Reconstructing Flood Inundation Probability by Enhancing Near Real-Time Imagery With Real-Time Gauges and Tweets

Article *in* IEEE Transactions on Geoscience and Remote Sensing · May 2018

DOI: 10.1109/TGRS.2018.2835306

CITATIONS
2

READS
252

3 authors:

Xiao Huang
University of South Carolina
6 PUBLICATIONS **7** CITATIONS
SEE PROFILE

Cuizhen Wang
University of South Carolina
113 PUBLICATIONS **1,577** CITATIONS
SEE PROFILE

Zhenlong Li
University of South Carolina
62 PUBLICATIONS **654** CITATIONS
SEE PROFILE

Reconstructing Flood Inundation Probability by Enhancing Near Real-Time Imagery With Real-Time Gauges and Tweets

Xiao Huang, *Student Member, IEEE*, Cuizhen Wang, and Zhenlong Li

***Abstract*— Flood inundation probability is critical for situation awareness, flood mitigation, emergency response, and postevent damage assessment. Current flood inundation mapping approaches can be categorized into real-time (RT) and near-RT (NRT) processes based on the timing of data acquisition. However, the intrinsic limitations of each category largely hamper their applications for flood mapping. Taking the 2015 South Carolina flood in downtown Columbia as a case study, this paper proposes a flood inundation reconstruction model by enhancing the NRT normalized difference water index (NDWI) derived from remote sensing imagery with the RT data including stream gauge readings and social media (tweets). Splitting into three modules: water height module, global enhancement module, and local enhancement module, the proposed model first incorporates the gauge readings and the NDWI image to reconstruct a macroscale flood probability layer, which is then locally enhanced using the verified flood-related tweets. The final output of the model matches well with the U.S. Geological Survey inundation map and its surveyed high-water marks. Results suggest that by enhancing NRT imagery with RT data sources, the proposed flood inundation probability reconstruction model renders a more robust, spatially enhanced flood probability index for emergency responders to quickly identify areas in need of urgent attention.**

***Index Terms*— Geographic information science, modeling, remote sensing.**

I. Introduction

FLOODS are among the most powerful forces on earth, posing devastating threats for all population in the world. Despite the advances in flood studies and enforcement of hazard reduction policies, the flood-induced damages follow an increasing trend [1]. In the United States, extreme rainfalls and floods have accounted for an annual loss of 82 lives and an economic loss averaging U.S. $7.96 billion each year in the past 30 years [2]. In 2005 alone, direct flood damage claimed 43 lives and a total of U.S. $55.33 billion [2]. As the severity and frequency of flood events have noticeably increased, there is a growing need of rapid responses from local authorities to reduce flood-related fatalities and economic losses [3]. Flood inundation maps are the prerequisite information that is critical for flood mitigation, emergency response, and postevent damage assessment.

Intensive studies on flood innudation mapping have been conducted and could be generally split into the real-time (RT) and near-RT (NRT) mapping based on data acquisition time: the RT mapping utilizes RT data from water gauge sensors, timely official flooding reports, or crowdsourcing and the NRT mapping utilizes data shortly after the flooding event, usually with a lag of several days. The postevent flood mapping, i.e., utilizing data long after the flooding event for long-term damage assessment after a severe flood, is not the focus of this paper.

The RT mapping techniques produce a relatively coarse but timely flood assessment that is crucial for rapid flood mitigation and response. Water height readings collected from the stream gauges are one of the most commonly used RT data [4]. The U.S. Geological Survey (USGS), for example, provides water height information at each stream gauge at 15-min intervals [5]. However, limitations do exist because those gauges are dispersedly distributed, resulting in insufficient information acquisition and, when water levels rise beyond the measuring limits, saturation of water height records [6]. Volunteered geographical information (VGI) provides an alternative source about a flood at the exact time of its occurrence [7], [8]. Defined by Goodchild [9], VGI describes the concept of citizens as sensors, allowing rich voluntary information to be provided in the form of text, images, and videos to aid geospatial and temporal analysis. With increasingly sound crowdsourcing platforms being built, there is an up-surging interest of utilizing VGI for better flooding situation awareness [10]–[13]. However, compared to authoritative information, the largely untrained VGI providers and its built-in assertive characteristic strongly restrict its functional utility [14]–[16].

The NRT flood mapping deals with data within a few days of delay, for example, remotely sensed (RS) imagery when weather conditions become capable. Compared to the RT data, RS images render a synoptic view in a large geographic area, significantly contributes to a holistic understanding of flooding situation. The RS imagery has been widely used to monitor flooding extent and to assess its damages [17]–[20]. Among all, flood-related indicators derived from RS images, the Normalized Difference Water Index (NDWI) has been extensively used as an indicator of land surface wetness [21].

Manuscript received January 28, 2018; revised April 3, 2018 and May 2, 2018; accepted May 4, 2018. This work was supported by the Office of the Vice Provost Research at the University of South Carolina under the ASPIRE-I Grant 13540-18-47767. *(Corresponding author: Xiao Huang.)*

The authors are with the Department of Geography, University of South Carolina, Columbia, SC 29208 USA (e-mail: xh1@email.sc.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2018.2835306

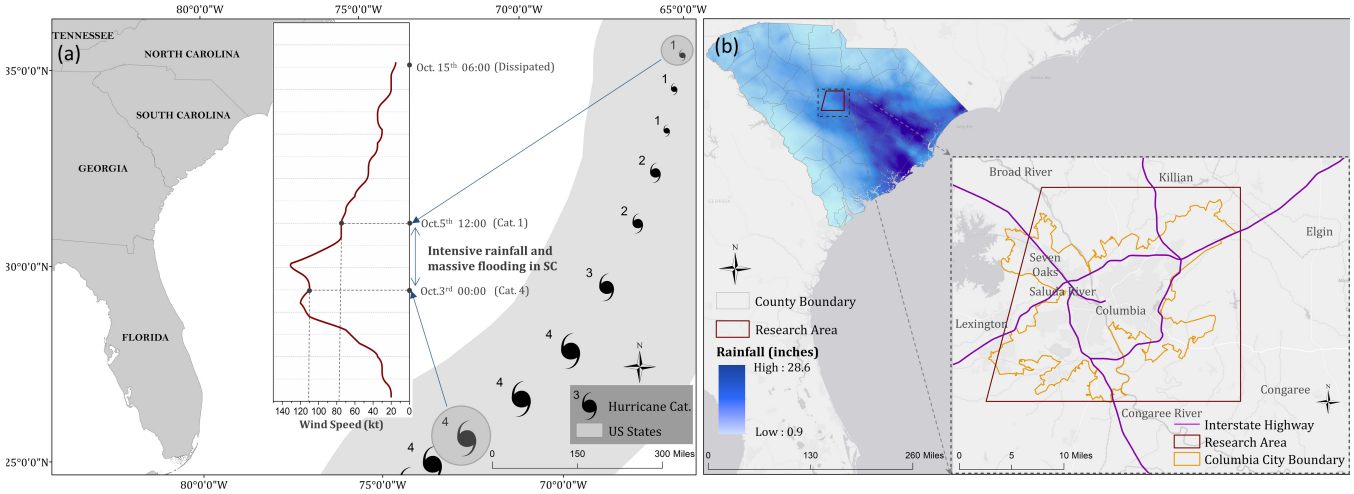

Fig. 1. (a) Hurricane Joaquin (NOAA NHC, 2016). (b) Rainfall status in SC (FEMA MOTF, 2016) and the research area of this paper.

A broad range of NDWI applications for flood-related investigations have been conducted, including water coverage delineating [22], flood hazard mapping [23], and flood-prone area identification [24]. The NDWI provides rich wetness information that is more valuable in flooding analysis than do traditional water delineation approaches. However, the limited RS data availability due to long revisit cycle and extreme weather conditions heavily prevents RS sensors from acquiring timely images [25]. For images acquired after a flood event, the information in the temporal gap between flooding peaks and image acquisitions is lost [26].

These inherent limitations from both RT and NRT sources can be reciprocally compensated. The spatially isolated RT data can be used to enhance the delayed image observations, and the RS imagery provides a spatially continuous view beyond the isolated stream gauge records and the uncertain tweeter posts. The combined information leads to the reconstruction of flood extent and flood risks during an event. Guided by this idea, attempts have been made to integrate the NRT images with RT sources. In our previous paper, we found that the satellite-extracted surface wetness serves as a great weighting factor for flood-related tweets [27]. Wang *et al.* [18] found that the integration of Landsat TM images and river gauge readings overcame the flooding underestimation issue in vegetation canopies. Schnebele and Cervone [28] improved the RS flood assessment by combining satellite imagery with high temporal-resolution ground data. More fusion of NRT and RT data can be found in a number of flood studies [11], [29], [30].

This paper builds a probability reconstruction model of flood inundation by fusing the RT stream gauge data and social media (tweets) to enhance the near-RT NDWI layer derived from RS imagery. The flood probability is reconstructed by offsetting the information loss in the time gap while preserving the distributions of land surface wetness. The model output is validated via the USGS survey points and the official inundation map released four months after the 2015 SC flood event [31]. The term "flooding probability," or FP later used in this paper, refers to the probability of a certain area being inundated during a flood event.

II. Research Area and Data Sets

A. Hurricane Joaquin and the 2015 Sc Flood in Columbia, Sc, Usa

Hurricane Joaquin in 2015 is the strongest Atlantic hurricane of nontropical origin in the satellite era [32]. Along its development path [Fig. 1(a)], it contributed to a record rainfall in SC during October 1–October 6. The widespread record-breaking rainfall (more than 20 in in some areas) caused catastrophic floods from the central SC to the coast, resulting in 19 fatalities and U.S. $1.5 billion damage losses in the state [33]. The capital city of SC, Columbia, experienced one of the most widespread and devastating floods in this event [34]. This densely populated area covers Richland county and Lexington County in the central SC [Fig. 1(b)]. Congaree River, joined by Broad River and Saluda River in the north, is the major flowing waterbody across its metropolitan area. During the fifth-day flood, Columbia was significantly impacted in all aspects. In this paper, we chose the City of Columbia as our research area.

B. Data Sets and Preprocessing

The data sets used in this paper are broken into four categories based on the data acquisition period, which include RT data, NRT data, postevent data, and accessory data.

1) Real-Time Data Set:

a) Water height readings at five stream gauges: The five gauges in the research area are numbered u02162093, u02169000, u02169500, u02169506, and u02169570 [Fig. 2(a)]. The USGS provides water height data for these gauges at a 15-min interval [35]. The maximum and reference water height readings at each gauge were extracted (Table I). It should be noted that the maximum water height readings do not necessarily represent the maximum water level in streams. Sensors in some gauges failed to record water height due to a variety of reasons [36], leading to the missing records for a certain period of time [as shown in Fig. 2(b)]. Given the incomplete height readings, the maximum water height readings in Table I denote the highest readings available for each gauge. The reference water height was

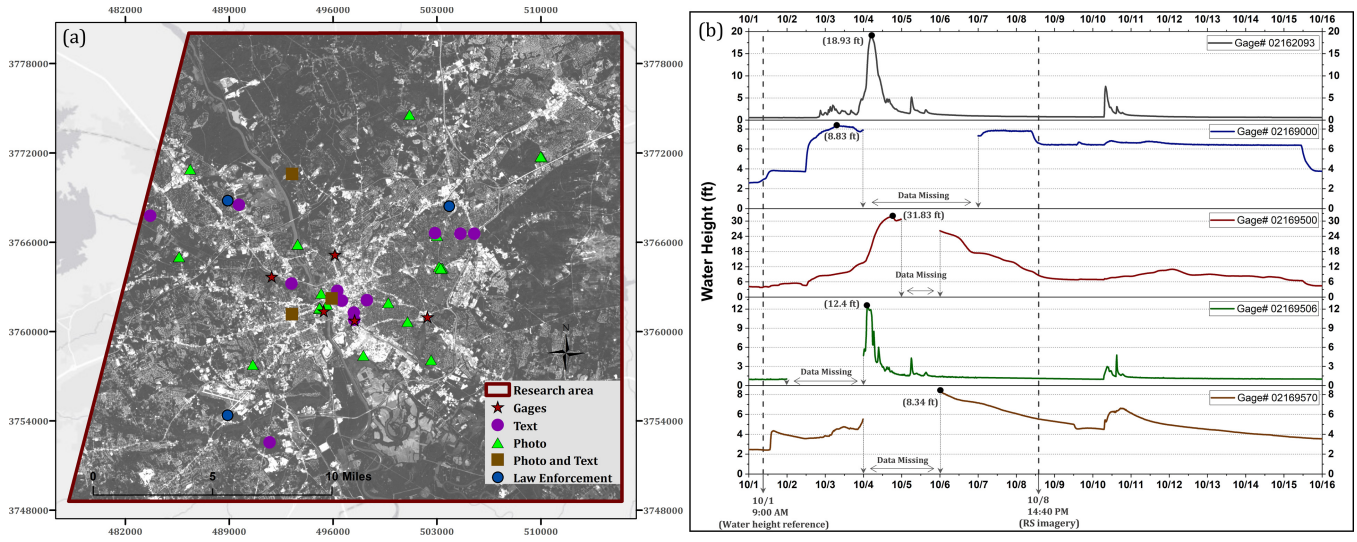

Fig. 2. (a) ALI image (panchromatic) in the research area and locations of the verified tweets and stream gauges used in this paper. (b) Water height readings at the five stream gauges on October 1–October 16. The stream gauge data were obtained from the USGS National Water Information System (https://waterdata.usgs.gov/sc/nwis/rt).

TABLE I
DETAILED STATISTICS OF FIVE GAUGES

Gauge Number	Gauge base height (ft)	Datum shift (ft)	Maximum water height reading (ft)	Maximum water height Reached time	Reference water height reading (ft)[a]
# 02162093	199.10	-0.794	18.93	Oct 4th 5:07:00 AM	0.56
# 02169000	149.46	-0.787	8.33	Oct 3rd 7:45:00 AM	2.91
# 02169500	113.02	-0.787	31.83	Oct 4th 5:52:00 PM	3.91
# 02169506	165.55	-0.781	12.40	Oct 4th 2:22:00 AM	0.98
# 02169570	137.38	-0.778	8.34	Oct 6th 0:00:00 AM	3.17

Note. The datum for gauge base height is NGVD 29 while the datum for DEM used in this study is NAVD 88. A conversion tool in VERTCON (*https://beta.ngs.noaa.gov/cgi-bin/VERTCON/vert_con.prl*) is used to convert them to a uniform datum (NAVD 88).
[a] Reference water height readings from all five gauges were obtained at Oct 1st 9:00 AM

assumed the water height reading at 9:00 AM, October 1 (a stable stage before the flood). Both the maximum and reference water height readings are later translated to the elevation based on their base height. (A datum conversion process is involved.)

b) Verified tweets: The tweet pool used in this paper has been generated using twitter stream air position indicator (API) and REST API in [6]. In these efforts, all geotagged tweets within the research area between October 3 and October 6 were downloaded, and then manually checked to make sure that their information was coordinate relevant and flood relevant. After the spatial restraint and keywords restraint, a total of 49 flood-related tweets with content covering text, photo, or both were selected within the research area [Fig. 2(a)]. Their contents are flood related and matched well with their intrinsic longitude and latitude. Tweets labeled "law enforcement" are official flash flood warnings issued by local authorities.

2) Near Real-Time Satellite Image: The NRT satellite image was acquired from the EO-1 Advanced Land Imager (ALI) on October 8, 14:40 PM, the earliest cloud-free multispectral image available in this flood to our best knowledge. The ALI image has 30-m resolution in multispectral bands and 10-m resolution in panchromatic band [Fig. 2(a)]. Given the slight haze existence and atmosphere interference, an atmospheric correction together with haze removal function was applied in the ATCOR2 module of ERDAS/IMAGINE. The NDWI is then calculated from the green- and short-wave infrared (SWIR) bands: NDWI $= (\rho_{\text{green}} - \rho_{\text{SWIR}})/(\rho_{\text{green}} + \rho_{\text{SWIR}})$ [37].

3) Postevent Data: Set Usgs Survey Points and Inundation Map: After the 2015 SC flood, the USGS conducted field

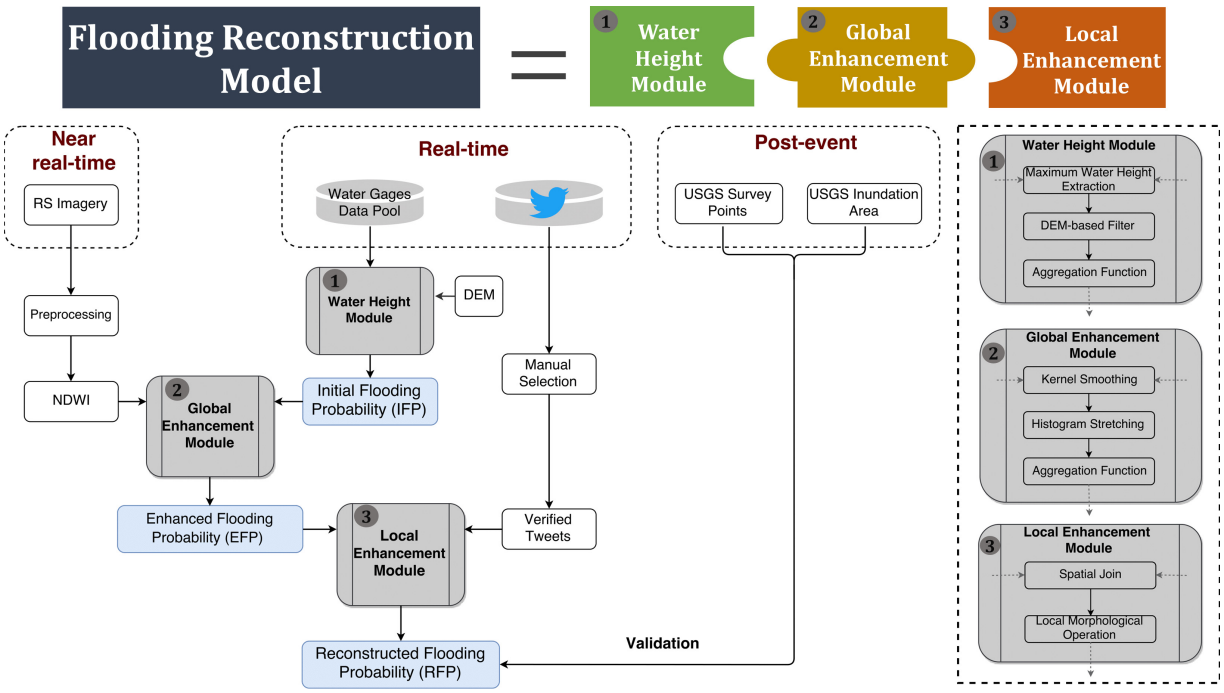

Fig. 3. Methodology overview. The flood reconstruction model is composed of three modules: water height module, global enhancement module, and local enhancement module.

TABLE II
THREE DEM-RECLASSIFIED CATEGORIES IN THE RESEARCH AREA BASED ON THE REFERENCE WATER HEIGHT AND MAXIMUM WATER HEIGHT AT EACH GAUGE

Gauge Number	DEM range (ft)		
	Water Body	Flooded areas	Non-flooded areas
# 02162093	<198.866	[198.866, 217.236)	≥217.236
# 02169000	<151.583	[151.583, 157.003)	≥ 157.003
# 02169500	<116.143	[116.143, 144.063)	≥ 144.063
# 02169506	<165.749	[165.749, 177.169)	≥ 177.169
# 02169570	<139.772	[139.772, 144.942)	≥ 144.942

surveys to collect water height marks to aid in documenting the high-water events [38]. Downloaded through USGS short-term network portal (https://stn.wim.usgs.gov/STNDataPortal/#). The survey point data set contains 574 high water marks (HWMs) in SC and 337 within our research area. To ensure data quality, we only selected the HWMs with quality remarks "good" and "excellent." A total of 277 good-quality HWMs were extracted within the research area. The inundation map in the research area was acquired from the USGS Flood Inundation Mapping Program (https://water.usgs.gov/osw/flood_inundation). It should be noted that USGS only surveyed the area within the flood zone of the research area. This binary flood extent and field-surveyed HWMs are the only official postevent data available at the time of writing.

4) Accessory Data Set: The digital elevation model (DEM) elevation data at 3-m resolution was obtained from the SC Department Natural Resources (http://www.dnr.sc.gov/GIS/lidar.html). The high-resolution Google Earth images acquired on October 7 was later available in the research area. Due to their small coverage, they were only visually compared with our modeled results for verification purposes. Other spatial data sets (Shapefiles) including city, county, and state boundaries were retrieved from local authorities.

III. METHODOLOGY

The proposed flood reconstruction model consists of three modules:

1) generating an initial FP (IFP) merely based on water height readings at five gauges and DEM (water height module);
2) generating an enhanced FP (EFP) by globally enhancing the satellite-extracted NDWI surface with the IFP via kernel smoothing, standardization and aggregation (global enhancement module);
3) generating reconstructed FP (RFP) by locally enhancing the EFP with the verified tweet points via a morphological operation (local enhancement module).

Detailed flowchart is outlined in Fig. 3.

A. Water Height Module

This module integrates the DEM and water height points to generate an IFP layer by building an initial flood water surface. The maximum readings of water height at a stream gauge represent this point's RT water height during the flood event. Data from all five gauges are used in the model to compensate for the uncertainty from local topological unevenness.

With the readings at a given gauge, the elevation at each pixel of the research area is compared against the reference water height and maximum water height (listed in Table I), and is classified as one of the three categories: water body, flooded, and nonflooded. If the elevation of a pixel is below the reference water height, it is more likely natural water body. If its evaluation is above the maximum height, it is not likely to be flooded (nonflooded). For a pixel with its elevation in between, we assume that it is flooded. Different elevation thresholds are applied when compared with readings at different gauges (Table II). Different weights are given to the three categories to approximate their proneness to flood. Natural water bodies are assigned a weight of 2. The nonflooded areas have a weight of 0 because areas higher than the maximum water height are not likely to be flooded. Areas in between are assigned a weight of 1.

Five weight layers (W_Layer) are extracted since we have five gauges in the research area. An aggregation function was applied to extract the IFP layer

$$\text{IFP} = \sum_{i=1}^{n} W\_\text{Layer}_i \tag{1}$$

where $n = 5$ denoting five separate layers from five gauges within the research area. The resulted IFP has a weight range of [0, 10].

B. Global Enhancement Module

This module incorporates the land surface wetness and topographic characteristics by integrating the NDWI surface with the IFP from the Water Height Module. The integration aids in: 1) compensating the spatial contiguity of wetness that is not captured from gauge readings and 2) providing additional flooding awareness for areas with high elevation where the DEM-based IFP fails to cover. Even the NDWI layer represents the wetness conditions a few days after the flood event, areas remaining high wetness indicates that they are prone to flooding during the event.

The global enhancement module begins with a quartic kernel smoothing function, which places a moving 2-D kernel over the data layer to estimate the density at the kernel center. In this module, a kernel smoothing function is applied to both NDWI and IFP layers. The kernel-smoothed IFP layer, namely, IFPs is mathematically defined as

$$\text{IFPs}_{(x,y)} = \frac{1}{nh^2} \sum_{i=0}^{n} K\left(\frac{\text{IFP}_x - x_i}{h}\right) K\left(\frac{\text{IFP}_y - y_i}{h}\right) \tag{2}$$

where h and n denote the bandwidth and sample size of the kernel, respectively. The $\text{IFPs}_{(x,y)}$ represents the density at location (x, y). Bandwidth h of kernel K is initially set to 1000 m for calculational convenience. The kernel function K used in this paper is the quartic kernel, which is calculated as

$$K(u) = \frac{15}{16}(1 - u^2)^2 \tag{3}$$

where variable u has to meet a cutoff requirement: $|u| \leq 1$.

The ALI-derived NDWI is smoothed using the same kernel function and parameter settings. The result is named NDWI_s. Both NDWI_s and IFP_s are then normalized to the [0, 1] using a maximum–minimum normalization, namely, IFS_n and NDWI_n, respectively. The normalization process makes these layers mathematically comparable. We notice that the IFS_n is distributed consistently and distinguishably in its value range. However, the NDWI_n values are mostly clustered in its middle range (around 0.5). To better spread out these most frequent values, a modified logistic stretch function is applied to NDWI_n. The contrast enhancement is defined as follows:

$$S\_\text{NDWI}_n = \frac{L}{1 + e^{-k(\text{NDWI}_n + a)}} \tag{4}$$

where S_NDWI_n represents the stretched NDWI_n, k is the coefficient that controls the steepness of the curve, a is the x-value of midpoint, and L measures the maximum value of the curve. L, a, and k are set as 1, 0.5, and 10, respectively.

Finally, an EFP is produced by aggregating the IFS_n and S_NDWI_n as follows:

$$\text{EFP} = \text{IFS}_n + S\_\text{NDWI}_n. \tag{5}$$

C. Local Enhancement Module

The EFP layer could be locally enhanced around the verified flood-related tweets locations, which are the local areas that were actually flooded during the 2015 SC flood event. According to Tobler's law, areas closer to a verified tweet point are more likely to be flooded. Therefore, pixels surrounding the tweet point could be locally enhanced for their FP, and those closer to the tweet point receive stronger enhancement.

It is reasonable to assume that this local enhancement follows a morphological dilation pattern. Assigning a domain D_s with a search radius centered at a tweet point, at a pixel (x, y), the dilation $g(x, y)$ is the maximal value within the research area adjusted by the distance-decay effect

$$\begin{aligned} &g(x, y) \\ &= (\text{EFP} \oplus e)(x, y) \\ &= \max\{\text{EFP}(x-a, y-b) | (x-a, y-b) \in D_s; (a, b) \in D_e\} \end{aligned} \tag{6}$$

where D_e denotes the domain of a structuring element e, which is set to be square shaped with a length of 10 pixels (300 m) in this paper. The final probability $\text{RFP}(x, y)$ is then calculated by introducing a distance-related coefficient c to the dilation

$$\text{RFP}(x, y) = \frac{g(x, y) - \text{EFP}(x, y)}{c} + \text{EFP}(x, y) \tag{7}$$

where c is defined as $c = (r/r - d)$, in which r denotes the radius of the search area and d denotes the distance from (x, y) to the tweet point. The search radius is set to 1000 m in this paper.

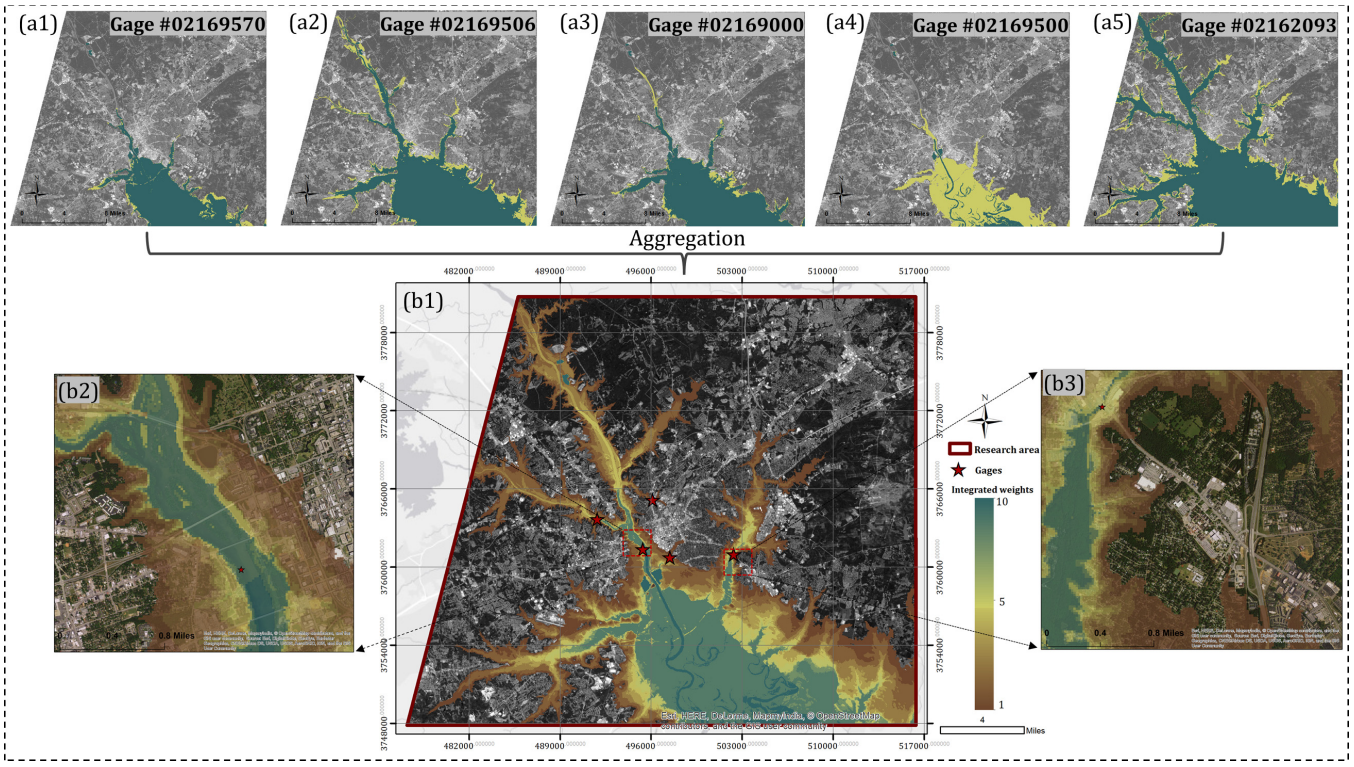

Fig. 4. IFP derived from DEM and gauges. The background is the black-and-white display of the ALI image.

The RFP layer is the final RFP of the research area after three consecutive steps: initial construction, global enhancement, and local enhancement.

IV. Results and Discussion

A. Initial Flooding Probability (Ifp) Map

The integration of DEM and water height readings at five gauges generated the IFP for the entire research area. Each flood extent weight layer provides its own estimation based on a single gauge reading, rendering a local FP prediction. Fig. 4(a1)–(a5) demonstrates five weight layers extracted from each gauge. The estimated flood extents are different from each other due to the unevenness of local terrain at the five gauges. Exceptionally, the weight layer based on Gauge 02169500 [Fig. 4(a4)] varies significantly from others.

The integration of all weighted layers greatly reduces the uncertainties from local DEM variations and provides an overall representation of FP in the research area. The integrated IFP layer in Fig. 4(b1) has integer values ranging from 0 to 10. An IFP value of 0 represents the areas at elevations higher than the maximum water height of all gauges, and therefore, is not possibly flooded. An IFP value of 10 represents the areas lower than the maximum water height of all gauges, and therefore, has the highest potential of being flooded. Areas with IFP $= 0$ are left transparent to show the background image in Fig. 4(b1). Fig. 4(b1) reveals a massive flood occurrence in the south of the research area (IFP value $= 10$) and high fps along the Congaree River [Fig. 4(b2)] and Gills Creek [Fig. 4(b3)] in downtown Columbia.

B. Globally Enhanced Flooding Probability (Efp) Map

The EFP enhances the IFP layer by fusing in the spatially continuous wetness information derived from remote sensing imagery. Fig. 5 shows the comparison of the distributions of fps before and after the NDWI-implemented global enhancement.

In Fig. 5(a), the normalized IFP (IFP_n) successfully identifies the low-elevation flooded areas along the major stream channels. The extensive area of high FP in the south of the research area was actually flooded due to its low elevation during the event. The flooded areas in the south end of Fig. 5(a) are not well identified in Fig. 5(b) due primarily to the time lag of image acquisition. The image was taken three days after the flooding peak, and these local floods have been retreated. Therefore, the IFP_n has a distinctive RT advantage over the satellite-extracted S_NDWI_n.

On the other hand, with the spatially continuous NDWI, Fig. 5(b) reveals the hot spots of high wetness in high-elevation zones (highlighted by circles). Areas in high elevations are also likely to be flood prone due to its local topographic unevenness and land use types. In Fig. 5(b), these local areas are fairly identified by their high land surface wetness. Therefore, the S_NDWI_n renders the holistic wetness information, largely contributing to identifying areas at higher elevations where IFP_n fails to cover.

The integrated EFP obviously provides a better estimation by taking advantage of both aspects [Fig. 5(c)]. Spatially, it not only keeps high FP in lower elevations and alongside river channels demonstrated by the IFP_n but also highlights the potentially flooded regions in higher elevations as suggested

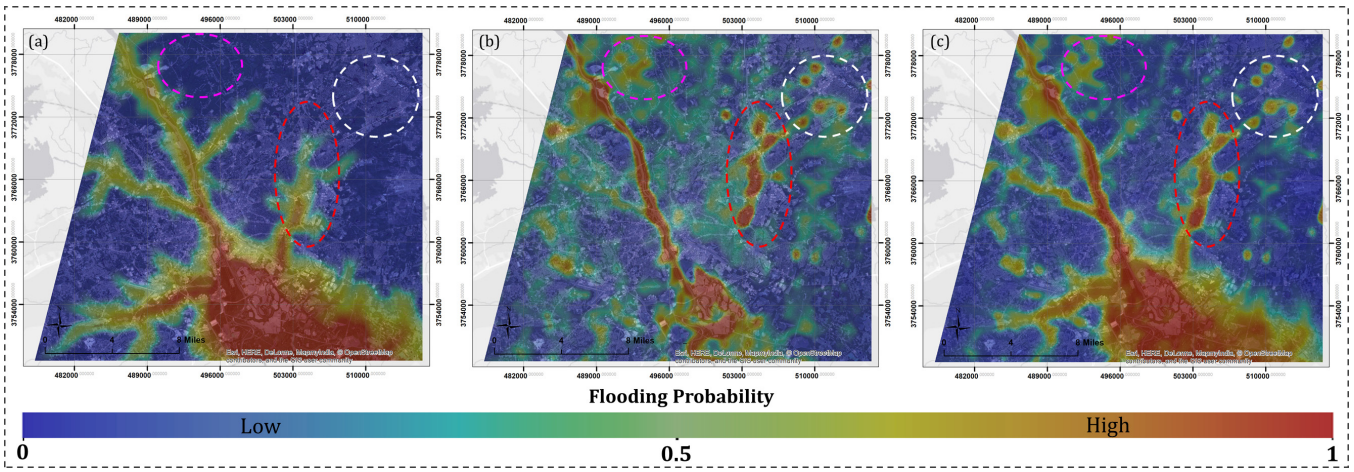

Fig. 5. Comparison of the flood probability before and after integrating the NDWI. (a) Kernel-smoothed and normalized IFP ($\mathbf{IFP_n}$). (b) Kernel-smoothed, normalized, and stretched NDWI ($\mathbf{S\_NDWI_n}$). (c) Integrated EFP.

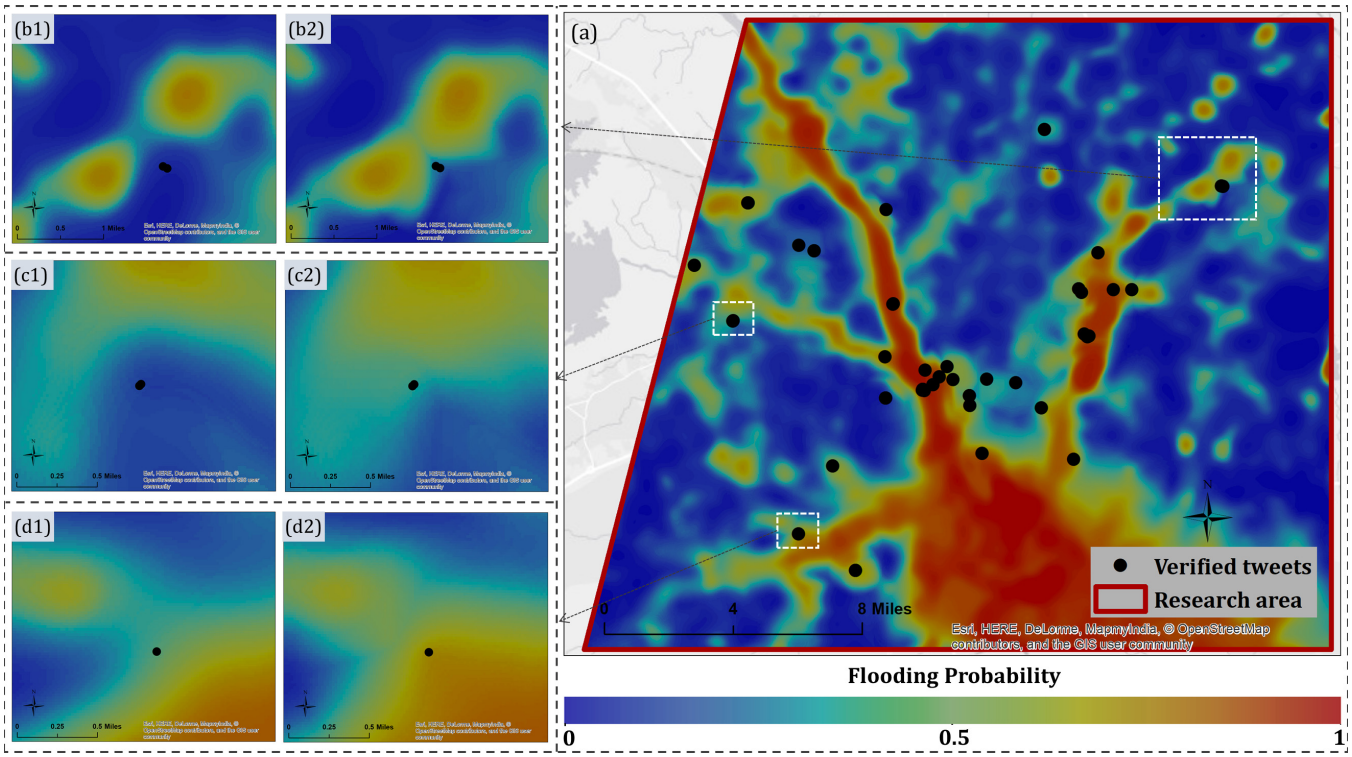

Fig. 6. Final RFP and subset comparisons with the EFP. (a) Final RFP layer for the entire research area. (b2), (c2), and (d2) Results after local enhancement of (b1), (c1), and (d1), respectively.

by the high wetness. Temporally, the EFP compensates the time lag between image acquisition and gauge readings. Fig. 5 demonstrates that by considering DEM/gauges and wetness together, a comprehensive, global FP estimation is achieved.

C. Reconstructed Flooding Probability (Rfp) Map

The RFP is a result of local enhancement from the EFP via a morphological dilation process. It utilizes the spatial and temporal advantages provided by the verified flood-related tweets (Fig. 6). The *in situ*, RT occurrences of floods at these tweets help to enhance the local flood surfaces. Comparison between the EFP and RFP is demonstrated in several subsets (marked in Fig. 6). For subsets Fig. 6(b)–(d), fps around the tweets (black dots) in the RFP maps Fig. 6(b2), (c2), and (d2) are much higher than those in the EFP maps Fig. 6(b1), (c1), and (d1), respectively. These modifications indicate that supplemental RT data source like tweets can greatly aid in identifying local floods ignored in the EFP.

In Fig. 7, the before and after the tweet-implemented flood probabilities (EFP versus RFP) are visually compared with the high-resolution Google Earth images [Fig. 7(a1) and (b1)]

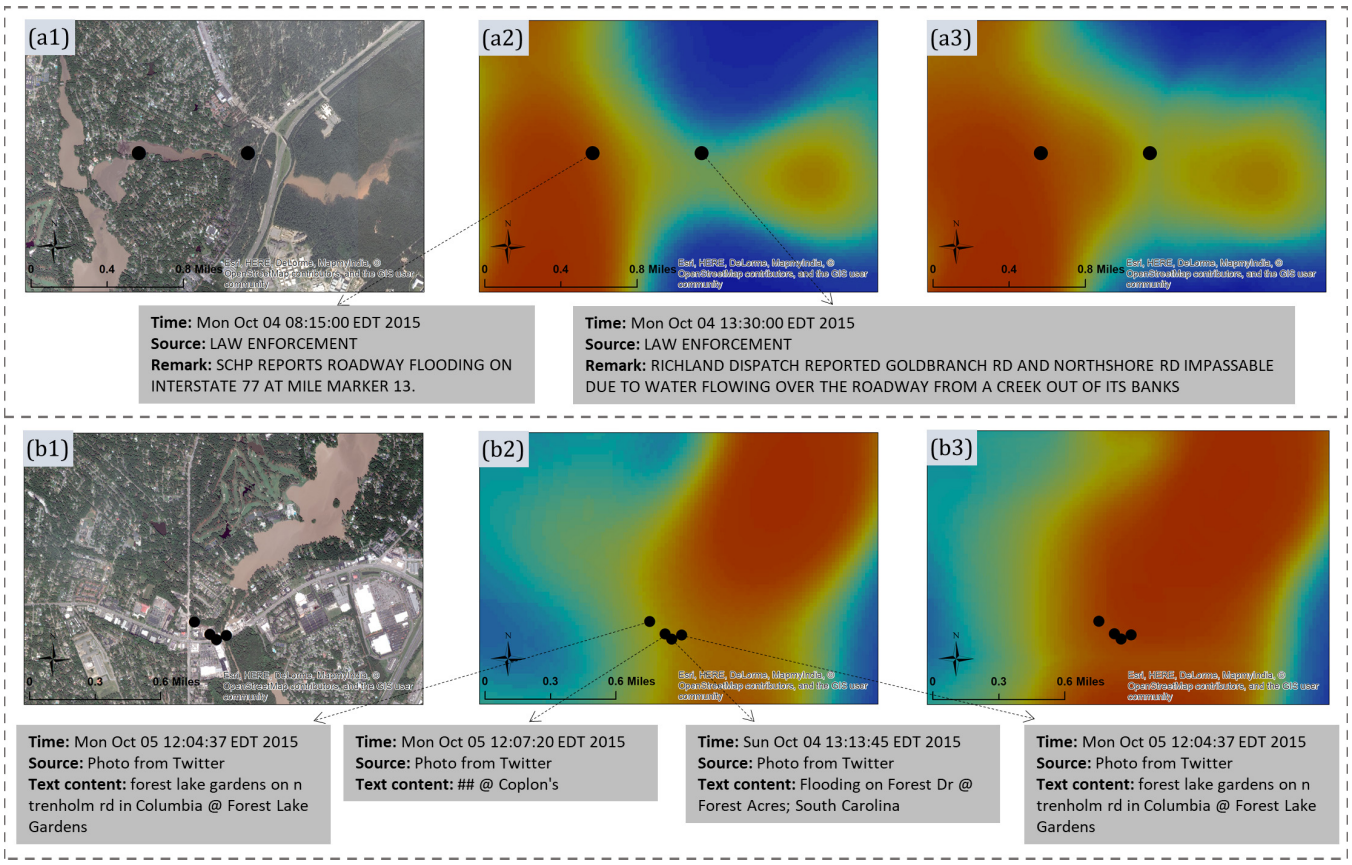

Fig. 7. Tweet examples and visual comparison of EFP and RFP with high-resolution Google Earth images acquired on October 7.

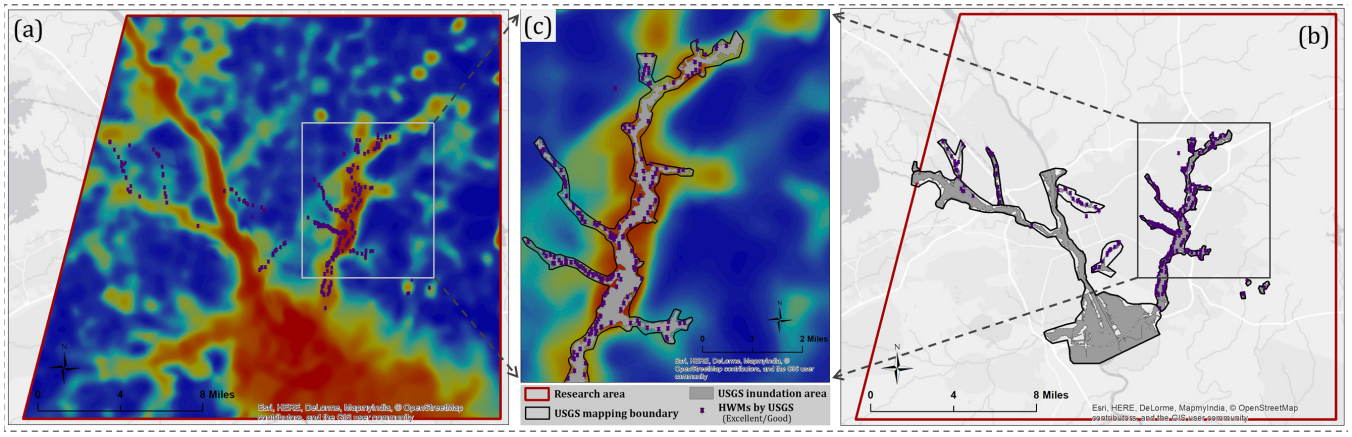

Fig. 8. RFP compared with USGS HWMs and USGS Inundation map. (a) RFP. (b) USGS inundation boundary and inundation area. (c) (a) and (b) overlaid in Grill Creek region.

acquired on October 7, two days after the flooding. Fig. 7(a2) and (b2) shows the EFP in the imaged subsets, and Fig. 7(a3) and (b3) shows the locally enhanced RFP from Fig. 7(a2) and (b2), respectively, using those tweets (black dots) in each subset.

The Google Earth images in Fig. 7(a1) and (b1) show two flooded subsets that are also identified in the EFP, as shown in Fig. 7(a2) and (b2), respectively. However, some flooded areas indicated by the verified tweets are not well spotted on either Google Earth images or the EFP, for instance, the four tweets marked in Fig. 7(b1). This is partly due to the time discrepancy between the tweets and images as those high-resolution images were taken on October 7, while the flood reached to its peak around October 5. Compared with the EFP, the RFP [Fig. 7(a3) and (b3)] boosts the FPs surrounding a certain tweet, resulting in a reasonable probability distribution. Examples in Fig. 7 illustrate that even a small amount of RT supplemental VGI from social media can significantly adjust the FP in a local scale, thus improving the reconstruction the flood probability during a flood event.

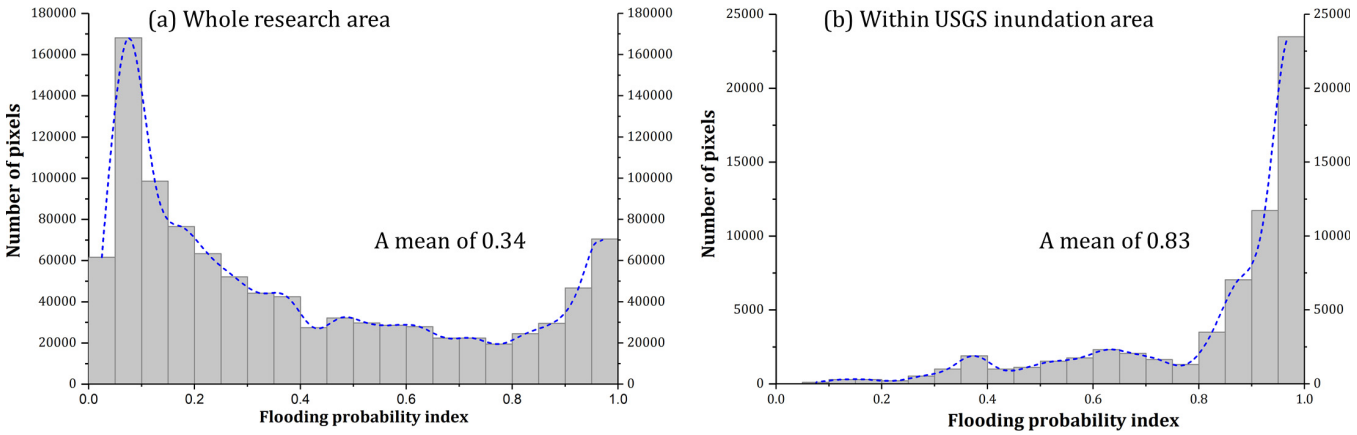

Fig. 9. EFP in (a) whole research area and (b) area within USGS inundation area.

TABLE III
RFP AND HWMS WITHIN THE USGS INUNDATION AREA

RFP	Number of pixels (%)	Number of HWMs (%)
0 ~ 0.2	1153 (01.84%)	20 (7.22%)
0.2 ~ 0.4	4905 (07.80%)	27 (9.75%)
0.4 ~ 0.6	7406 (11.78%)	44 (15.88%)
0.6 ~ 0.8	10167 (16.17%)	69 (24.91%)
0.8 ~ 1.0	39236 (62.41%)	117 (42.24%)
Total	62876 (100.00%)	277 (100.00%)

D. Comparison With the Usgs Hwms and Inundation Map

The USGS inundation map [Fig. 8(b)] shows similar patterns as the RFP [Fig. 8(a)]. In the RFP, areas with higher probabilities match well with the USGS inundated areas. More importantly, our results extracted the areas with high probabilities all over the research area that is superior to the boundary restricted, binary USGS inundation map. For example, in the subset along Gills Creek [Fig. 8(c)], areas closer to the creek are identified with relatively high flooding potentials although it is beyond the USGS survey boundary.

An RFP histogram comparison was conducted between the whole research area and those within the USGS inundated boundary (Fig. 9). The result suggests a "U-"shaped histogram for the FP within the research area where the histogram peaks occur in both low and high ends, suggesting a bimodal distribution pattern with a mean flood probability of 0.34 [Fig. 9(a)]. After confining the RFP within the USGS survey boundary, it reveals a monomodal pattern with a peak in high end with a mean of 0.83. Detailed statistics can be seen from Table III, where 62.41% of pixels constrained within the boundary have the fps larger than 0.80. This indicates that areas within USGS survey boundary tend to have significantly higher flooding potentials than those beyond the boundary.

We also extracted the FP at the USGS-surveyed HWMs. It is suggested that 67.15% of the HWMs have the fps higher than 0.6 (Table III). Given the fact that these HWMs are not likely to be distributed inside water bodies where the highest fps exist, but rather mostly alongside waterbodies, the RFP in this paper matches well with the HWMs.

In short, this paper develops a flood inundation reconstruction model that logistically combines the DEM, gauge readings, RS images, and social media to achieve a near RT FP prediction. The initial flooding surface derived from RT DEM/gauge readings successfully identified the low-elevation floods. The satellite-derived NDWI contributes to identifying areas in high-elevation zone where DEM/gauge data analysis fails to cover. Further integration of RT social media indicates that even a small amount of tweets data can significantly enhance the prediction of local floods. Superior to the official USGS map that has a delay of four months, the resulted RFP in this paper is much less time consuming and labor inexpensive and is not restricted to the survey boundary. Its ability to provide continuous FPs largely contributes to a rapid and more accurate understanding of areas in need of urgent attention.

V. CONCLUSION

Taking the 2015 SC flood as the study case, this paper proposed a rapid flooding inundation reconstruction model to enhance the delayed remote sensed observations with spatially isolated RT river gauges and twitter data. The primary findings are as follows.

1) The inclusion of multiple gauges compensates the uncertainties from local topological unevenness and the flaws of gauge data itself, and successfully identifies the low-elevation flooded areas.

2) Satellite-derived wetness contributes to identifying the high-elevation flooded areas, and its information loss due to delayed observation is compensated by integrating with the gauge and DEM-based flood layer.
3) Local enhancement with RT tweets proves that even a small amount of crowdsourcing data can largely improve the identification of high flood probability areas during a flood event.

This paper provides a spatially continuous probability surface that measures the likeliness of a certain area being inundated during a flood event in a near-RT manner. The results in this paper can greatly benefit local authorities and first responders for a rapid and comprehensive understanding of flooding situations. The proposed model could be generalized to other flooding cases. Other crowdsourcing databases could also be involved to provide supplemental information, aiding in a more robust local awareness. The methodology used in this paper could seed a wide range of future flood studies for rapid and improved flood situational awareness in a city as well as at a regional level.

REFERENCES

[1] R. A. Pielke, Jr., and M. W. Downton, "Precipitation and damaging floods: Trends in the United States, 1932–97," *J. Climate*, vol. 13, no. 20, pp. 3625–3637, 2000.

[2] (2018). *NOAA's National Weather Service*. Accessed: Jan. 26, 2018. [Online]. Available: http://www.nws.noaa.gov/hic/

[3] A. Sarhadi, S. Soltani, and R. Modarres, "Probabilistic flood inundation mapping of ungauged rivers: Linking GIS techniques and frequency analysis," *J. Hydrol.*, vols. 458–459, pp. 68–86, Aug. 2012.

[4] B. Sanders, "Evaluation of on-line DEMs for flood inundation modeling," *Adv. Water Resour.*, vol. 30, no. 8, pp. 1831–1843, 2007.

[5] (2018). *USGS Surface-Water Data for the Nation*. Accessed: Jan. 26, 2018. [Online]. Available: https://waterdata.usgs.gov/nwis/sw

[6] Z. Li, C. Wang, C. T. Emrich, and D. Guo, "A novel approach to leveraging social media for rapid flood mapping: A case study of the 2015 South Carolina floods," *Cartogr. Geograph. Inf. Sci.*, vol. 45, no. 2, pp. 97–110, 2017.

[7] M. Triglav-Čekada and D. Radovan, "Using volunteered geographical information to map the November 2012 floods in Slovenia," *Natural Hazards Earth Syst. Sci.*, vol. 13, no. 11, pp. 2753–2762, 2013.

[8] K. McDougall and P. Temple-Watts, "The use of LiDAR and volunteered geographic information to map flood extents and inundation," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 4, pp. 251–256, Aug. 2012.

[9] M. F. Goodchild, "Citizens as sensors: The world of volunteered geography," *GeoJournal*, vol. 69, no. 4, pp. 211–221, 2007.

[10] F. E. A. Horita, J. P. de Albuquerque, L. C. Degrossi, E. M. Mendiondo, and J. Ueyama, "Development of a spatial decision support system for flood risk management in Brazil that combines volunteered geographic information with wireless sensor networks," *Comput. Geosci.*, vol. 80, pp. 84–94, Jul. 2015.

[11] E. Schnebele, G. Cervone, and N. Waters, "Road assessment after flood events using non-authoritative data," *Natural Hazards Earth Syst. Sci.*, vol. 14, no. 4, pp. 1007–1015, 2014.

[12] J. Fohringer, D. Dransch, H. Kreibich, and K. Schröter, "Social media as an information source for rapid flood inundation mapping," *Natural Hazards Earth Syst. Sci.*, vol. 15, no. 12, pp. 2725–2738, 2015.

[13] J. Huerta, S. Schade, and C. Granell, *Connecting a Digital Europe Through Location and Place*. Cham, Switzerland: Springer, 2014, pp. 55–71.

[14] S. P. Jackson, W. Mullen, P. Agouris, A. Crooks, A. Croitoru, and A. Stefanidis, "Assessing completeness and spatial error of features in volunteered geographic information," *ISPRS Int. J. Geo-Inf.*, vol. 2, no. 2, pp. 507–530, 2013.

[15] R. Feick and S. Roch, "Understanding the value of VGI," in *Crowdsourcing Geographic Knowledge*. Amsterdam, The Netherlands: Springer, 2012, pp. 15–29.

[16] B. Haworth and E. Bruce, "A review of volunteered geographic information for disaster management," *Geogr. Compass*, vol. 9, no. 5, pp. 237–250, 2015.

[17] L. C. Smith, "Satellite remote sensing of river inundation area, stage, and discharge: A review," *Hydrol. Process.*, vol. 11, no. 10, pp. 1427–1439, Aug. 1997.

[18] Y. Wang, J. D. Colby, and K. A. Mulcahy, "An efficient method for mapping flood extent in a coastal floodplain using Landsat TM and DEM data," *Int. J. Remote Sens.*, vol. 23, no. 18, pp. 3681–3696, 2002.

[19] P. A. Brivio, R. Colombo, M. Maggi, and R. Tomasoni, "Integration of remote sensing data and GIS for accurate mapping of flooded areas," *Int. J. Remote Sens.*, vol. 23, no. 3, pp. 429–441, 2002.

[20] K. E. Joyce, S. E. Belliss, S. V. Samsonov, S. J. McNeill, and P. J. Glassey, "A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters," *Prog. Phys. Geogr., Earth Environ.*, vol. 33, no. 2, pp. 183–207, 2009.

[21] G. Mallinis, I. Z. Gitas, V. Giannakopoulos, F. Maris, and M. Tsakiri-Strati, "An object-based approach for flood area delineation in a transboundary area using ENVISAT ASAR and LANDSAT TM data," *Int. J. Digit. Earth*, vol. 6, pp. 124–136, Apr. 2011.

[22] S. K. Mcfeeter, "The use of the normalized difference water index (NDWI) in the delineation of open water features," *Int. J. Remote Sens.*, vol. 17, no. 7, pp. 1425–1432, 1996.

[23] S. K. Jain, A. K. Saraf, A. Goswami, and T. Ahmad, "Flood inundation mapping using NOAA AVHRR data," *Water Resour. Manage.*, vol. 20, no. 6, pp. 949–959, 2006.

[24] S. K. Jain, R. D. Singh, M. K. Jain, and A. K. Lohani, "Delineation of flood-prone areas using remote sensing techniques," *Water Resour. Manage.*, vol. 19, no. 4, pp. 333–347, 2005.

[25] H. R. Fazeli, M. N. Said, S. Amerudin, and M. Z. A. Rahman, "A study of volunteered geographic information (VGI) assessment methods for flood hazard mapping: A review," *J. Teknol.*, vol. 75, no. 10, pp. 127–134, 2015.

[26] G. Andre, R. Guillande, and F. Bahoken, "Flood mapping using spatial radar and optical imagery and digital elevation model: Limits and capacities," *Houille Blanche*, no. 1, pp. 49–54, Jan. 2002.

[27] X. Huang, C. Wang, and Z. Li, "A near real-time flood-mapping approach by integrating social media and post-event satellite imagery," *Ann. GIS*, vol. 24, no. 2, pp. 113–123, 2018.

[28] E. Schnebele and G. Cervone, "Improving remote sensing flood assessment using volunteered geographical data," *Natural Hazards Earth Syst. Sci.*, vol. 13, no. 3, pp. 669–677, 2013.

[29] E. Schnebele, G. Cervone, S. Kumar, and N. Waters, "Real time estimation of the calgary floods using limited remote sensing data," *Water*, vol. 6, no. 2, pp. 381–398, 2014.

[30] G. Cervone, E. Sava, Q. Huang, E. Schnebele, J. Harrison, and N. Waters, "Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study," *Int. J. Remote Sens.*, vol. 37, no. 1, pp. 100–124, 2015.

[31] J. W. Musser, K. M. Waston, J. A. Painter, and A. J. Gotvald. (2016). *Flood-Inundation Maps of Selected Areas Affected by the Flood of October 2015 in Central and Coastal South Carolina*. [Online]. Available: https://pubs.usgs.gov/of/2016/1019/ofr20161019.pdf

[32] R. Berg. (2016). National Hurricane Center Tropical Cyclone Report. Hurricane Joaquin (AL112015), NOAA. [Online]. Available: https://www.nhc.noaa.gov/data/tcr/AL112015_Joaquin.pdf

[33] T. D. Feaster, J. M. Shelton, and J. C. Robbins. (2015). *Preliminary Peak Stage and Streamflow Data at Selected USGS Streamgaging Stations for the South Carolina Flood of October 2015*. [Online]. Available: https://pubs.usgs.gov/of/2015/1201/ofr20151201.pdf

[34] J. D. Murphy. (2015). The Historic South Carolina Floods of October 1–5, 2015. National Weather Service, Silver Spring, Maryland. [Online]. Available: https://www.weather.gov/media/publications/assessments/SCFlooding_072216_Signed_Final.pdf

[35] (2018). *USGS Current Water Data for South Carolina*. Accessed: Jan. 26, 2018. [Online]. Available: https://waterdata.usgs.gov/sc/nwis/rt

[36] (2018). *USGS Statistics Web Service*. Accessed: Jan. 26, 2018. [Online]. Available: https://waterservices.usgs.gov/rest/Statistics-Service.html

[37] B. Gao, "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 257–266, 1996.

[38] U. Howard Perlman. (2018). *High-Water Marks and Stream Stage, USGS Water-Science School*. Accessed: Jan. 26, 2018. [Online]. Available: https://water.usgs.gov/edu/highwatermarks.html

Xiao Huang (S'18) received the bachelor's degree in remote sensing information and engineering from Wuhan University, Wuhan, China, in 2015, and the M.S. degree in geographic information science and technology from the Georgia Institute of Technology, Atlanta, GA, USA, in 2016. He is currently pursuing the Ph.D. degree with the University of South Carolina, Columbia, SC, USA.

His research interests include the application of remote sensing, geographic information science, and the technical development of rapid response to natural disasters incorporating social media.

Mr. Huang is a member of the Association of the American Geographers and a student member of the International Association of Chinese Professionals in Geographic Information Sciences.

Cuizhen Wang received the Ph.D. degree in geography from Michigan State University, East Lansing, MI, USA, in 2004, and the Ph.D. degree in photogrammetry and remote sensing from the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China, in 1999.

She is currently a Professor with the Department of Geography, University of South Carolina, Columbia, SC, USA, with her expertise in bioenvironmental remote sensing. Her research interests include optical/radar radiative transfer model and synergy, vegetation mapping, environmental stress monitoring, bioenergy land use change, and coastal environment change.

Dr. Wang is a USGIF Subject Matter Expert of remote sensing and image analysis, and serves on the Editorial Boards of the *International Journal of Digital Earth* and the *International Journal of Biometeorology*.

Zhenlong Li received the Ph.D. degree in earth systems and geoinformation sciences from George Mason University, Fairfax, VA, USA, in 2015.

He is currently the Founding Director with the Geoinformation and Big Data Research Laboratory, University of South Carolina, Columbia, SC, USA, where he is an Assistant Professor with the Department of Geography. His research interests include GIScience with a focus on big data analytics, spatial computing, and geospatial cyberinfrastructure with applications to disaster management and climate analysis.

Dr. Li serves on the Editorial Board of the *ISPRS International Journal of Geo-Information* and as the Vice Chair of the Association of the American Geographers CyberInfrastructure Specialty Group.