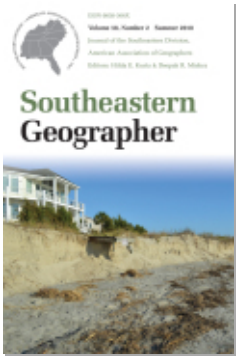
Geospatial Assessment of Wetness Dynamics in the October 2015 SC Flood with Remote Sensing and Social Media



Cuizhen Wang, Zhenlong Li, Xiao Huang

Southeastern Geographer, Volume 58, Number 2, Summer 2018, pp. 164-180 (Article)

Published by The University of North Carolina Press
DOI: https://doi.org/10.1353/sgo.2018.0020

Geospatial Assessment of Wetness Dynamics in the October 2015 SC Flood with Remote Sensing and Social Media

CUIZHEN WANG
Department of Geography, University of South Carolina

ZHENLONG LI, DEPARTMENT OF GEOGRAPHY
University of South Carolina

XIAO HUANG, DEPARTMENT OF GEOGRAPHY
University of South Carolina

Real-time data on flood extents and dynamics are important for risk assessment and emergency response during the event. While real-time imagery is often unavailable due to heavy cloud cover during a flood, remote sensing platforms can be used to monitor its development through a synoptic view. Record rainfall occurring October 1–5, 2015 in coastal South Carolina caused the October 2015 South Carolina Flood. The Congaree River Watershed downstream of Columbia, SC experienced historic flooding. This study utilizes two satellite images acquired on October 8th (EO-1 ALI) and 18th (Landsat8 OLI) to examine flood dynamics. Using a normalized difference wetness index (NDWI), the flooded and highly wet areas were extracted. Social media, such as Twitter, from public users allows quick awareness of floods in an area, but geolocation is not necessarily accurate. Assisted with real-time Twitter data, satellite images after a flood helps to assess water retreat and potential risks for emergency responders. Since social media data sets are big, highly unstructured and noisy in nature, sophisticated data mining algorithms are needed for the verification process from millions of tweets in a region. When automatic tweets verification approaches are available, integrating social media into geospatial science could become important data sources for disaster assessment and management.

Información en tiempo real de las extensiones y dinámicas de las inundaciones son importantes para la evaluación de riesgos y la respuesta a emergencias durante un evento. Mientras que las imágenes de tiempo real suelen no ser disponibles debido a una cubierta de nubes pesadas durante una inundación, las plataformas de teledetección pueden ser usados para seguir su desarrollo por una vista sinóptica. Una cantidad histórica de lluvia entre el 1-5 de octubre, 2015 en Carolina del Sur causó la inundación de Carolina del Sur de octubre 2015. La cuenca del Río Congaree abajo de Columbia, Carolina del Sur pasó una inundación histórica. El presente estudio usa dos imágenes satelitales adquiridos el 8 de octubre (EO-1 ALI) y 18 de octubre (Landsat8 OLI) para examinar las dinámicas de las inundaciones. Usando un índice normalizado de diferencia de humidad (NDWI), los lugares inundados y mojados fueron sacados. Los medios de comunicación social, como Twitter, de usuarios públicos permite una transmisión rápida de información, pero la geo-locación no siempre es precisa. Junto a información de tiempo

vivo de Twitter, imágenes satelitales después de las inundaciones ayuda entender el retiro de agua y los posibles riesgos para los personales de emergencia. Como los conjuntos de datos de los medios de comunicación son grandes y difíciles de manejar, es necesario tener algoritmos sofisticados para verificar los millones de tweets de una región. Cuando están disponibles sistemas de verificación de tweets, integrar los medios de comunicación social a la ciencia geoespacial podría convertirse en una fuente de información importante para la evaluación y manejo de desastres.

KEYWORDS: October 2015 SC Flood; satellite imagery; Twitter; flood dynamics; GEOINT

PALABRAS CLAVE: Inundación de octubre 2015 en Carolina del Sur; imágenes satelitales; Twitter; dinámicas de inundaciones; GEOINT

INTRODUCTION

Spatial extents and temporal dynamics of a flood are important for rapid risk assessment and post-disaster damage evaluation. In the October 2015 South Carolina Flood, the Congaree River Watershed downstream of Columbia, SC was severely flooded due to intensive precipitation from Hurricane Joaquin. According to the U.S. Geological Survey (USGS) stream flow data, Congaree River in the downstream reached its peak flow at 185,000 cubic feet per second (cfs) on October 4th (Musser et al. 2016). Water level data, however, were only available at a limited number of stream gauges, some of which were saturated by the severe floods. Traditional survey-based flood mapping usually takes place long after the event. For example, the first flood inundation map of the 2015 SC Flood was built upon hundreds of high-water marks surveyed by USGS in the pre-selected flood zone and was released in February 2016, a few months after the flood (Musser et al. 2016). The information is important for long-term assessments, but could be too late for decision makers to evaluate flood damages and risks for rapid emergency response. Moreover, a recent study showed that the 100-year floodplain maps from the Federal Emergency Management Agency (FEMA) failed to capture 75 percent of flood damages claimed in Houston suburbs in 1999–2009 (Blessing et al. 2017). Since U.S. officials use these flood zone maps to determine flood risk and insurance premiums, the study raises high concerns on the efficiency and accuracy of damage assessment during a flood, e.g., the catastrophic record-breaking flood in Houston from Hurricane Harvey in August 2017.

Satellite observations provide a synoptic view of Earth surfaces and changes in a large spatial extent. When atmospheric conditions allow (e.g. without heavy clouds), satellite imagery can provide spatially continuous coverage of flooded areas. Multi-temporal observations easily reveal flood development. Additionally, beyond binary outputs such as flooded or not, multi-spectral imagery could extract a variety of indices to quantify flooding conditions across the spatial coverage. Ji et al. (2009), for example, reviewed a set of water indices for extracting water surfaces from reflective bands including green, red, green, near-infrared (NIR) and shortwave-infrared (SWIR). They found that the green-SWIR normalized difference wetness index (NDWI) (Gao 1996) optimally reflected moisture conditions and extracted waterbodies in medium-resolution

satellite images. Past studies also showed that SWIR is most sensitive to surface moisture because of water absorption in this spectral region (Wang et al. 2007).

One major challenge for image-based flood assessment is that data availability is restricted by long revisit cycles of spaceborne satellites, and restricted viewing angle. For example, the most commonly applied Landsat imagery is only available in every 16 days. While more spaceborne sensors in recent years (e.g., the commercial small satellites) have the capability of changing view angles to acquire imagery on nearby orbits, heavy cloud cover and storms make these sensors infeasible to image land surfaces during a flood event. The earliest image is often acquired a few days after a flood, which hinders the real-time flood assessment from remote sensing imagery. Aerial image acquisition from low-altitude platforms such as airplanes and emerging drone technologies can occur on demand, and cover area of interest, yet is often operationally restricted during severe storms. Regardless platform used to acquire imagery, the image-extracted information after an event must be calibrated and corrected using real-time observations to better reflect flood characteristics.

With the recent development of Citizen Science, volunteered geographical information (VGI) drawn from the concept of crowdsourcing (Goodchild 2007) is becoming widely used in monitoring extreme events (Poster and Dransch 2010; Triglav-Čekada and Radovan 2013). Social media - such as Twitter - captures localized, real-time information in forms of photos, videos and texts about an event, and the geotagged tweets contain geographic information (Imran et al. 2013). Integrating tweets, elevation and stream flow data in hydrological modeling, flood probabilities in a geographic area have been mapped in stochastic approaches (de Albuquerque et al. 2014; Smith et al. 2017). Specifically, to study the 2015 SC Flood, Li et al. (2017) utilized the verified flood-related tweets and developed a geostatistical model for rapid flood mapping based on topographic effects on water inundation in flood zone. Social media for disaster analysis, however, has been criticized for the reliability of the non-authoritative data collected by the public, which usually do not comply with standard spatial data assurance procedures (Haklay et al. 2010; Poser and Dransch 2010; Schnebele 2014). Studies have been conducted to assess VGI's uncertainties in both locations and attributes (See et al. 2013; Roberts and Doyle 2017). For locational uncertainties, Haklay et al. (2010) tested with the OpenStreetMap dataset and found that VGI followed the Linus Law - with an increased number of contributors, the data would hold an intrinsic quality assurance - although the relationship between the contributor size and data quality was not linear. For informational uncertainties, Fonte et al. (2015) developed a quality control framework and suggested that VGI under good practices could be reliable for integrating with authoritative maps. In 2017, the Associated Programme on Flood Management (APFM 2017) released the Crisis Mapping and Crowdsourcing in Flood Management Tool to provide guidance materials for crowdsourcing practitioners in perception of risk awareness from social actors for improved flood management. Rosser et al. (2017) performed rapid probabilistic assessment of flood inundation by fusing geotagged Flicker photographs, satellite

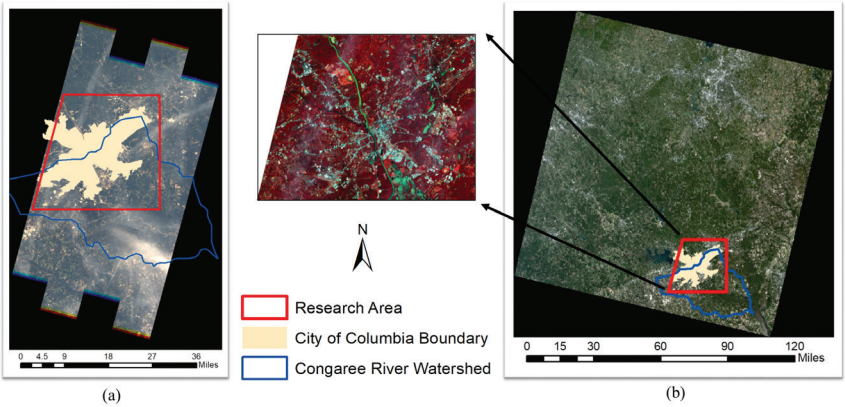


Figure 1. Two satellite images (true color display) in the study area: the EO-1 ALI on October 8 (a) and Landsat OLI on October 18 (b). The inset is the ALI image in the study area (standard false color).

imagery and elevation data in a statistical model, revealing good potentials of integrating VGI with remote sensing for near-real-time flood hazards mapping (Schnebele and Cervone 2013).

This study evaluated the spatiotemporal dynamics of wetness and flood risks from post-event satellite images using Twitter data as the real-time evaluation source. The NDWI is extracted to assess the spatial extents and temporal dynamics of the SC Flood from two satellite images acquired on October 8 and 18, 2015. The wetness levels are ranked to identify the flood risks in the study area, and are evaluated at the locations of the verified flood-related tweets, hereafter referred to as tweet points. Changes of land surface wetness in the 10-day period are also examined to assess the spatial transition of high-risk areas after the peak event.

STUDY AREA AND DATA SETS

The study area is the upper Congaree River Watershed, an urban watershed in Columbia, SC, which includes most of the urban lands of this capital city (Figure 1). Affected by Hurricane Joaquin on October 2 - 5, 2015, SC experienced widespread, record-breaking rainfall of 20-25 inches in four days (Jonathan et al. 2016). A large number of flood events erupted across central and eastern SC, devastated thousands of homes, infrastructures and agricultural lands, and caused direct losses of more than 300 million dollars and around one billion dollars for post-flood recovery (Feaster et al. 2015). Columbia was among the most damaged areas of the state, reaching a 1000-year Flood level while other parts of the state were at a 500-year Flood level (Li et al. 2017). Based upon the 2011 National Land Cover Database (NLCD) product (Homer, et al. 2011), land cover within the Columbia Metropolitan area is mostly developed lands with impervious surfaces (e.g. pavement, buildings). Outside the city, land covers are mostly herbaceous, agricultural and forest lands. The downstream parts of the watershed are woody wetlands.

Two medium-resolution satellite images covering the watershed were downloaded from the USGS Data Clearinghouse (EarthExplorer). The EO-1 ALI image (Figure 1a) was acquired on October 8 and the Landsat8 OLI image was acquired on October 18 (Figure 1b). These images represented two post-flood stages of water retreat and land surface moisture movement. Both images have a 30-m resolution in multispectral bands. Taking the OLI image as reference, the ALI image was geometrically and atmospherically corrected. Image pre-processing steps are described in next section.

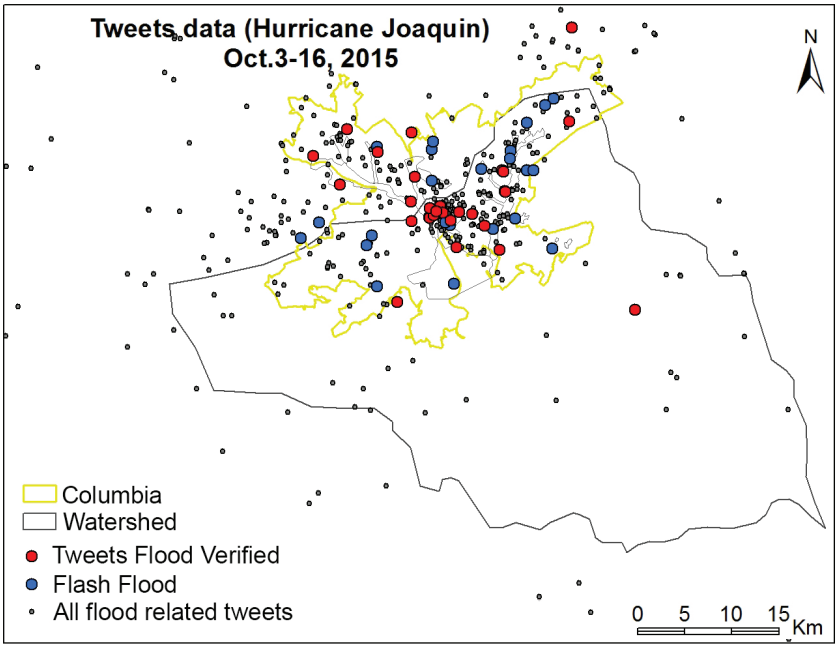
Official ground truthing data were limited due to physical access difficulties caused by the flood. In February 2016 USGS released the inundation map based on field surveys within a pre-selected boundary in Columbia (Musser et al. 2016). Although this inundation map did not survey our entire study area, this official inundation map represents the flooded areas during the event. Additionally, in the Flash Flood Observation Database at the National Weather Service (NWS), the USGS Flash Flood data points were updated to extend through May 2016 (FLASH 2017). Within the study area, we extracted 23 geo-referenced flash flood points that served as ground truthing points (blue marks in Figure 2a).

As an emerging data source, social media allows us to quickly gain a sense of floods at multiple sites, although their geo-locations may not necessarily be accurate. Following Hurricane Joaquin, about 1.3 million geotagged flood-related tweets were collected in the whole state of South Carolina using Twitter Stream API and REST API (Li et al. 2017). Over 2,000 geotagged flood-related tweets were collected in the Columbia Metro area (Figure 2a) using the commonly applied text-match technique, with case-insensitive keywords such as "flood" or "Joaquin". The wildcard "*" was applied to include their variants such as "flooding" or "floods". Tweets not having these keywords were not included. Therefore, some flood-related tweets could be missed if they did not have the matching texts. This has been commonly recognized as a limitation in text-match approaches to filtering social media data (de Albuquerque et al. 2015). To verify these flood-related tweets, colleagues in our research team manually checked them one by one to extract those showing evidence of flooding select tweets by finding matching photos of the tweeted message. For example, if a photo attached on a tweet showed that a street was flooded, then the photo was compared against different mapping sources such as the road map, Google Earth, and street views. Only those with the verified photos were treated as the verified flood-related tweets. Among more than 2,000 geotagged tweets in Columbia, only 33 tweets were verified to show the evidence of flooding (red marks in Figure 2a). In Figure 2b, the posting dates of these tweets peak on October 4 (19 posts) and 5 (5 posts), the two dates with the heaviest rainfall in Columbia (Berg et al. 2015). In this study, these 33 points were also used as ground "truthing" sources of floods to evaluate our satellite-extracted results.

METHODOLOGY

Data preprocessing

The Landsat8 OLI image downloaded from the USGS Data Clearinghouse is the surface reflectance product that has been atmospherically corrected. The ALI image



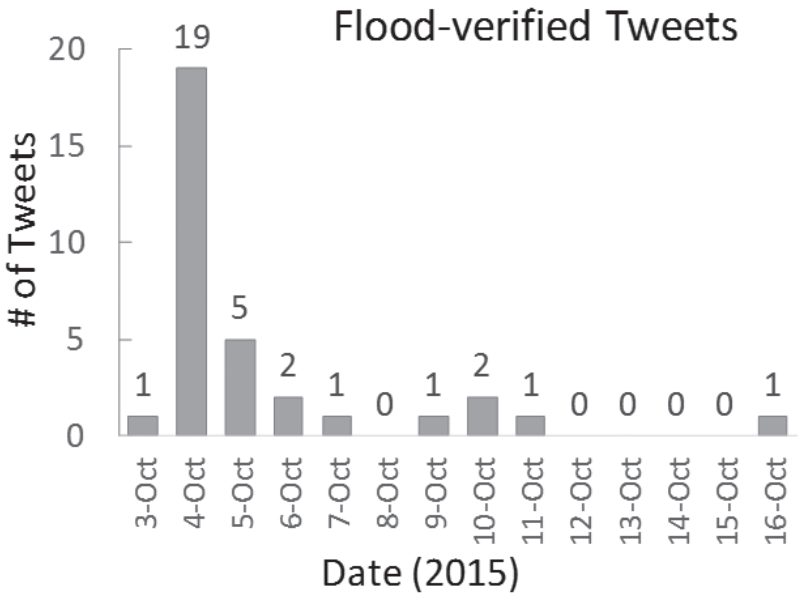


Figure 2. Spatial distributions of tweets and Flash Flood points in the study area (a) and the posting dates of tweets (b).

is not corrected and is contaminated by haze and thin cloud (as shown in Figure 1a). Using the ATCOR2 algorithm, an extension in ERDAS/IMAGINE 2016 software package, we performed the haze reduction to the ALI image. Atmospheric correction was then performed to convert the digital numbers of the ALI image to surface reflectance.

Taking the OLI image as the reference, geometric correction of the new ALI image was performed, reaching the root-mean-square-error (RMSE) less than one pixel. Even after atmospheric correction, surface reflectance values between the two images may not match exactly due to different sensor calibrations, atmospheric conditions and, or correction algorithms. The histogram match was then performed to match the histograms of all ALI bands to the corresponding OLI bands. These preprocessing steps reduce the noises in the ALI image that are not related to land surface conditions. In this way, surface reflectance of the two images becomes comparable so that it is possible to examine surface changes between the two images at different times after the 2015 Flood event.

NDWI and wetness levels

Land surface changes in a flood are mostly water inundation in flooded areas and wetness changes of land surfaces in non-flooded areas. Here we adopt the NDWI to represent land surface wetness in the study area. As discussed in Ji et al. (2009), it is the most stable index for extracting water surfaces from satellite images. With the green and SWIR bands, the NDWI is calculated as:

$$NDWI = \frac{\rho_{green} - \rho_{SWIR}}{\rho_{green} + \rho_{SWIR}} * 1000 + 1000 \quad (1)$$

where ρ_{green} and ρ_{SWIR} are surface reflectance of green and SWIR band, respectively. A scale factor of 1000 and a shift of 1000 are used to scale up the NDWI from [−1,1] to [0, 2000].

The NDWI is positively related to land surface wetness. Pixels with higher NDWI represent moister conditions. Waterbodies have the highest NDWI that could be easily delineated in the NDWI image. The average NDWI value of water is around 1,000 in randomly selected water pixels across the study area. Basic statistics show that the histogram of ΔNDWI (OLI-ALI) is normally distributed with a mean of 0.557 (close to 0) and standard deviation of 57.799. Therefore, the two NDWI images are highly comparable, and the large standard deviation indicates reasonable wetness transition between the two dates.

Based on the NDWI values in each image, we define five wetness levels, i.e. Low, Medium, Wet, Highly Wet, and Water (Table 1). Natural water bodies and flooded areas have NDWI > 1,000. Pixels in Wet and Highly Wet levels represent land surfaces that are not flooded but are under high risk of flooding because of their high moisture content. The same thresholding criteria are applied to the ALI and OLI images. This thresholding approach is somewhat subjective, but it accurately reveals the spatial patterns of wetness in the study area and their dynamics between the two images. Only the last three levels (Wet, Highly Wet, Water) are examined in this study. The Low and Medium levels are not under flooding risks because of their low NDWI values.

Table 1. The NDWI-extracted wetness levels in the two images.

Rank	Wetness level	NDWI range (ALI & OLI)
1	Low	<800
2	Medium	800–900
3	Wet	900–950
4	High Wet	950–1,000
5	Water	>1,000

Wetness dynamics evaluation

Surface wetness changes are compared among the three wetness levels in the two images. Areal changes of the wetness level reveal flood retreat, and changes of the Wet and Highly Wet levels indicate the transition of high-risk areas.

The 33 tweets represent the real-time flood at each tweeted location in the posting date. To apply the Twitter data in geographical analysis, however, two challenges need to be addressed. Firstly, the geo-locations of tweets are not accurate because, for example, people most likely take photos in distance from the flooded area depicted by the photo. It is thus difficult to match the tweet points to a specific pixel in the image. To spatially examine the wetness levels of the tweeted floods, we draw a 150-m buffer centered at each tweet point, and assign the maximum NDWI within the buffer as the wetness level of this point. We assume that this buffer size fairly defines the distance limit for a person to take pictures of a flooded area, especially in urban environments such as Columbia. The maximal NDWI indicates the highest wetness condition within the buffer, and represents the highest flooding possibility if a flood event occurs in this area. The 23 Flash Flood points reported by USGS are assumed to have accurate positions, so the buffer analysis is not performed.

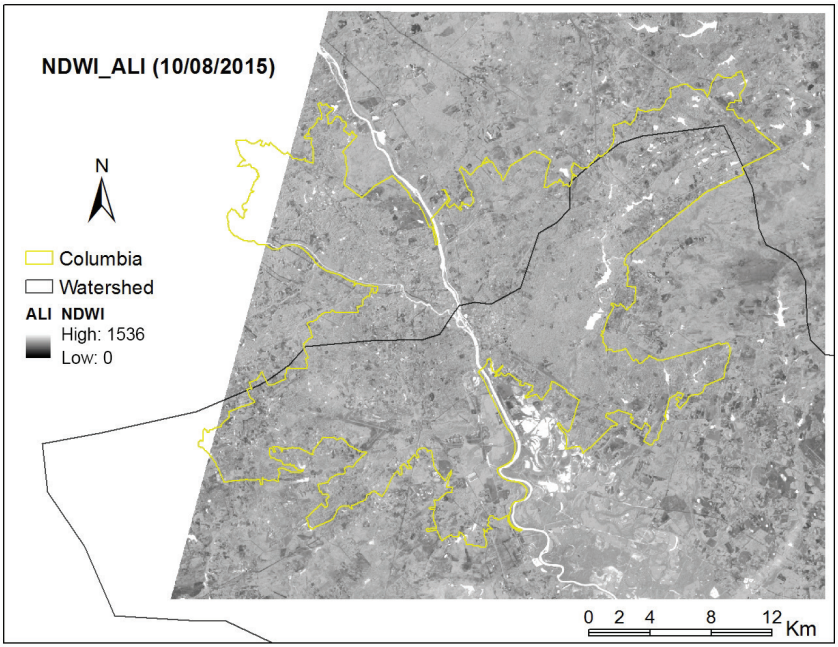
The second challenge is the temporal lags between satellite image acquisition and tweeted dates. As shown in Figure 2b, majority of tweets were posted on October 4 and 5. The earliest image (ALI), however, was acquired on October 8. Water in flooded areas during peak precipitation may retreat by the time of image acquisition, which is especially true for flash floods. Here our hypothesis is that even though water in flooded areas may have retreated, surface wetness in these areas remains high and therefore, could still be detected from satellite images acquired after the peak event. The changes of wet pixels (in the three wetness levels) between the ALI and OLI images captured the flood retreat after the 2015 Flood event.

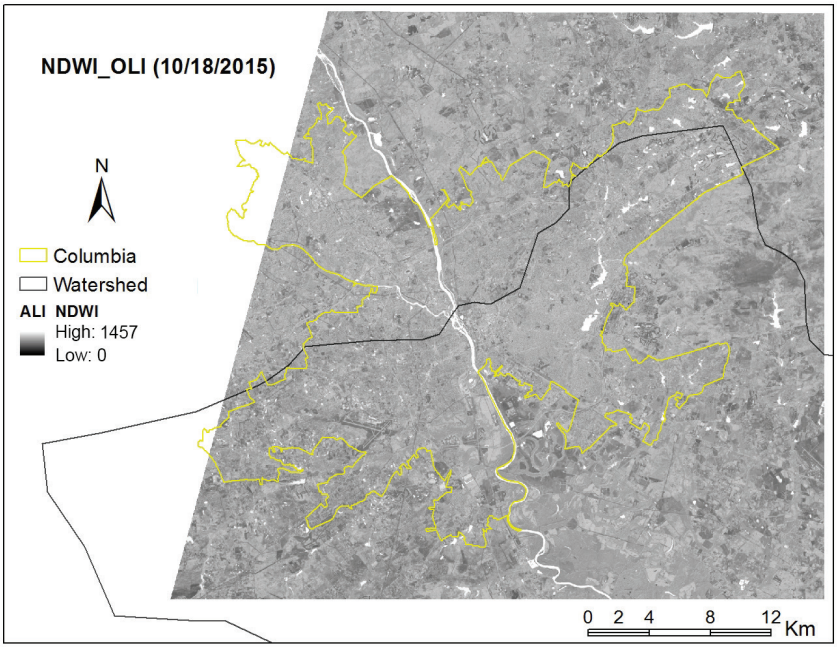
RESULTS AND DISCUSSION

Wetness level distributions and dynamics

Water clearly stands out in the two NDWI maps calculated from both the ALI (Figure 3a) and OLI (Figure 3b) images. A number of small, natural waterbodies along Mills Creek in the northeast of the city were easily identified in both maps. Large patches of flooded areas on October 8, for example those highlighted in a red dashed circle in southern Columbia, were also apparent. Water in these flooded areas retreated by October 18.

The wetness level maps were extracted from the two images based on the NDWI ranks in Table 1. In Figure 3c and 3d, wet pixels (the three wet levels) are clustered along the Congaree River in southern Columbia and in Mills Creek northeast of the city. Areas along Saluda River in the northwest of the city had minimal flooding.


NDWI_ALI (10/08/2015)
N
Columbia
Watershed
ALI NDWI
High: 1536
Low: 0
0 2 4 8 12 Km


NDWI_OLI (10/18/2015)
N
Columbia
Watershed
ALI NDWI
High: 1457
Low: 0
0 2 4 8 12 Km

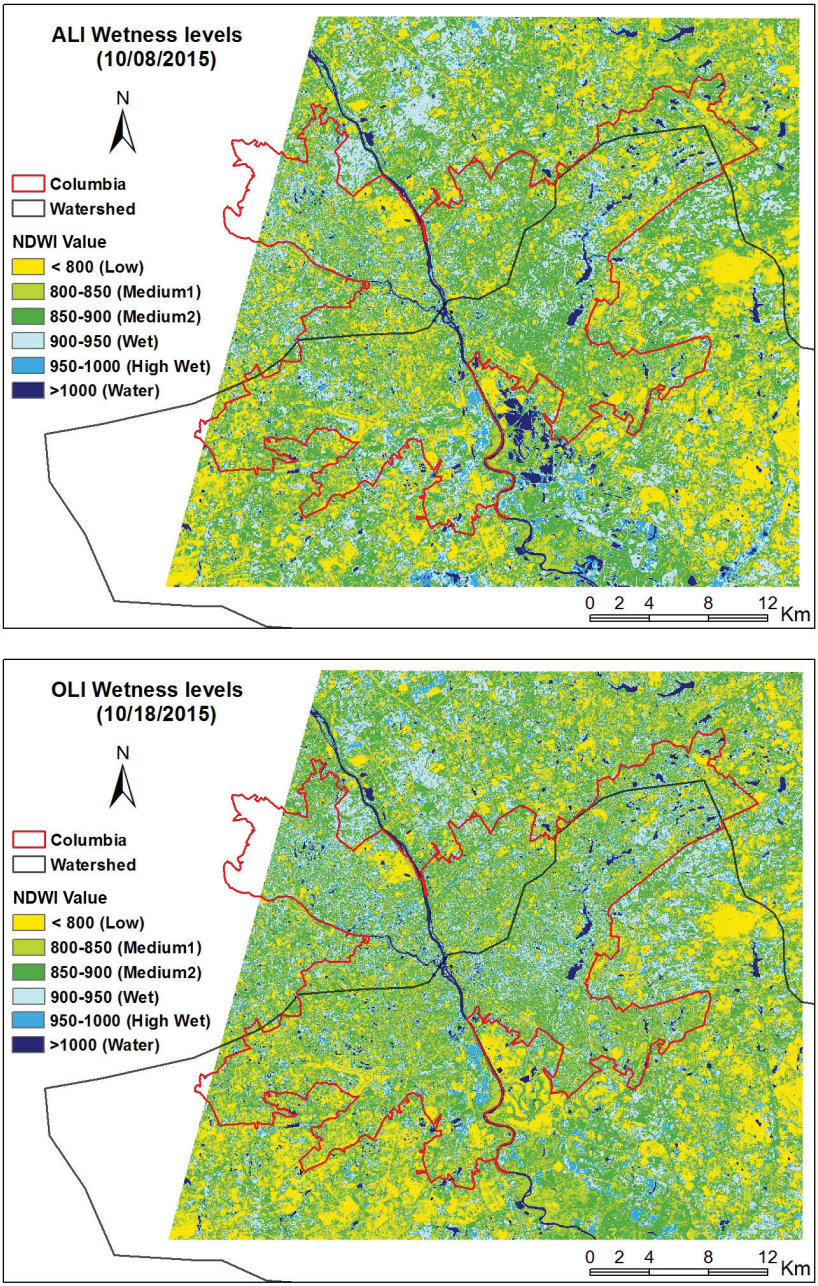


Figure 3. The NDWI distributions and Wetness Level maps on October 8 (a, c) and 18 (b, d). A red dashed circle in (a) and (b) demonstrated flood retreat in the 10-day period.

Similarly, from October 8 to 18 water retreat was clear. According to the tweets in Figure 2, most flood events occurred on October 4-5. For flooded patches in the same circled area, the wetness levels on October 18 (Figure 3d) reduced to Medium and were no longer under flooding risk. Similar patterns were observed around smaller waterbodies along Mills Creek in the northeast.

Aside from these apparent water retreat patterns, the Wet and Highly Wet areas also showed interesting transition between the two dates. Overall wet clusters were fragmented in both maps. On October 8 they were mostly clustered in south of the city along the downstream of Congaree River. On October18, patches of wet areas became smaller but were more widespread over the city. It may indicate that, in a short period after heavy precipitation, flood risk could remain high in developed lands even though water in flooded areas began to retreat. If more precipitation occurs as the initial flood waters are receding, these areas appear at higher risks of potentially re-flooding.

Wetness dynamics over the 10 days between the satellite images are demonstrated in Figure 4. Areas remain in Wet, Highly Wet, and Water levels, and those with their levels increased (new) or decreased (retreat) are displayed in the figure. Inside the city boundary are mostly developed lands under various intensities of development (open lands, low, medium, and dense urban). As expected, large flooded patches were mostly

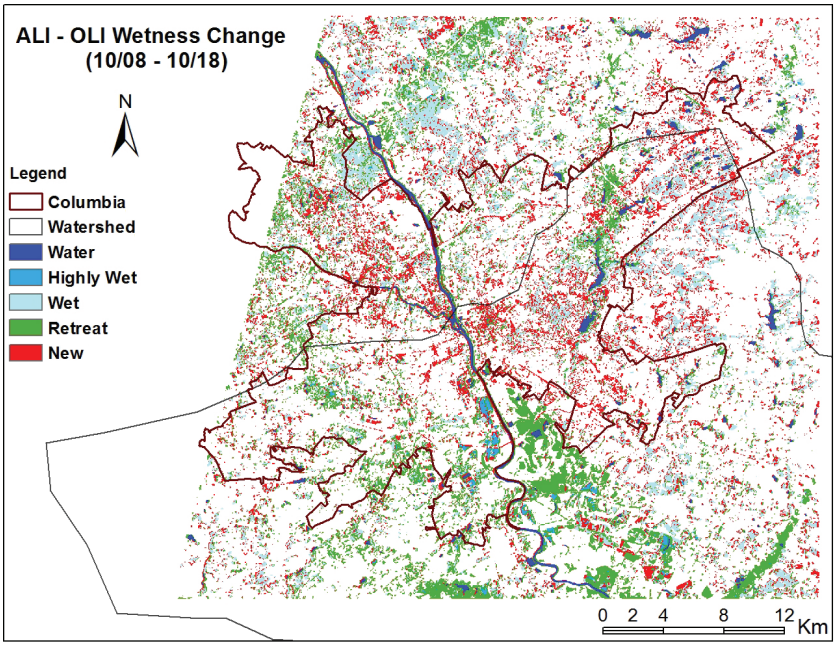


Figure 4. Wetness dynamics among the three wetness levels. The "Retreat" represents areas with their wetness levels decreased, while the "New" are those with their levels increased.

retreated by October 18. However, wetness in Columbia City was still a concern. Developed lands within the city boundary had increased wetness. In vegetated lands outside of the city, wetness dramatically decreased, indicating reduced risk on these pervious surfaces. Areal coverage of the three wetness levels between October 8 (ALI) and 18 (OLI) shifted from wetter classes to less wet classes; open water decreased from 40.07 to 33.33 km^2, the Highly Wet areas increased from 44.87 to 56.33 km^2, and the Wet areas slightly increased from 310.20 to 313.35 km^2. For risk assessment purpose, we suggest higher attention to these Highly Wet areas. Soil water content is approaching to its saturation point in these areas, which may result in flash flood within the short term or other damages in the long run.

Wetness evaluation against ground truthing

The two wetness level maps in Figure 3 are evaluated against the tweets with verified geolocations and the USGS Flash Flood points. Table 2 compares the wetness levels at the exact locations of tweets and Flash Flood points and the maximal NDWI values of the 150-m buffers at tweets.

The satellite-extracted wetness levels did not match with the posted flood at the exact locations of tweets. Only at two tweets occurred at points that were classified as Water on October 8 and one on October 18. At most tweet points, the images were classified as either Wet or non-wet (low to medium) levels. Note that most of these tweets were dated October 4 or 5. Although satellite images were acquired either 3-4 (ALI) or 13-14 (OLI) days later after the tweets, the maximal wetness in the buffers was still in higher levels (Wet, High Wet and Water). This suggests that, within a 150-m buffer, it is very likely that wetness remains high a few days after a flood event.

Spatial distributions of high wetness levels on October 8 were also overlaid on the tweets and Flash Flood points and the official USGS Inundation map (Figure 5). The USGS Inundation map provides the

Table 2. Satellite-extracted wetness at tweeted and Flash Flood points.

Ground Truth	Image	Wetness levels				
		At Tweeted locations				
		Note Wet	Wet	High Wet	Water	Total
Tweet points	ALI (10/08)	17	11	2	2	33
	OLI (10/18)	17	9	6	1	33
		In 150-m buffers at tweeted locations				
	ALI (10/08)	2	14	12	5	33
	OLI (10/18)	0	10	16	7	33
Flash Flood points		At Flash Flood reported locations				
	ALI (10/08)	9	12	0	2	23
	OLI (10/18)	10	8	3	2	23

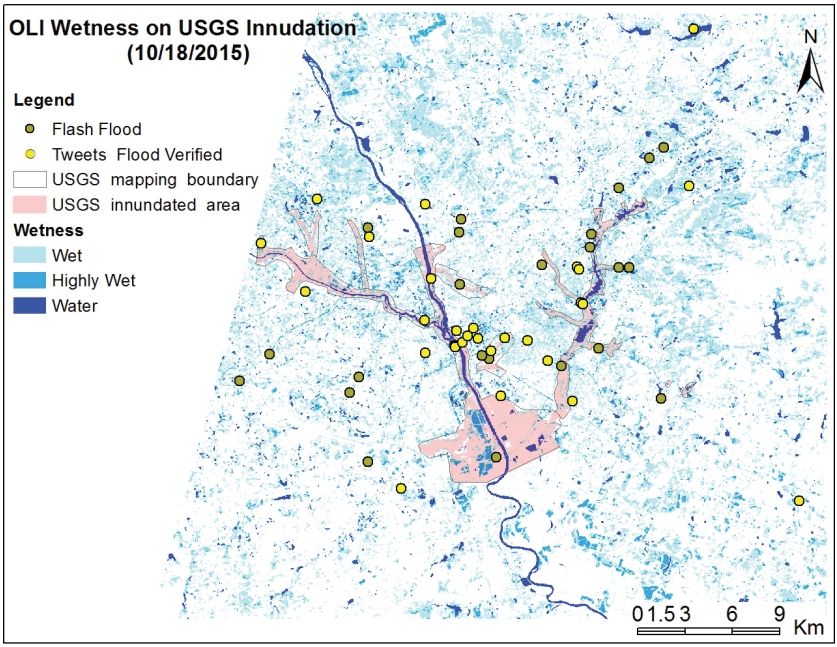


Figure 5. The three wetness levels on October 8 overlaid on three ground truthing sources.

flooded extents within its survey boundary. As shown in the figure, the USGS only surveyed in the predefined flood zone close to the major river channels. The inundation area agreed with some tweets and Flash Flood points, especially along the Gills Creek that connects waterbodies in the east branch of the survey zone, indicating that the floods were probably caused by water overflow in these areas.

Constrained by its survey zone, the USGS map could not reveal the complete inundation of the 2015 SC Flood in the study area. Most tweets and Flash Flood points were located outside of the inundation map survey area. Further away from waterbodies, these points may represent flash floods that occurred when soil surfaces reached saturation and could not hold additional moisture from excessive rainfall. When a flash flood quickly developed and retreated, areas around this flood may still hold high wetness levels that represented high risk of flooding. With a synoptic view, satellite imagery accurately picks up the high wetness levels at these points, which makes it possible to assess flash flood risks. Not all flash floods, however, could be effectively detected from post-event satellite imagery. Considering the limits on timing and size, satellite assessment of the localized flash floods is more difficult than water overflow.

Potential of integrating remote sensing and social media in extreme weathers

Extreme weather events such as floods have become more frequent in the southeastern US (Ingram et al. 2013). Due to

high cloud cover and rainfall, optical imagery from airborne or satellite platforms could not be acquired during the flood event, which makes the real-time imagery assessment impossible. This study extracts surface wetness patterns from the post-flood satellite imagery and evaluates their temporal dynamics during the 2015 SC Flood. The image acquired a few days after the flood easily picks up the remaining water in flooded areas. With more images acquired about 10 days later, the spatio-temporal changes of surface wetness will help to understand the pathways of water retreat and to identify the high flood-risk areas in case of continuous rainfall. Combined with the infrastructure network in Columbia, this information is of great value to quickly interpret flood severity, quantify flood damage, and assess the security along roadways, bridges and dams for emergency responders.

This study preliminarily explores the applicability of social media to assist satellite remote sensing in rapid flood assessment. The flood-related tweets were mostly posted in real time during flood events. While their geo-locations contain high spatial uncertainties, these tweets help us gain quick knowledge about the occurrences and developments of local flood events. When authoritative ground truth data are not available during a flood event, this study indicates that social media provides valid "truthing" source for evaluation of satellite-extracted flood information. To reduce the effects of spatial uncertainties, this study simply calculates the maximum NDWI in a 150-m buffer to represent the flood event at a specific tweet point. The assumption, however, could be questionable when the tweet points are close to natural waterbodies such as rivers and lakes. Flood development is highly related to topography. Past studies have developed various statistical weighting approaches to estimate flood inundation with elevation data (Apel et al. 2006; Li et al. 2017). In the future, we will test similar probability-based weighting approaches at the corresponding tweet points to better assessing flood risks from satellite-extracted wetness across the study area.

Challenges of Twitter data application also come from the verification of tweets related to a specific flood event. For the example in this study, among over two thousands of flood-related geotagged tweets posted after Hurricane Joaquin in City of Columbia, only 33 tweets were verified to show evidence of local floods. Manual selection of these tweets is time-consuming and subjective. Tweets not verified did not necessarily indicate that these locations were not flooded. Rather, the tweets did not show the evidence of flooding, e.g. without a flood photo attached or the photo not matching the searching keywords; therefore, we could not treat them as our "truthing" points in this study. Since social media data sets are big, highly unstructured and noisy in nature, sophisticated data mining algorithms are needed to automate the verification process from millions of tweets in a region. When automatic tweets verification approaches are available, social media could become important data sources for disaster assessment and management.

This study takes advantage of real-time tweets and large-coverage satellite imagery to compensate their drawbacks of locational uncertainty and the lack of spatial contiguity (for tweets) and delayed acquisition (for imagery) for rapid flood assessment. Integrated with social media,

the satellite-observed wetness maps help us better assess the flood severity in a timely manner, and thus quickly assist the resilience of society and environment responding to this extreme disaster. Nowadays, more guiding materials have become available to guide stakeholders or practitioners during disaster events (APFM 2017). VGI could thus be more involved in decision-making of disaster management by integrating authoritative reports with extended, crowdsourced databases. With improved crowdsourcing practices, the approaches tested in this study and future work could be applied for quick assessment of extreme events in a large geographic extent such as the southeastern US.

CONCLUSION

This study conducts satellite image analysis to evaluate the spatio-temporal wetness dynamics in Columbia, SC in the upper Congaree River Watershed after the 2015 Flood event. With a normalized difference wetness index, surface wetness levels are categorized from two satellite images acquired on October 8 and 18, 2015, in which high wetness (Wet, High Wet, Water) areas represent high risks for flood watch. The peak flood occurred on October 4 and 5 according to Twitter data. Within ten days, water retreated in flooded areas, but wetness within the City of Columbia remained high. These highly wet areas in the watershed deserve further attention for assessment of immediate flood risks and long-term impacts to the watershed. This study suggests that the post-event satellite imagery is useful for flood risk assessment because the surface wetness often remains high even after water retreats. The study also suggests that tweets could serve as a useful real-time source to leverage the post-event satellite assessment. It contributes to the literature of integrating crowdsourcing with remote sensing to support geospatial assessment and rapid response of extreme weather events.

ACKNOWLEDGMENTS

This study was supported by the 2015 SC Floods Research Initiative, University of South Carolina (USC). We thank the Hazards & Vulnerability Research Institute (HVRI) at USC for providing the verified flood-related tweets data for this study.

REFERENCES CITED

Associated Programme on Flood Management (APFM). 2017. Integrated Flood Management Tools Series: crisis mapping and crowdsourcing in flood management, Issue 26. Available at http://www.floodmanagement.info/publications/tools/APFM_Tool_26_e.pdf, last accessed on November 2, 2017.

Apel, H., Thieken, A. H., Merz, B., and Blöschl, G. 2006. A probabilistic modeling system for assessing flood risks. *Natural hazards,* 38(1):79–100.

Berg, R., 2016. Hurricane Joaquin, 28 September – 7 October, 2015. National Hurricane Center Tropical Cyclone Report (AL112015). Available at http://www.nhc.noaa.gov/data/tcr/AL112015_Joaquin.pdf. Last accessed on November 2, 2017.

de Albuquerque, J. P., Herfort, B., Brenning, A. and Zipf, A., 2015. A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management, *International Journal of Geographical Information Science*, 29(4), 667–689.

Blessing, R., Sebastian, A., S. M. ASEC, and Brody, S. D., 2017. Flood risk delineation

in the United States: how much loss are we capturing? *National Hazards Review*, 18(3):04017002, 1–10.

Feaster, T. D., Shelton, J. M., and Robbins, J. C., 2015. *Preliminary peak stage and streamflow data at selected USGS streamgaging stations for the South Carolina flood of October 2015* (No. 2015–1201). U.S. Geological Survey.

FLASH - Flooded Locations and Simulated Hydrographs, 2017. Improving the Science behind flash flood prediction. Available at https://blog.nssl.noaa.gov/flash/. Last accessed on August 10, 2017.

Fonte, C. C., Bastin, L., Foody, G., Kellenberger, T., Kerle, N., Mooney, P., Olteanu-Raimond, A. M., and See, L., 2015. VGI quality control. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-3/W5: 317–324.

Gao, B. C., 1996. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment,* 58(3):257–266.

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

Homer, C. G., Dewitz, J. A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N. D., Wickham, J. D., and Megown, K., 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing,* 81:345–354.

Imran, M., Elbassuoni, S. M., Castillo, C., Diaz, F., Meier, P., and Müller, T., 2013. Extracting information nuggets from disaster- related messages in social media. In T. Comes, F. Fiedrich, S. Fortier, J. Geldermann and T. Müller, (Eds.), *Proceedings of the 10th International ISCRAM Conference*, pp. 791–800. Karlsruhe, Germany: KIT.

Ingram, K., K. Dow, L. Carter, J. Anderson, eds, 2013. *Climate of the Southeast United States: Variability, Change, Impacts, and Vulnerability*. Washington DC: Island Press.

Li, Z., Wang, C., Emrich, C. T., and Guo, D., 2017. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods. *Cartography and Geographic Information Science,* DOI: 10.1080/15230406.2016.1271356.

Ji, L., Zhang, Z. and Wylie, B., 2009. Analysis of dynamic thresholds for the normalized difference water index. *Photogrammetric Engineering and Remote Sensing,* 75(11):1307–1317.

Musser, J. W., Watson, K. M., Painter, J. A. and Gotvald, A. J., 2016. *Flood-Inundation Maps of Selected Areas Affected by the Flood of October 2015 in Central and Coastal South Carolina*. Open-file Report 2016–2019, U.S. Dept. of the Interior, U.S. Geological Survey. DOI: 10.3133/ofr20161019.

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

Roberts, S. and Doyle, T., 2017. Understanding crowdsourcing and volunteer engagement: case studies for hurricanes, data processing, and floods. In *Flood Damage Survey and Assessment: New Insights from Research and Practice* (eds D. Molinari, S. Menoni and F. Ballio), John Wiley & Sons, Inc., Hoboken, NJ, USA.

Rosser, J. F., Leibovici, D. G., and Jackson, M. J., 2017. Rapid flood inundation mapping using social media, remote sensing and topographic data. *Natural Hazards*, 87:103–120.

Schnebele, E. and Cervone, G. 2013. Improving remote sensing flood assessment using volunteered geographic data. *Natural Hazards and Earth System Sciences,* 13(3), 669–677.

Schnebele, E. and Waters, N., 2014. Road assessment after flood events using non-authoritative data. *Natural Hazards and Earth System Sciences,* 14(4):1007–1015.

See L, Comber A, Salk C, Fritz S, van der Velde M, Perger C, Schill, C., McCallum I., Kraxner, F. and Obersteiner, M., 2013. Comparing the quality of crowdsourced data contributed by expert and mon-experts. *PLoS ONE*, 8(7): e69958.

Smith, L., Liang, Q., James, P., and Lin, W., 2017. Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. *Journal of Flood Risk Management,* 10(3): 370–380.

Triglav-Čekada, M. and Radovan, D., 2013. Using volunteered geographical information to map the November 2012 floods in Slovenia. *Natural Hazards and Earth System Sciences*, 13(11): 2753–2762.

Wang, C., Lu, Z. and Haithcoat, T. L., 2007. Using Landsat images to detecting forest dynamics responding to oak dieback in the Mark Twain National Forest, Missouri. *Forest Ecology and Management,* 240(1–3):70–78.

DR. CUIZHEN WANG (CWANG@MAILBOX.SC.EDU) is a professor in Department of Geography at the University of South Carolina, Columbia, South Carolina 29208. Her research interests include bio-environmental remote sensing and spatial analysis. Example applications include vegetation mapping, soil/water quality assessment, and coastal environment monitoring.

DR. ZHENLONG LI (ZHENLONG@MAILBOX.SC.EDU) is an assistant professor in Department of Geography at the University of South Carolina, Columbia, South Carolina 29208. His research focuses on spatial high-performance/cloud computing, big data processing/mining, and geospatial cyberinfrastructure within the area of data and computational intensive GISciences.

XIAO HUANG (XHI@EMAIL.SC.EDU) is a second year PhD student in Department of Geography at the University of South Carolina, Columbia, South Carolina 29208. His research is on geospatial development on rapid flood mapping and risk assessment of a disaster event.