | | | | | | |
|-------------------------|-----------|-------|-------|-------|-------|-------|------|------|-----|----|-----|
| | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| Ground | | | | | | | | | | | |
| Tweets | 191 | 22432 | 57840 | 20923 | 22295 | 13419 | 7878 | 3401 | | | |
| Flickr Photographs | 2 | 313 | 37 | 334 | 20 | 517 | 8 | | | | |
| Aerial Images | | | | | | | | | | | |
| CAP | | | 1595 | 33 | 1572 | 894 | 67 | 654 | 104 | | 788 |
| Falcon UAV | many | many | many | | | | | | | | |
| Satellite Images | | | | | | | | | | | |
| WorldView2 | | | 2 | 5 | | | 3 | | | | |
| Landsat 8 | | | | 1 | | | | | | | |

The specific geographical locations, hashtags, and keywords identified by CarbonScanner are used to collect additional Twitter data using applications developed at The Pennsylvania State and the University of Wisconsin at Madison. Specifically, for the current research the following three criteria were used:

- Tweets extending from 105°4'53" – 105°17'50" W and 40°5'58" – 39°57'12" N and containing the hashtag '#boulderflood' from 12 to 16 September 2013;
- Tweets extending from 105°01'56" – 105°25'49" W and 40°5'58" – 39°56'01" N and containing all hashtags from 11 to 17 September 2013; and
- Tweets extending from 12 to 16 September 2013 and containing the hashtag '#boulderflood'.

About 150,000 tweets have been analysed in the current research (see [Table 1](#)).

2.2. Ground photographs

Ground photographs contributed both by official sources and by ordinary citizens were collected to document the flooding in the Boulder area. For this study, a total of 1232 images relative to the period from 11 to 17 September 2013 were harvested using Flickr, a popular image-sharing portal (<http://www.flickr.com>). Flickr provides an API to query for spatial and temporal extents, making it possible to fully automate the analysis. Only geotagged Flickr photographs were used.

2.3. WorldView2

Ten, full-resolution GeoTIFF WorldView2 multispectral images collected by Digital Globe on 13 September 2013 (two images), 14 September 2013 (five images), and 17 September 2013 (three images) provide high-resolution data of Boulder and the surrounding counties. The WorldView2 data used were tasked using the CarbonScanner application developed for this research.

WorldView2 is a commercial Earth observation satellite owned and operated by DigitalGlobe that was launched into orbit on 8 October 2009. It provides commercially available panchromatic images of 0.46 m resolution, and eight-band multispectral images with 1.84 m resolution.

The eight multispectral bands at a 1.84 m resolution cover the following part of the electromagnetic (EM) spectrum: coastal (0.400–0.450 μm), blue (0.450–0.510 μm), green (0.510–0.580 μm), yellow (0.585–0.625 μm), red (0.630–0.690 μm), red edge (0.705–0.745 μm), near-IR 1 (0.770–0.895 μm), and near-IR 2 (0.860–0.900 μm).

2.4. Landsat 8

Two resolution multispectral Landsat 8 OLI/TIRS images collected on 12 May 2013 and on 17 September 2013 provide data of the Boulder County area before and after the flooding, respectively. The data were downloaded from the United States Geological Survey (USGS) Hazards Data Distribution System (HDDS).

Landsat 8 consists of nine spectral bands with a resolution of 30 m: Band 1 (coastal aerosol, useful for coastal and aerosol studies, 0.430–0.450 μm); Bands 2–4 (optical, 0.450–0.510, 0.530–0.590, 0.64–0.67 μm), Band 5 (near-IR, 0.850–0.880 μm), Bands 6 and 7 (shortwave-IR, 1.570–1.650, 2.110–2.290 μm), and Band 9 (cirrus, useful for cirrus cloud detection, 1.36–1.38 μm). In addition, a 15 m panchromatic band (Band 8, 0.500–0.680 μm) and two 100 m thermal IR (Bands 10 and 11, 10.600–11.190, 11.500–12.510 μm) were also collected from Landsat 8 OLI/TIRS.

2.5. Civil Air Patrol

Satellite remote-sensing data may be insufficient as a function of spatial resolution, revisit time, or may be obstructed due to clouds or vegetation. Therefore, data from other sources can be used to provide supplemental information.

The Civil Air Patrol (CAP) is a congressionally funded, non-profit corporation that functions as an auxiliary to the United States Air Force. The CAP conducts a variety of missions in support of federal, state, local, and non-governmental entities, including search and rescue, disaster relief support, and aerial reconnaissance for homeland security.

CAP images were captured from 14 to 17 September 2013 in the areas surrounding Boulder (105°32'11' – 104°59'33' W and 40°15'37' – 39°56'10' N) and provide a third source of remote-sensing data. The georeferenced CAP red-green-blue (RGB) composite images were downloaded from the USGS HDDS.

2.6. Falcon UAV

Falcon UAV images were collected over the cities of Lyons and Longmont from 12 to 14 September 2013. Falcon UAV collected valuable RGB composite images while other aircraft were grounded due to weather conditions (<http://www.falconunmanned.com/>). The mosaicked, georeferenced images were downloaded as a .kmz file from the Falcon Unmanned website.

2.7. Open source geospatial data

Shapefiles defining the extent of the City of Boulder and Boulder County were downloaded from the City of Boulder (<https://bouldercolorado.gov>) and the Colorado Department of Local Affairs (<http://www.colorado.gov>) websites, respectively. In addition, a 2012 TIGER/line[®] shapefile of road networks for Boulder County was downloaded from the United States Census Bureau (<http://www.census.gov>).

3. Methodology

The proposed methodology is based on two separate tasks:

- (1) Remote-sensing data collection
- (2) Damage assessment

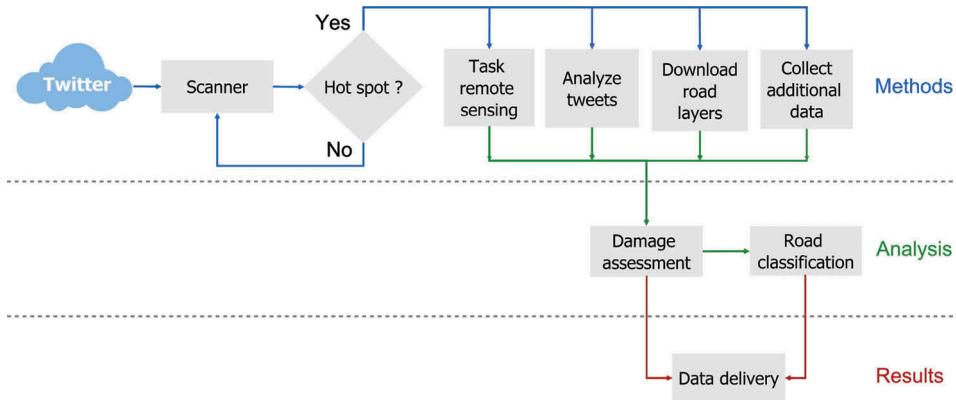


Figure 1. Flow chart of the proposed methodology.

Figure 1 shows the flow chart of the proposed methodology, indicating three main phases. The first phase consists of a series of methods to ingest and process Twitter data to identify regions of interest and to task remote-sensing data collection. Data are also added to characterize the transportation infrastructure network (e.g. roads, tunnels, bridges). Additional VGI data such as photographs and aerial images are also downloaded. In the second phase, Twitter data are fused with the tasked remote-sensing data, and the additional VGI to estimate the damage. In the final and third phase the data are made freely accessible through web services.

3.1. Remote-sensing data collection

The first task consists of the analysis of Twitter data to identify the location of natural hazards, or 'hot spots', and the tasking of remote-sensing data collection for these areas.

3.1.1. Geolocation of tweets

The proposed methodology is heavily dependent on the availability of geolocated data. Only between 1% and 2% of all tweets contain geolocation information. Because of the huge volume of Twitter data, usually even such a small percentage can be used to extract meaningful spatial patterns. However, during emergencies it is paramount to maximize the amount of spatial information, and therefore it is necessary to find alternative methodologies for geolocating tweets.

For each tweet entry that does not contain geolocation information collected by the CarbonScanner application, all non-words are removed (punctuation, special characters, URLs, emoticons, and white space). The remaining text is then tokenized into uni-, bi-, and tri-grams, which correspond to one, two, and three consecutive words after stop words (e.g. a, an, and are). All of the stop words are removed. Each n -gram is queried against a gazetteer database. A gazetteer is a geographic index or dictionary to help identify the geographic location associated with a place name. In general, gazetteer entries consist of place names, locations, and other descriptive information. For this work, a database was created using data from the USGS and from GeoNames (<http://www.geonames.org/>), and containing approximately 2.1 million entries. If a match is

returned for a place name query, the associated geometry is used to derive the location of the tweet where it was sent from. If the location of a place is represented as a polyline or polygon feature, its centroid is used to georeference the tweet. The CarbonScanner constantly processes Twitter streams and matches the tweet content to the gazetteer database in real time.

This same methodology is employed to geolocate the tweets used for damage assessment (Section 3.2). As a result, an additional 8% of tweets were given location information, increasing the percentage of geolocated tweets used in this research from the original 2% (circa 3000 tweets) to 10% (circa 15,000 tweets).

3.1.2. Identification of hot spots

The CarbonScanner scanning application was developed to access and harvest social media data by browsing Twitter in real time. This application 'scans' the United States each hour to assess and generate alerts by clustering in space and time tweets whose text contains specific keywords for natural hazards (e.g. floods, tornadoes). These alerts or 'hot spots' are areas where natural hazards may be potentially occurring. The application system is highly flexible with filter settings managed by a portal service enabling manager-level users to quickly adjust keyword and other settings as events develop.

A potential event is detected when a significant amount of activity (in space and time) is identified in a region. In this work, the threshold for significant activity is set to 10 tweets within a 100 km² area. A 100 km² extent was selected because it is sufficient to capture significant transportation infrastructure in many urban areas as well as being the minimum ordering area for DigitalGlobe orthorectified images. The threshold settings can be adjusted to user requirements or preferences. When the threshold is reached, an alert box is generated, cueing the collection of images for that region. The request for data collection is automatically issued to DigitalGlobe, which is satisfied pending the constraints of the satellites and their orbits.

Figure 2(a) shows the CarbonScanner application during the 2013 Boulder floods identifying several tweets that clustered in space and time, and which caused an alert to be generated. The alert boxes generated are shown in Figure 2(b). An additional task performed by CarbonScanner is the identification of most common hashtags among the tweets that are used to generate an alert. These hashtags are used to query additional Twitter data for damage assessment.

Once the data are acquired, they are made available as open data services. The imagery is deployed in near-real-time using OpenImageMap (<http://www.cubewerx.com/solutions/openimagemap/>) services as well as other open mapping services including Open Geospatial Consortium Web Map Service (OGC WMS), Web Map Tile Services (WMTS), Web Coverage Services (WCS), Google Maps API, Google Earth KML overlays, and Open Source Geospatial Foundation Tile Map Service (OSGeo TMS). Providing the images in an open source format ensures rapid deployment as well as open access to the data. System settings are adjustable and managed by a portal service enabling manager-level users to quickly add images, manage services, and access controls. Images may also be accessed by the WMS standard and combined with National Spatial Data Infrastructure (NSDI) framework data in any GIS supporting this popular standard.

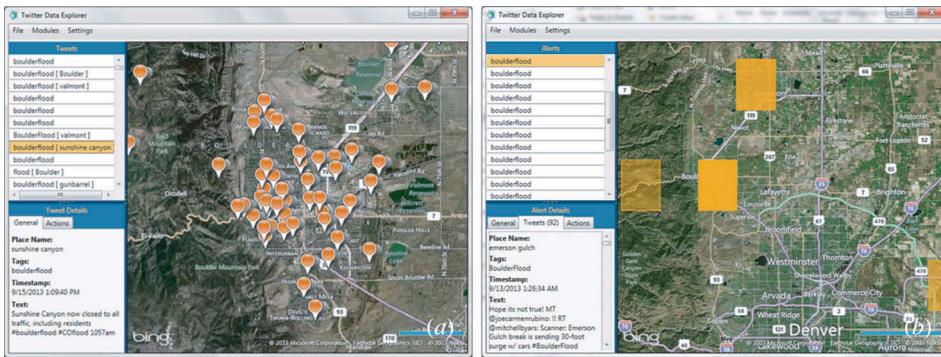


Figure 2. Screenshot of CarbonScanner during the 2013 Boulder floods showing filtered and geolocated tweets captured by scanning application in real time (a), and the alert box generated for which remote-sensing data collection was tasked (b).

3.2. Damage assessment

Once a potential 'hot spot' is identified and remote-sensing data tasked and acquired, these images are fused with Twitter and additional available data for damage assessment. For the Boulder flood test case described, the most important task consisted in identifying the flood extent. This task called for the automatic detection of water in the images available (satellite, aerial, ground), as well as the identification of tweets that specifically mentioned flooded areas. This damage assessment is tailored to classify transportation infrastructure during and after an event to determine which roads are impassable.

3.2.1. Classification of satellite images

A decision tree supervised machine learning classifier is employed to semi-automatically identify water regions in each of the satellite images (Ripley 2008).

Cervone and Haack (2012) describe the general rule induction methodology and its implementation used in this study.

Several control areas of roughly the same are manually identified as examples and counter-examples of water pixels. This task involves visually inspecting the images using graphics software to identify the flooded and non-flooded regions. These regions are then used as training events by the classifier, which learns a binary decision tree. This is an upside-down tree-like structure, where each node is an attribute-value condition (e.g. $\text{band}_7 < .088$), and each branch represents the path that is taken if the condition is satisfied (left) or not (right). A classification is made by transversing the entire tree starting at the root, and assigning the class (water or no water) associated with the leaf.

Figure 3 shows a sample decision tree built using multispectral satellite data for the classification of pixels. In this particular example, class 1 is water, class 2 is an urban area, and class 3 is a vegetated area.

The learned tree is used to classify all pixels in all images available. The decision tree algorithm uses a combination of bands to determine a classification. Most important is the near-IR band because in this part of the EM spectrum, water is easily distinguished from soil and vegetation due to its strong absorption (Smith 1997).

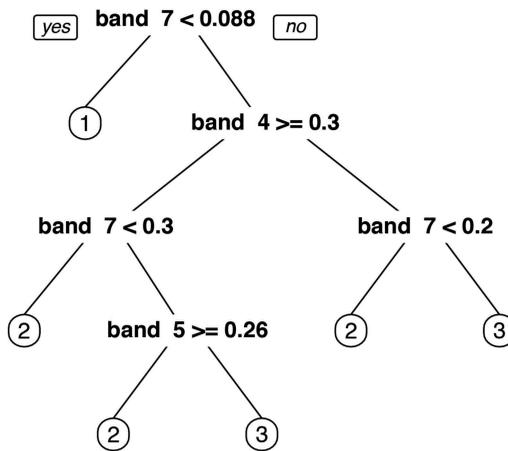


Figure 3. Decision Tree learned for the classification of pixels of a satellite image into one of three classes. In this particular example, class 1 is water, class 2 is an urban area, and class 3 is a vegetated area.

3.2.2. Classification of RGB images

Working with aerial and ground images poses challenges because the data are usually collected using conventional digital cameras that can only capture information in visible light. Therefore, each image contains only three spectral channels in RGB colour space. Conversely, advanced multispectral sensors available through various remote-sensing platforms are structured to collect data in different parts of the EM spectrum usually ranging from ultraviolet (UV) to emissive infrared (IR). For the identification of water, near-IR is particularly important (see [Section 3.2.1](#)).

The aerial (CAP and Falcon UAV) and ground images are classified using an ensemble of a supervised decision tree machine learning classifier and a Maximum Likelihood Classifier. Two classifiers are used because of the additional difficulty introduced by the lack of IR information.

Furthermore, because of the lack of IR data, a 2D wavelet transformation using the Haar mother wavelet is first run to extract texture information for all the images, and thus expands the search space beyond RGB. Then for each transformed image representation (RGB + wavelet coefficient components) a decision tree learning algorithm is used to classify each pixel as water in a similar manner to that described in [Section 3.2.1](#).

There are thousands of aerial images, and thus training regions are chosen only in a small subset of the total data. Once the decision tree is learned using this small training set, it is applied to all images in the data set. An identical process is repeated for the ground images.

The overall goal of the proposed method is not to achieve a 100% accuracy in pixel classification, but rather to perform a quick assessment to determine which areas are

more likely to be flooded. Even if large errors might occur, they are naturally filtered and smoothed when thousands of points across the different data sets are fused together.

3.2.3. Spatial interpolation

The data fusion problem consists of merging together heterogeneous data with different temporal and spatial resolutions. Furthermore, especially in the case of tweets, the amount of data to be analysed is very large although it is usually only 1% of the entire data stream. This is usually referred to as the 'Big Data paradox', where very large amounts of data to be analysed are only a small sample of the total data. This sample might or might not be representative of the distribution of the entire population.

The absence of data in some parts of the region is likely to underestimate the total damage. To compensate for missing data, a non-homogeneous weighting is employed (Tobler 1970, 236). This assumes some level of dependence among spatial data as well as to closely examine spatial information found to be inconsistent with its surroundings (Waters, forthcoming). For these reasons a punctual representation of data may not be sufficient to provide a complete portrayal of the hazard; therefore a spatial interpolation is employed to estimate flood conditions and damage from point sources.

Spatial interpolation consists of estimating the damage at unsampled locations by using information about the nearest available measured points. This process can be implemented by using two different approaches: deterministic models and statistical techniques. The first method does not provide an indication of the extent of possible errors, whereas the second method supplies probabilistic estimates. Deterministic models include Inverse Distance Weighted (IDW), Rectangular and Spline (Waters 2009). Statistical methods include Kriging and kernel (Myers 1994).

Kernel interpolation is the most popular non-parametric density estimator that is a function $\hat{p} : \mathfrak{R} \times (\mathfrak{R})^N \rightarrow \mathfrak{R}$, where \mathfrak{R} indicates the set of all real numbers. In particular the density estimate $\hat{p}(x)$ for a given value of x is equal to

$$\hat{p}(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right), \quad (1)$$

where K is the kernel function, h is the bandwidth (Raykar and Duraiswami 2006), N is the number of observed points, x_i are the observed data points, and i is the point number. The density estimator chosen for this work is a Gaussian kernel with zero mean and unit variance having the following form:

$$\hat{p}(x) = \frac{1}{N\sqrt{2\pi}h^2} \sum_{i=1}^N e^{-\frac{(x-x_i)^2}{2h^2}}, \quad (2)$$

where N , h , and x_i are scalars and, as for Equation (1), respectively, represent the number of observed points, the bandwidth, and the observed data points. Kernel interpolation is often preferred because it provides an estimate of error as opposed to methods based on radial basis functions. In addition, it is more effective than a Kriging interpolation in case of small data sets (e.g. the data set of photographs in this project) or data with non-stationary behaviour (all data sets used) (Mühlenstädt and Kuhnt 2011).

The result of a kernel density estimation will depend on the kernel K and the bandwidth h chosen. The former is linked to the shape, or function, of the curve centred

over each point whereas the latter determines the width of the function. The choice of bandwidth will exert greater influence over an interpolation result than the kernel function. Indeed, as the value of the bandwidth decreases, the local weight of single observations will increase. There are different methods for choosing an appropriate bandwidth. For example, it can be identified as the value that minimizes the approximation of the error between the estimate $\hat{p}(x)$ and the actual density $p(x)$ as explained in Raykar and Duraiswami (2006). In this work, a spatial kernel estimation is used because of the geographical nature of the data, and an optimal bandwidth matrix is chosen using the method illustrated in Duong and Hazelton (2005).

Because confidence in data may vary with source characteristics, bandwidth selection can be adjusted for each data type. The basic idea is that the more certain the information given by a kind of data, the higher the chosen bandwidth. For example, aerial or ground images can be weighted according to the amount of water pixels identified by the machine learning classification. Some layers might be weighted more highly because they originated from official sources, or can be quantified and verified. On the other hand, tweets are very subjective and often reflect users' feelings. Without performing a sentiment analysis of the tweets (Mohammad, Kiritchenko, and Zhu 2013), it is safer to weigh them lower than quantitative data.

3.3. Transportation damage assessment

The assessment of transportation facilities, such as roads, railways, and bridges, is performed by overlaying the satellite flood extent and the contoured interpolation over a street layer. The facilities that fall within the flooded area and the contoured interpolation are predicted to be affected and potentially impassable. The degree to which they are affected is a function of the interpolation value and the satellite flood estimation. Generally, all roads that are inside the satellite flood extent are marked as having severe damage.

4. Results

The CarbonScanner application detected a 'hot spot' of Twitter activity in Colorado beginning 11 September 2013 for which remote-sensing imagery was tasked. Ten WorldView2 high-resolution satellite images were acquired as a result of the request. The damage assessment of transportation infrastructure was performed by fusing these satellite images with Twitter data and the other sources available as described in Section 2.

4.1. Classification of images

The supervised machine learning classification discussed in Section 3.2.1 was used to identify water pixels from the WorldView2 and Landsat8 satellite data. A total of six control regions of roughly 150×150 pixels were identified for each image, three as examples of flooded areas and three as counter-examples of non-flooded areas. The decision tree algorithm was trained on these control areas, and the results were used to classify all remaining pixels. This process was repeated for each of the 12 images (10

WorldView2 and two Landsat8). The supervised classification was applied to each scene individually to account for differences in brightness values and the results were merged together in a single layer.

Remote-sensing pixel classification is a very common data analysis process, and it does not present particular problems. However, it is intrinsically limited by the spatial resolution of the images, and the presence of clouds. The use of supplemental data sources, such as the CAP aerial photographs and the Falcon UAV data, shows flooding in areas that are not captured by the satellite remote sensing. In the case of the Boulder flooding, the cloud cover over the front range and western parts of the City of Boulder made the identification of water from satellite platforms difficult. The ability of planes to fly below cloud cover as well as to collect data without the revisit limitation common to space-borne platforms allowed the CAP to capture flooding and damage that was not visible in the satellite images.

The supervised machine learning classification discussed in Section 3.2.2 was applied to the CAP data. First, 40 images were selected at random, and for each image three control regions were identified as examples of flooded pixels and three regions as counter-examples of non-flooded pixels. All the control regions were kept the same at 150×150 pixels. The models learned by the decision tree and maximum likelihood algorithms were then applied to all of the data to quantify the amount of flooded pixels.

Specifically, each pixel was classified as one of three classes: No water (NW), water in one classifier (C1), and water in both classifiers (C2). By using an ensemble of classifiers, a measure of uncertainty can be derived. For example, Figures 4(a) and (b) show a classified image (a) and the original RGB version (b). In the figure, the pixels classified

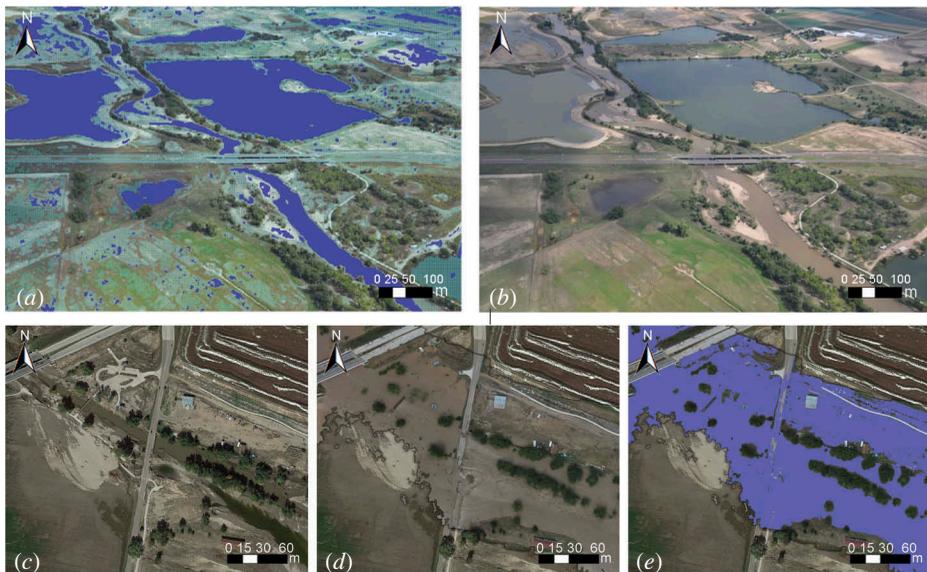


Figure 4. Water pixel classification in aerial images: (a) CAP classified, (b) CAP original, (c) Digital Globe pre-flood, (d) Falcon UAV, and (e) classified Falcon UAV. CAP image classification by both classifiers (blue) or by a single classifier (teal). The Falcon UAV classified image shows significant damage sustained by roads and bridges.

as NW are left in the original colour, those classified as C1 are shown in transparent teal, and those classified as C2 as solid blue. An identical process was applied to the classification of the Falcon UAV images. For this task 10 control regions were chosen throughout the scene acquired, five for flooded and five for non-flooded pixels, and used as input for the decision tree algorithm. The learned decision tree was used to classify the entire acquired scene.

Figures 4(c)–(e) show a change detection analysis performed using the methodology for a region with significant transportation infrastructure. The figure shows high-resolution Digital Globe data as ‘before’ images (c) and the Falcon UAV photographs as ‘after’ images (d), and the results of the classification performed (e). This figure illustrates how aerial data collected from a UAV provide valuable, timely, high-resolution details of flooding and affected transportation infrastructure. The use of these high-resolution images allows flooding and damage to be identified within a few metres in near-real-time.

4.2. *Damage assessment*

Following the collection of remote-sensing data, contributed data were also collected, geolocated, and interpolated. The WorldView2 images were manually georectified using the high-resolution road layer obtained by the City of Boulder website. All of the data are overlaid on the remote-sensing classification to provide an enhanced indication of flood activity.

Figure 5 provides a summary of the data available for damage assessment and Figure 6 a zoom over the cities of Boulder and Longmont, CO, which were studied in detail. The CAP data were acquired along major rivers, reservoirs, and areas that were at risk of flooding. The contributed tweets and photographs had a clear spatial distribution that corresponded to the population density. The highest concentration of tweets and photographs were found in the City of Boulder.

The high concentration of VGI was due to a high population density, composed in part by young and highly educated students and young professionals. In fact, Boulder CO is home to the University of Colorado, and several research institutions (e.g. National Center for Atmospheric Research (NCAR), National Renewable Energy Laboratory (NREL), National Institute of Standards and Technology (NIST)) and high technology companies (e.g. Google, Oracle). The Falcon UAV footprint is shown in the NE corner of Figure 6 corresponding to the city of Longmont. Although the UAV data were available only for a small area, they proved crucial because the WorldView2 image contained several clouds exactly in this region, and thus could only be partially used for damage assessment.

All the satellite images available were co-registered, and classified using their multispectral channels. The satellite classified layer was compared with the FEMA flood estimation available for Boulder County. An area analysis between the FEMA and the satellite flood estimation shows a spatial correlation of 87%. The errors show a tendency of both WorldView2 and Landsat to underestimate the flood extent, which can be explained by the presence of clouds. This correlation was performed only on about 65% of Boulder County because the satellite imagery limited spatial coverage extent. The kernel interpolation was performed using all of the contributed data to refine the satellite flood estimation. The kernel result extends the flood area well beyond the FEMA map, including several downtown

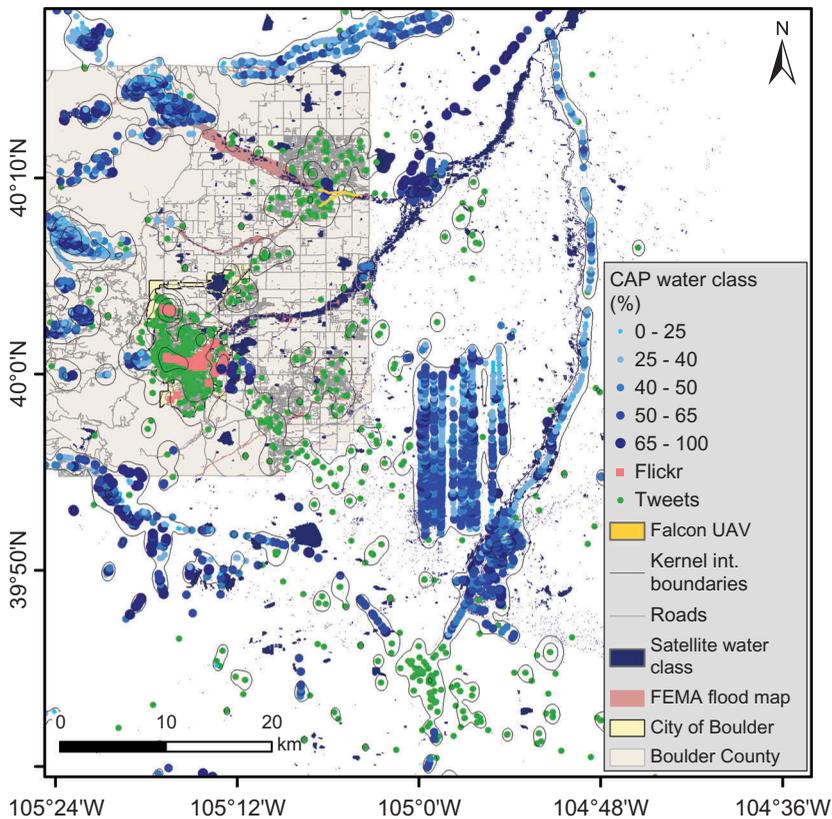


Figure 5. Data sources available, including Flickr ground photographs, tweets, CAP, Falcon UAV, and the satellite water classification. The road network is also shown.

areas. Results are shown for the cities of Boulder and Longmont, where different combinations of data are available.

4.2.1. Road damage assessment: case 1

A transportation damage assessment was performed for the City of Boulder, CO. The supervised classification of remote-sensing data classified very few water pixels in the City of Boulder. This is likely due to the images being collected about a week after the flood event began, when waters might have already partially or fully receded. The presence of obstructing vegetation and cloud cover might also have prevented a clear view of the ground.

Figure 7(a) shows a map of the satellite water classification and the FEMA flood map for Boulder City and County. There is a good agreement between the two flood maps, and large areas classified as water in downtown Boulder are absent. This lack of flooding contradicts the contributed data that indicate different levels of flooding throughout the City of Boulder. Figure 7(b) shows the distribution of the contributed data (tweets, photographs, and CAP images), and the kernel interpolation generated.

The kernel interpolation was used to predict which roads were more likely to be damaged and impassable. Figure 7(c) shows the roads classified as having medium, moderate, and severe damage, along with roads predicted not to have been affected.

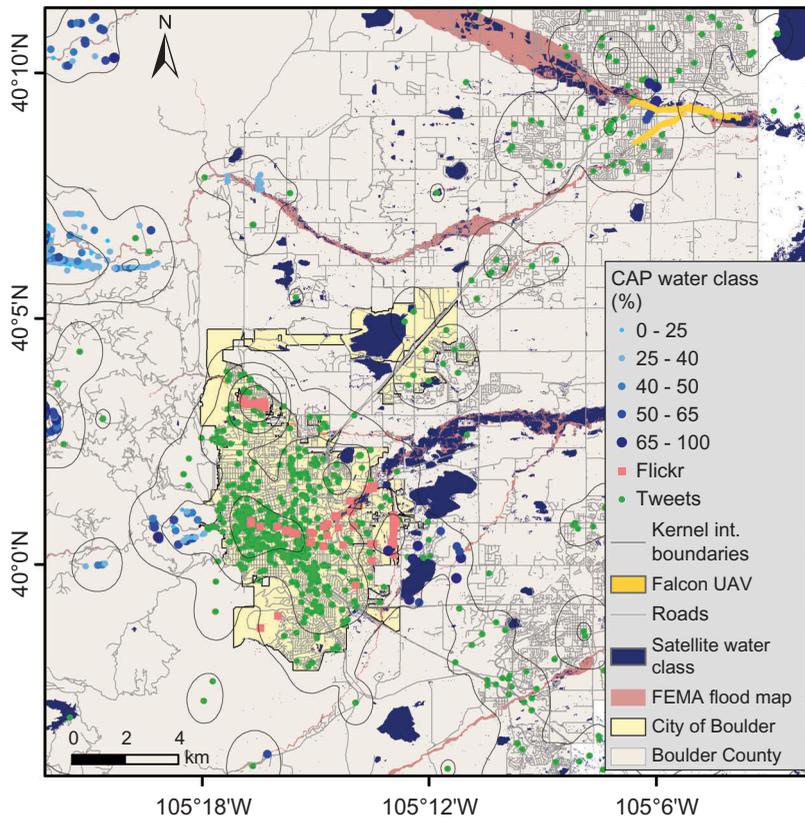


Figure 6. Data sources over the cities of Boulder and Longmont, CO, used for damage assessment, including Flickr ground photographs, tweets, CAP, Falcon UAV, and the satellite water classification. The road network is also shown.

While there may be uncertainties associated with the information obtained from tweets, the presence of flooding and damage is more easily verified in photographs. This was captured by the kernel interpolation by using a larger bandwidth to weight these data more heavily. Finally, [Figure 7\(d\)](#) overlays the predicted closures with roads that were closed by the Boulder EOC. The kernel prediction included all roads that were effectively closed, but overestimated two regions predicting moderate damage that did not occur.

The results obtained are very important because they show that transportation damage assessment can be effectively performed even in the absence of remote-sensing classification. Furthermore, the contributed data captured the subjective perception of the people who experienced the flood event. Fusing remote-sensing data with Twitter and other VGI presents the additional challenge of fusing quantitative data acquired through precise measurements of EM radiation, with qualitative data that are very subjective and only based on perceptions.

[Table 2](#) shows a small sample of tweets collected on 12 and 13 September 2013 during the peak of the flooding. One of the tweets described the event as being the worst ever seen. While the flooding of downtown Boulder caused severe damage, in terms of flood area it was rather limited, and in fact it was barely detected by the satellites.

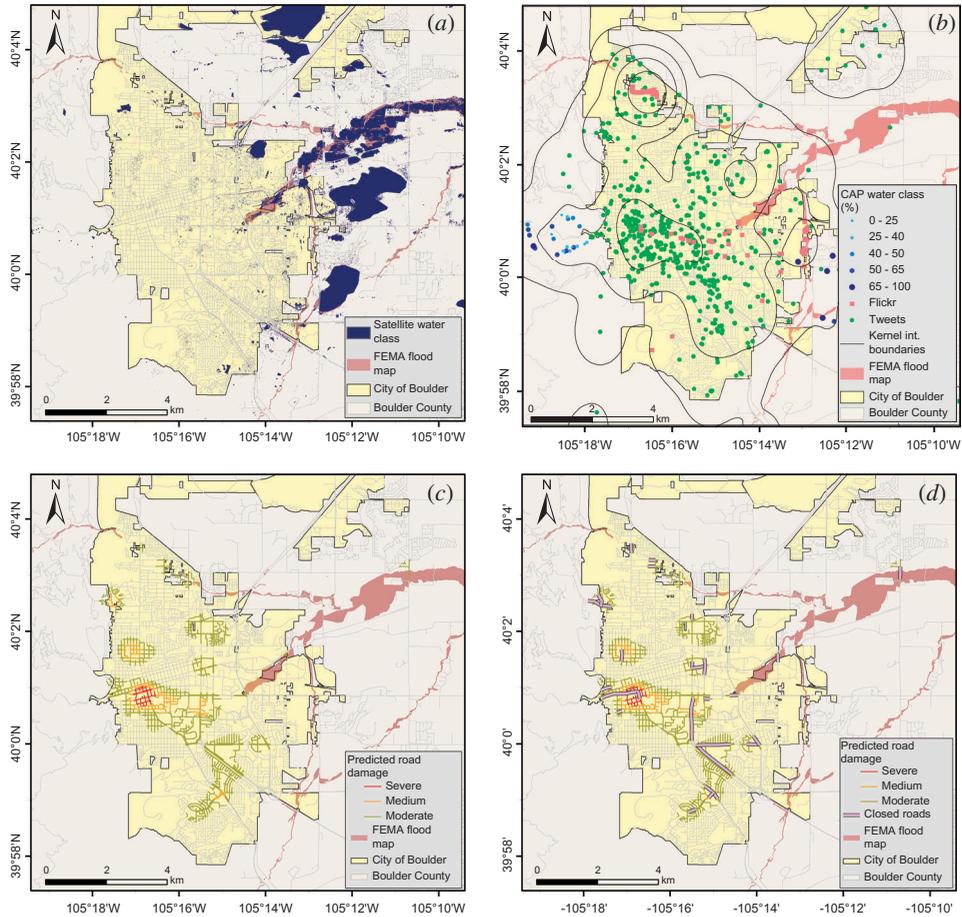


Figure 7. Map of Boulder City and County showing the satellite water classification and the FEMA flood map (a), interpolated damage surface created using CAP, ground photographs, and Twitter data (b), predicted roads damage based on the interpolation, classified as moderate, medium, and severe, classifications (c), roads that were closed by the Boulder Emergency Operations Center (EOC) (d). The road network is also shown in each figure.

Table 2. Sample of Tweets located in Boulder illustrating the subjective nature of the tweets.

| Latitude N | Longitude W | Date | Time | Tweet text |
|------------|-------------|-------------------|-------|--|
| 39°59'56" | 105°15'02" | 12 September 2013 | 17:36 | I have never seen anything like this. Dear God. #boulderflood |
| 40°01'03" | 105°16'48" | 12 September 2013 | 20:45 | Left my wife and brother on the other side of the newly formed crick #BoulderFlood |
| 40°00'57" | 105°16'16" | 13 September 2013 | 2:01 | Worst comes to worst I get submerged in the water and become a mermaid. #IAintEvacuating |
| 40°00'60" | 105°16'01" | 13 September 2013 | 2:34 | Surviving a small/medium sized tsunami in #Boulder right about now. |
| 39°58'34" | 105°15'50" | 12 September 2013 | 17:39 | Woah just got a text from CU Boulder saying 'wall of water coming down canyon' #boulderflood |
| 40°00'40" | 105°15'29" | 13 September 2013 | 18:21 | This water is insane #boulderflood http://t.co/WCJfq3hlyv |

Other tweets described the event as a ‘tsunami’, the water being ‘insane’, and creeks being formed downtown. Some users used humorous language to describe the situation, while at the same time stating that they did not intend to evacuate. All this information, while being qualitative and very subjective, could still be harvested for damage assessment, as seen for the prediction of closed roads. Finally, some of the tweets were annotated with links to a picture (see the last event in [Table 2](#)). In this work all tweets have been weighted equally; however, it is possible to give higher weights to messages that link to photographs.

4.2.2. Road damage assessment: case 2

A transportation damage assessment was performed for the city of Longmont, CO. Longmont is located about 20 km NE of Boulder, and suffered extensive damage as a result of the floods. This damage assessment presented different challenges than those from Boulder. First of all, there was a smaller number of contributed data due to the lower population density and different demographics. Second, this region was observed using the Falcon UAV, which provided a timely high-resolution view of the flooded area. Furthermore, the WorldView2 data contained a large amount of clouds exactly where most of the damage was recorded.

[Figure 8](#) shows a map for the study area (*a*), and the footprint of the Falcon UAV data used (*b*). There were only a limited number of tweets and CAP images for this area. A comparison between the satellite water classification and the FEMA flood map shows that the former underestimated the flood extent due to the presence of thick clouds in the centre of the image. When clouds were not present, the classification agreed with the FEMA map, and several roads were identified as being impassable. Five labels, i–iv, are included in [Figure 8\(a\)](#), which identify four tweets and one location where both CAP and UAV data were available. The tweets shown in [Table 3](#) describe extensive damage to both roads and railways, and include links to pictures that can be used to assess the damage.

[Figures 8\(c\)–\(f\)](#) show the pictures associated with the tweets (i–iv) shown in [Table 3](#) and in [Figure 8\(a\)](#). [Figure 8\(c\)](#) shows extensive damage to highway 36, which was not captured by the satellite classification. Mostly because of this image, the kernel interpolation predicted that this stretch of the road was impassable. [Figure 8\(d\)](#) shows extensive damage to a railway, which was totally occluded by clouds, and just at the edge of the data collected by the UAV, and thus not visible in either of the data sets. [Figures 8\(e\)](#) and [\(f\)](#) are two images associated with the same number of tweets that show two impassable roads. The satellite classification did not include these roads in the flooded area, and thus they were predicted as damaged only because of the cluster of tweets (see [Figure 8\(a\)](#)).

[Figures 8\(e\)–\(f\)](#) are from the same region, and illustrate damage to a bridge and roads in the surrounding areas. The first two images were collected by the Falcon UAV on 13 September 2013. The last image shows the same bridge collected by CAP aerial imagery on 19 September 2013. While the bridge is almost entirely submerged in the UAV imagery, it appears driveable in the CAP image. In fact, the CAP imagery was acquired six days after the UAV data, when the water might have already receded in this region. This result suggests the critical role that UAV data play during emergencies, because UAVs can be deployed quickly and in adverse atmospheric conditions when manned

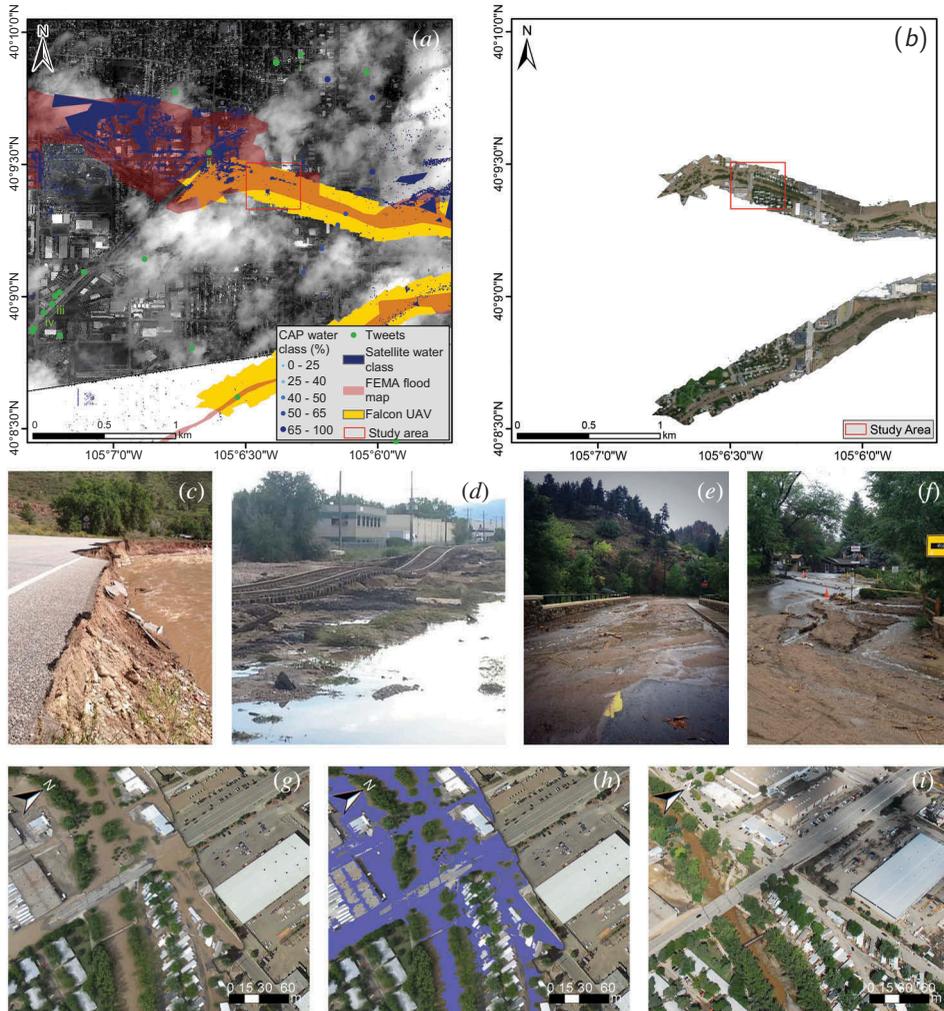


Figure 8. Map of the City of Longmont showing as basemap WorldView2 data band 3 and the data used for damage assessment, including CAP, tweets, Falcon UAV, and the FEMA flood map (a), the Falcon UAV data used in this study (b), images associated with tweets showing different damage to roads and railways (c–f), view of a damaged bridge in Falcon UAV imagery (g), Falcon UAV classified image (h), and the same bridge viewed in CAP data (i).

Table 3. Tweets used for the damage assessment in Longmont.

| Latitude N | Longitude W | Date | Time | Label | Tweet text |
|------------|-------------|-------------------|-------|---------|--|
| 40°09'55" | 105°06'18" | 17 September 2013 | 15:20 | i | Hwy 36 in #Lyons. #stvrainflood #coflood lyons bridge at hwy 36 http://t.co/ksN1Uo60kA |
| 40°09'32" | 105°06'40" | 14 September 2013 | 16:02 | ii | Wow! Look what the #coflood did to the BNSFRailway looks more like a rollercoaster! http://t.co/4AKdZmhkmc |
| 40°08'56" | 105°07'16" | 12 September 2013 | 19:04 | iii, iv | #boulderflood Boulder is in trouble folks. My employer isn't even attempting to open today. Massive flooding. 100 year flood? Looks like it. http://t.co/L8mvtzVFwE http://t.co/Fmbh6vPwrX |

airplanes are grounded. Additionally, if data with enough temporal and spatial coverage are available, it is possible to map the damage over time.

4.2.3. Road damage assessment: case 3

A road transportation assessment was performed for another part of Longmont CO, where an unobstructed satellite view led to the classification of several areas as flooded. [Figure 9\(a\)](#) shows a map for a specific region where the kernel interpolation identifies one road as being potentially damaged due to the area being classified as flooded both

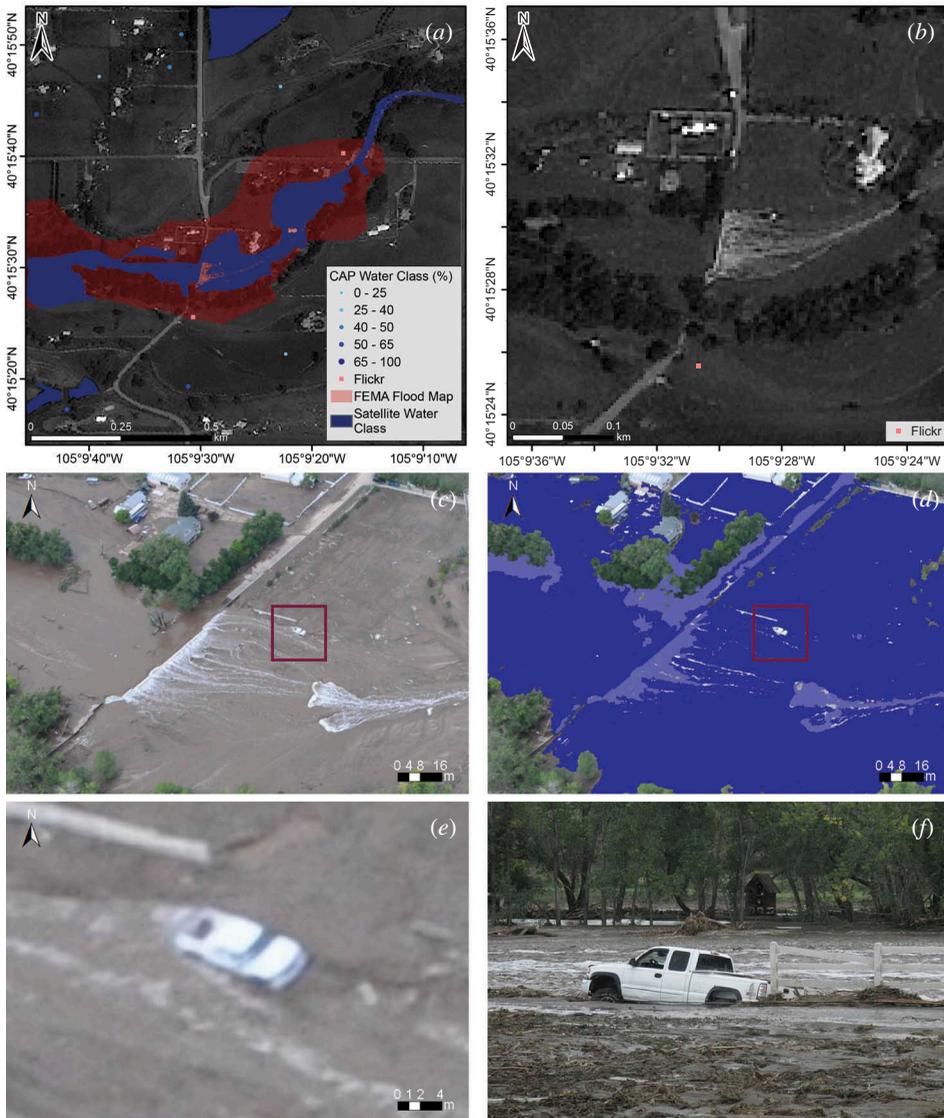


Figure 9. Map showing WorldView2 band 3 for an additional area northeast of Longmont, which suffered severe damage (a), and a close-up view of a submerged road (b), CAP image showing the same submerged road and identifying with a square a stranded truck (c), the same CAP image classified for water pixels (d), a CAP image close-up view of the stranded truck (e), and the same truck in a Flickr post (e).

in the satellite and in the CAP images, as well as being tagged flooded in two Flickr photographs. The figure shows the WorldView2 band 3 as background along with the FEMA flood map. In this particular case the damage assessment was accomplished by fusing satellite and contributed data, both of which were collected at roughly the same time.

Figure 9(b) is a close-up view that illustrates a road that is totally submerged. The cluster of white pixels to the right of the water flowing over and past the road is likely to be a stranded truck. Figure 9(c) shows a CAP image where it is possible to discern the submerged road and a stalled truck, highlighted with a square, and (d) shows the same image automatically classified by the machine learning classifiers with the extensive flooded area correctly identified. Figure 9(c) shows a close-up for the area with the stranded truck, and (d) shows a picture taken on the ground and posted on Flickr of the same partially submerged truck. The CAP images and the ground Flickr photograph were collected only a few hours apart. The kernel interpolation reflects the fact that satellite, CAP, and Flickr images all suggest the area is compromised, and correctly identifies this area as flooded and this particular road as being impassable.

5. Conclusions

This article presents a methodology that uses Twitter data to prioritize remote-sensing tasking and VGI data collection. A transportation damage assessment is performed by fusing the acquired remote-sensing data, Twitter data, and other contributed data such as aerial and ground images. Satellite data and available images are classified for water pixels using supervised machine learning. A spatial interpolation is performed to fuse the data by weighting each source according to a perceived trust level. Transportation damage assessment is performed by overlaying the interpolation results and the satellite data flood extent over a road layer.

Twitter is shown as an effective source of data to identify 'hot spots' and to task remote-sensing data collection for the 2013 Boulder floods. While only satellite data were tasked for this event, the same methodology can be used to task the collection of aerial imagery through UAV and manned platforms, as well as ground imagery. Crucial to the proposed methodology is the geolocation of tweets and other contributed data. A method is proposed that tries to augment Twitter data with no geolocation information by performing a text analysis, and matching extracted place names with a geodatabase. Using this method an additional 8% of Twitter data were geolocated.

Results show roads, railways, and bridges damage assessment performed over the cities of Boulder and Longmont. In the case of Boulder, the road damage prediction was compared with a list of officially closed roads, and a good match between predictions and observations is found. Results for Longmont show that tweets and photographs are crucial to estimate road closures when clouds are present and satellite flood extent cannot be performed. Furthermore, the predicted road closures extend beyond the FEMA flood map. The use of UAV data is particularly suitable for the estimation of transportation infrastructure due to the high spatial resolution, and the ability of UAVs to fly below clouds and in adverse meteorological conditions when manned aircraft might be grounded. UAVs can

play a crucial role in generating temporal data by repeating the data acquisition to quantify damage over time.

While this article focuses on flood events, the same methodology can be applied to different natural hazards and emergencies. The steady deployment of more high-resolution satellite and aerial sensors to monitor the Earth and its environment, paired with the dramatic increase of VGI, will generate an unprecedented amount of data that can be used during emergencies. The automated analysis of this massive amount of geospatial data is paramount, suggesting that the proposed methodology and its extensions will find wide applicability for years to come.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Work performed under this project has been partially funded by the Office of Naval Research (ONR) award #N00014-14-1-0208 [PSU #171570].

References

- Acar, A., and Y. Muraki. 2011. "Twitter for Crisis Communication: Lessons Learned from Japan's Tsunami Disaster." *International Journal of Web Based Communities* 7 (3): 392–402. doi:10.1504/IJWBC.2011.041206.
- Cervone, G., and B. Haack. 2012. "Supervised Machine Learning of Fused RADAR and Optical Data for Land Cover Classification." *Journal of Applied Remote Sensing* 6 (1): 063,597/1–18. doi:10.1117/1.JRS.6.063597.
- Cervone, G., and G. Manca. 2011. "Damage Assessment of the 2011 Japanese Tsunami Using High-Resolution Satellite Data." *Cartographica: The International Journal for Geographic Information and Geovisualization* 46 (3): 200–203. doi:10.3138/carto.46.3.200.
- Crumpacker, D. W. (1985) The Boulder Creek Corridor Projects: Riparian Ecosystem Management in an Urban Setting. Riparian Ecosystems and Their Management: Reconciling Conflicting Uses USDA Forest Service General Technical Report RM-120, RR Johnson, CD Ziebell, DR Patton, PF Ffolliott, and R II I-lamre (Technical Coordinators), pp 389–392
- Cutter, S. L. 2003. "Giscience, Disasters, and Emergency Management." *Transactions in GIS* 7 (4): 439–446. doi:10.1111/tgis.2003.7.issue-4.
- Dashti, S., L. Palen, M. Heris, K. Anderson, S. Anderson, and T. Anderson. 2014. "Supporting Disaster Reconnaissance with Social Media Data: A Design-Oriented Case Study of the 2013 Colorado Floods." In *Proceedings of the 11th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, University Park, PA, May, edited by S. R. Hiltz, M. S. Pfaff, L. Plotnick, and A. C. Robinson, 630–639. ISCRAM.
- De Longueville, B., R. Smith, and G. Luraschi. 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 (LBSN '09)*, edited by O. Wolfson, 73–80. New York: ACM.
- Duda, K. A., and B. K. Jones. 2011. "USGS Remote Sensing Coordination for the 2010 Haiti Earthquake." *Photogrammetric Engineering & Remote Sensing* 77 (9): 899–907. doi:10.14358/PERS.77.9.899.
- Duong, T., and M. L. Hazelton. 2005. "Cross-Validation Bandwidth Matrices for Multivariate Kernel Density Estimation." *Scandinavian Journal of Statistics* 32 (3): 485–506. doi:10.1111/sjos.2005.32.issue-3.

- Earle, P., D. Bowden, and M. Guy. 2012. "Twitter Earthquake Detection: Earthquake Monitoring in a Social World." *Annals of Geophysics* 54 (6): 708–715.
- Elaine, P. 2014. "How Boulder, Colorado, Prepared for 'Biblical' Rainfall." *Emergency Management*. Accessed December 5, 2015. <http://www.emergencymgmt.com/disaster/Boulder-Colo-Prepared-Biblical-Rainfall.html>
- Elwood, S. 2008. "Volunteered Geographic Information: Key Questions, Concepts and Methods to Guide Emerging Research and Practice." *GeoJournal* 72 (3–4): 133–135. doi:10.1007/s10708-008-9187-z.
- Elwood, S., M. F. Goodchild, and D. Z. Sui. 2012. "Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice." *Annals of the Association of American Geographers* 102 (3): 571–590. doi:10.1080/00045608.2011.595657.
- Gochis, D., R. Schumacher, K. Friedrich, N. Doesken, M. Kelsch, J. Sun, K. Ikeda, D. Lindsey, A. Wood, B. Dolan, S. Matrosov, A. Newman, K. Mahoney, S. Rutledge, R. Johnson, P. Kucera, P. Kennedy, D. Sempere-Torres, M. Steiner, R. Roberts, J. Wilson, W. Yu, V. Chandrasekar, R. Rasmussen, A. Anderson, and B. Brown. 2014. "The Great Colorado Flood of September 2013." *Bulletin of the American Meteorological Society* 96: 1461–1487.
- Goodchild, M. 2007. "Citizens as Sensors: The World of Volunteered Geography." *GeoJournal* 69 (4): 211–221. doi:10.1007/s10708-007-9111-y.
- Heverin, T., and L. Zach. 2010. "Microblogging for Crisis Communication: Examination of Twitter Use in Response to a 2009 Violent Crisis in the Seattle-Tacoma, Washington, Area." In *Proceedings of the 7th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, Seattle, WA, May, edited by S. French, B. Tomaszewski, and C. Zober, 1–5. ISCRAM.
- Huang, Q., and Y. Xiao. 2015. "Geographic Situational Awareness: Mining Tweets for Disaster Preparedness, Emergency Response, Impact, and Recovery." *ISPRS International Journal of Geo-Information* 4 (3): 1549–1568. doi:10.3390/ijgi4031549.
- Huang, Q., and C. Xu. 2014. "A Data-Driven Framework for Archiving and Exploring Social Media Data." *Annals of GIS* 20 (4): 265–277. doi:10.1080/19475683.2014.942697.
- Joyce, K. E., S. E. Belliss, S. V. Samsonov, S. J. McNeill, and P. J. Glassey. 2009. "A Review of the Status of Satellite Remote Sensing and Image Processing Techniques for Mapping Natural Hazards and Disasters." *Progress in Physical Geography* 33 (2): 183–207. doi:10.1177/0309133309339563.
- Keilis-Borok, V. 2002. "Earthquake Prediction: State-Of-The-Art and Emerging Possibilities." *Annual Review of Earth and Planetary Sciences* 30: 1–33. doi:10.1146/annurev.earth.30.100301.083856.
- Laituri, M., and K. Kodrich. 2008. "On Line Disaster Response Community: People as Sensors of High Magnitude Disasters Using Internet GIS." *Sensors* 8 (5): 3037–3055. doi:10.3390/s8053037.
- McDougall, K., and P. Temple-Watts. 2012. "The Use of Lidar and Volunteered Geographic Information to Map Flood Extents and Inundation." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 1: 251–256. doi:10.5194/isprannals-1-4-251-2012.
- Mohammad, S. M., S. Kiritchenko, and X. Zhu. 2013. "Nrc-Canada: Building the State-Of-The-Art in Sentiment Analysis of Tweets." In *Second Joint Conference on Lexical and Computational Semantics (*SEM)*, Vol. 2, Atlanta, GA, edited by M. Diab, 321–327.
- Mühlenstädt, T., and S. Kuhnt. 2011. "Kernel Interpolation." *Computational Statistics & Data Analysis* 55 (11): 2962–2974. doi:10.1016/j.csda.2011.05.001.
- Myers, D. E. 1994. "Spatial Interpolation: An Overview." *Geoderma* 62 (1–3): 17–28. doi:10.1016/0016-7061(94)90025-6.
- Oxendine, C., E. Schnebele, G. Cervone, and N. Waters. 2014. "Fusing Non-Authoritative Data to Improve Situational Awareness in Emergencies." In *Proceedings of the 11th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, University Park, PA, May, edited by S. R. Hiltz, M. S. Pfaff, L. Plotnick, and A. C. Robinson, 760–764. ISCRAM.
- Oxendine, C., M. Sonwalkar, and N. Waters. 2012. "A Multi-Objective, Multi-Criteria Approach to Improve Situational Awareness in Emergency Evacuation Routing Using Mobile Phone Data." *Transactions in GIS* 16 (3): 375–396. doi:10.1111/tgis.2012.16.issue-3.
- Poser, K., and D. Dransch. 2010. "Volunteered Geographic Information for Disaster Management with Application to Rapid Flood Damage Estimation." *Geomatica* 64 (1): 89–98.

- Pultar, E., M. Raubal, T. Cova, and M. Goodchild. 2009. "Dynamic GIS Case Studies: Wildfire Evacuation and Volunteered Geographic Information." *Transactions in GIS* 13 (1): 85–104. doi:10.1111/tgis.2009.13.issue-s1.
- Raykar, V. C., and R. Duraiswami. 2006. "Fast Optimal Bandwidth Selection for Kernel Density Estimation." In *SDM*, edited by J. Srivastava and D. Skillicorn, 524–528. Bethesda, MD: SIAM.
- Ripley, B. 2008. *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press.
- Schnebele, E., and G. Cervone. 2013. "Improving Remote Sensing Flood Assessment Using Volunteered Geographical Data." *Natural Hazards Earth System Science* 13: 669–677. doi:10.5194/nhess-13-669-2013.
- Schnebele, E., G. Cervone, S. Kumar, and N. Waters. 2014a. "Real Time Estimation of the Calgary Floods Using Limited Remote Sensing Data." *Water* 6: 381–398. doi:10.3390/w6020381.
- Schnebele, E., G. Cervone, and N. Waters. 2014b. "Road Assessment after Flood Events Using Non-Authoritative Data." *Natural Hazards and Earth System Science* 14 (4): 1007–1015. doi:10.5194/nhess-14-1007-2014.
- Smith, L. 1997. "Satellite Remote Sensing of River Inundation Area, Stage, and Discharge: A Review." *Hydrological Processes* 11 (10): 1427–1439. doi:10.1002/(ISSN)1099-1085.
- Stryker, T., and B. Jones. 2009. "Disaster Response and the International Charter Program." *Photogrammetric Engineering and Remote Sensing* 75 (12): 1342–1344.
- Tate, C., and T. Frazier. 2013. "A GIS Methodology to Assess Exposure of Coastal Infrastructure to Storm Surge & Sea-Level Rise: A Case Study of Sarasota County, Florida." *Journal of Geography & Natural Disasters* 1: 2167–2587.
- Tatham, P. 2009. "An Investigation into the Suitability of the Use of Unmanned Aerial Vehicle Systems (Uavs) to Support the Initial Needs Assessment Process in Rapid Onset Humanitarian Disasters." *International Journal of Risk Assessment and Management* 13 (1): 60–78. doi:10.1504/IJRAM.2009.026391.
- Tobler, W. R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46: 234–240. doi:10.2307/143141.
- Triglav-Čekada, M., and D. Radovan. 2013. "Using Volunteered Geographical Information to Map the November 2012 Floods in Slovenia." *Natural Hazards and Earth System Science* 13 (11): 2753–2762. doi:10.5194/nhess-13-2753-2013.
- Tyshchuk, Y., C. Hui, M. Grabowski, and W. Wallace. 2012. "Social Media and Warning Response Impacts in Extreme Events: Results from a Naturally Occurring Experiment." In *Proceedings of the 45th Annual Hawaii International Conference on System Sciences (HICSS)*, Maui, HI, January, 818–827. IEEE.
- Uddin, W. 2011. "Remote Sensing Laser and Imagery Data for Inventory and Condition Assessment of Road and Airport Infrastructure and GIS Visualization." *International Journal of Roads and Airports (IJRA)* 1 (1): 53–67.
- Verma, S., S. Vieweg, W. Corvey, L. Palen, J. Martin, M. Palmer, A. Schram, and K. Anderson. 2011. "Natural Language Processing to the Rescue? Extracting "Situational Awareness" Tweets during Mass Emergency." In *Proceedings of the 5th International Conference on Weblogs and Social Media*, Barcelona, July, edited by N. Nicolov and J. Shanahan, 1–8. AAAI.
- Vieweg, S., A. Hughes, K. Starbird, and L. Palen. 2010. "Microblogging during Two Natural Hazards Events: What Twitter May Contribute to Situational Awareness." In *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, Atlanta, GA, April, edited by E. Mynatt, 1079–1088. ACM.
- Waters, N. 2009. "Representing Surfaces in the Natural Environment: Implications for Research and Geographical Education." *Representing, Modeling and Visualizing the Natural Environment: Innovations in GIS* 13: 21–39.
- Waters, N. 2014. "Social Network Analysis." In *Handbook of Regional Science*, edited by M. Fischer and P. Nijkamp, 725–740. Springer.
- Waters, N. forthcoming. "Tobler's First Law of Geography." In *The International Encyclopedia of Geography*, edited by D. Richardson. Washington, DC: Association of American Geographers.