

WATER DEPTH MAPPING OF HURRICANE HARVEY USING VOLUNTEERED
GEOGRAPHIC INFORMATION IN HARRIS COUNTY, TEXAS

by

Abdullatif Alyaqout, B.A.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Science
with a Major in Geography
August 2018

Committee Members:

Edwin T. Chow, Chair

Richard Dixon

Alexander Savelyev

COPYRIGHT

by

Abdullatif Alyaqout

2018

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

This work is protected by the Copyright Laws of the United States (Public Law 94-553, section 107). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgement. Use of this material for financial gain without the author's express written permission is not allowed.

Duplication Permission

As the copyright holder of this work I, Abdullatif Alyaqout, authorize duplication of this work, in whole or in part, for educational or scholarly purposes only.

DEDICATION

To my parents who spent their life and health who raised me to become the person I am. Whatever I do, I can never pay them back for all the time they spent to make me a responsible person:

“And lower to them the wing of humility out of mercy and say, My Lord, have mercy upon them as they brought me up [when I was] small” (17:24)

To my wife who stood by my side through good and bad times. Her support, patience, and dedication motivated me to complete my graduate studies.

To the joy and happiness of my life, Shuaib, my son.

To my brothers and sisters. Thank you.

To my parents and brothers-in-law. Thank you for your support.

ACKNOWLEDGEMENTS

Praise is to Allah by Whose grace good deeds are completed. I would like to express my deepest gratitude to my advisor, Dr. T. Edwin Chow, for his patience, guidance and encouragement throughout the time I spent during my master's study. His help was beyond the academic boundaries and always directed me to consider my family as the number one priority and to treasure the time with them. Simple words will not be enough and sufficient to reward his hard-work and enthusiasm for me to be able to finish my thesis.

I would like to appreciate the help and the guidance of Dr. Richard Dixon and Dr. Alexander Savelyev to complete my thesis. Their expertise, valuable comments and suggestions improved my knowledge and skills, which will be treasured during my PhD study in the future.

In addition, I would like to appreciate the help of Dr. Nathan Currit for his kindness and valuable knowledge. He welcomed me in his office and spent a lot of time to help with the remote sensing data. Also, I would like to thank Melissa Jurrens for her help with the authoritative data used in this study.

I would like to specially thank Ms. Allison Glass-Smith for her precious help during my master's study.

Finally, I'm grateful to have such a friend. Thank you, Faisal.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	v
LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
ABSTRACT.....	xi
CHAPTER	
I. INTRODUCTION.....	1
Flood Hazards.....	1
Flood Modeling.....	2
Volunteered Geographic Information.....	4
Purpose Statement.....	6
II. LITERATURE REVIEW.....	7
Remote Sensing and Flood Disasters.....	7
Social Media and Crowdsourcing for Flood Analysis.....	11
Text Messages in Social Media.....	11
Using Multimedia in Flood Modeling.....	14
Integration Between Multiple Data Modalities.....	15
Validation Against Authoritative Data.....	16
Gaps of the Literature.....	20
III. METHODOLOGY.....	22
Study Area.....	22
Data.....	23
VGI Data.....	24

Authoritative and RS Data	24
Preprocessing	25
WD Extraction	27
WD Comparison	35
IV. RESULTS	37
VGI Summary	37
VGI Validation and Comparison	43
Precision of VGI Data Modalities.....	43
Spatial Distribution of VGI Data Modalities	44
Temporal Distribution of VGI Data Modalities.....	44
VGI and RS Validation.....	45
VGI and Authoritative Data Comparison	47
VGI and USGS Stream Gauges	47
VGI and FEMA Modeled Depth Grids	49
V. DISCUSSION.....	51
Research Question Findings	52
Limitations	59
VI. CONCLUSION.....	64
REFERENCES	66

LIST OF TABLES

Table	Page
1. Literature Summary.	17
2. Data Summary.	25
3. List of Hashtags and Keywords Used for Relevant Tweets Classification.	26
4. Height of Physical Objects for Multimedia Water Level Extraction.....	29
5. Current and Expected Distribution of VGI Observations for The Chi-Square Test.	45

LIST OF FIGURES

Figure	Page
1. Example of Geotagged Information About Floods and Water Level Shared by Different Users in Multiple Data Modalities Using Social Media Platforms.	5
2. Map of The Study Area.....	23
3. An Example of Water Level Indicated in a Text VGI.	28
4. An Example of Extracting Water Level from a Picture Using the Surrounding Physical Objects.	29
5. An Example of Water Level Information Included in both Text and Multimedia for a Single Post.	30
6. Twitter and Crowdsourced Data Processing and WD Extraction.....	31
7. WD Extraction from RS Water Bodies.....	32
8. Remote Sensing Processing and WD Extraction.	33
9. Overall Research Design.....	34
10. Summary of Relevant and Non-Relevant Tweets in Harris County.....	38
11. Summary of Relevant Tweets for Text and Multimedia.	38
12. Summary of Tweets with Water Depth Information Compared to Relevant Tweets on a Daily Basis.	40
13. Summary of the Variation in Water Depth Extracted from the Three VGI Data Modalities (Text, Picture, and Video) on a Daily Basis.	41
14. Geographic Distribution of both VGI Sources, Social Media and Crowdsourcing in Harris County.....	42
15. Geographic Distribution of VGI Data Modalities in Harris County.....	43
16. WD Extraction from RS Images.	46

17. Geographic Distribution of USGS Stream Gauges in Harris County.....	48
18. Modeled Depth Grids in Meters from FEMA.....	49
19. Daily Comparison Between Total Relevant Tweets and Average Water Level from the USGS Gauges.	51
20. Mean Center and Directional Distribution of VGI Data Modalities in Harris County.....	53
21. VGI Data Modalities Kernel Density	55
22. Digital Divide Index (DDI) in Harris County.....	56
23. Power Outage Between August 27th and September 1st in Harris County.....	57
24. An Example of a Tweet with Relevant Hashtag But without Relevant Water Content.....	60
25. An Example of Relevant Tweets with Missing or Broken Links.	61
26. An Example of a Re-used Picture During the Event.	62

ABSTRACT

Natural hazards cause catastrophic damages to both population and economy. In the U.S., floods are the costliest hazard. In 2017, Hurricane Harvey made landfall along the Texas coast on August 25th and lasted for five days. It was one of the most destructive hurricanes in the history of the state. In order to enhance emergency response and management, it is essential to have a better understanding of the flood status, risks and conditions.

In flood modeling, conventional data sources include remote sensing, high water marks (HWMs) from field survey, and stream gauges are generally used. The availability of Volunteered Geographic Information (VGI), such as tweets and crowdsourced data, empowered the researchers to model flood (e.g. Water Depth (WD)) in near-real-time by integrating multi-sourced data available. Nevertheless, the quality of VGI and its reliability for flood analysis is not well understood and validated by empirical data. Therefore, the primary objective of this study was to evaluate the quality of multiple VGI data sources, especially the multimedia that include pictures and videos, against authoritative data for inundation mapping. This study collected the geospatial data from multiple sources to analyze the changing WD during Hurricane Harvey in Harris County, Texas. First, WD was generated from three VGI data modalities: (1) text, (2) pictures, and (3) videos, and they were compared against each other using Friedman test and Chi-square. Then, the VGI-derived WD was synthesized and consolidated to reconstruct the time-series of WD in Harris County. Finally, the quality of synthesized WD from VGI

was validated against remote sensing (RS) and two authoritative data: (1) water level records from stream gauges at discrete locations, and (2) modeled depth grids by Federal Emergency Management Agency (FEMA) using paired t -test. The results showed that there was no statistically significant differences among all VGI data modalities in terms of precision, while it showed significant difference in terms of spatial and temporal characteristics. In addition, the results showed that there was a statistically significant difference between VGI WD and RS WD. Finally, the analysis revealed that there was no statistically significant difference between VGI data and water records from the stream gauges, while it showed a statistically significant difference when VGI were compared with the depth grids from FEMA.

I. INTRODUCTION

Flood Hazards

Natural hazards are considered to be a major source of catastrophic damages and losses both to the population and to the economy at a global scale. It has been estimated that hazards during the first 12 years in the 21st century caused up to US\$1.7 trillion in losses to the economy and more than 1.4 million population casualties around the globe (Montz, Tobin, and Hagelman 2017). About 30% of the world's land area is flood prone, which affects approximately 80% of the world's population (Dilley et al. 2005).

In the U.S., floods are the costliest hazard, in term of human lives, property damage, and economic losses (Strömberg 2007). About 40% of hazard events occurring between 1900 and 2015 in the U.S. were major floods, which makes them the leading cause of natural disaster losses (United States Geological Survey 2016; Cigler 2017). These losses can destroy many businesses as well. According to Federal Emergency Management Agency (FEMA 2015), about 40% of small and individual businesses, small in term of size of employees and run individually, do not reopen following a disaster. Two of the most destructive flood events recorded in the U.S. were Hurricane Katrina in 2005 and Superstorm Sandy in 2012, with damages costing more than US\$24 billion for both hurricanes, US\$8.6 billion losses caused by Sandy and US\$16.3 billion caused by Katrina (FEMA 2018). In urban areas, land cover influences the magnitude and variation of floods in terms of infiltration, drain pipes and ditches, urban climate change, and stream channel alteration (Marsh and Grossa 2005).

Flood Modeling

One concern that has been repeatedly raised regarding the flood damage is associated with disaster planning in the flood-prone locations such as floodplains, which represent about 7% of the U.S. land mass (Marsh and Grossa 2005). Due to the calamitous impact of floods on the urban environment, various methods and models have been developed to assess and support emergency planning, risk analysis, and mitigation of damage related to floods. In particular, identifying flood-prone areas is critical for risk estimation and damage assessment during flood events.

In general, modeling is defined as a simplified representation of real-world phenomena to serve multiple purposes and answer questions about how such phenomena behave (Guinot and Gourbesville 2003). The goal of flood modeling is to generalize the relationships among rainfall intensity, watershed characteristics, surface and subsurface runoff, and river discharge. There are three approaches of flood modeling that can be used to generalize such relationships, including statistics, physical models, and machine-learning. A common example of statistical models is regression (Ahearn 2004), and the simulated hydrographs for flood probability estimation (SHYPRE) (Arnaud and Lavabre 2002). Physical model describes the hydrologic process of natural laws (e.g. conservation of mass, conservation of energy) operating on environmental variables, such as topography and fluvial system. Such models are applicable in a wide range of situations. Examples of physical models are the HEC-RAS and the SHE/MIKE SHE (Abbott et al. 1986; Brunner 2002). A main product of physical models during flood events is water depth (WD). WD represents the depth of the water relative to the point where it was measured or observed on the land surface. Some of these models involve the integration

of a geographic information system (GIS) and various physical (Casas et al. 2006), statistical (Pradhan 2010), and machine learning approaches (Kia et al. 2012). In an extension of that approach, Tsakiris (2014) combined GIS with the flood-modeling package FLOW-R2D to propose a new systematic paradigm in flood modeling and risk assessment. Other work on flood modeling has taken advantage of advanced machine learning methods by combining artificial neural networks (ANNs) and radar weather data for flash-flood modeling (Dinu et al. 2017).

Since landscape complexity in urban areas influences the accuracy of flood modeling, using detailed topographic data can reduce the uncertainties of inundation-mapping. In fact, remotely sensed (RS) data is used extensively in many flood modeling studies because of its spatiotemporal coverage and availability at a multi-geographic scale. Hese and Heyer (2016) integrated high-spatial resolution RapidEye multispectral data with ASTER GDEM V2 for rapid post-tsunami flood mapping after the 2011 Japan tsunami. The recent applications of photogrammetry and computer vision technologies in constructing 3D objects from 2D photographs are exploited extensively in flood modeling to extract detailed surface elevations using structure from motion (SFM) techniques (Meesuk et al. 2012). Terrestrial laser scanners (TLS) or light detection and ranging (lidar), has also been used to produce high-resolution digital elevation model (DEM) (Sampson et al. 2012). Papaioannou et al. (2016) used TLS and a multi-hydraulic modeling approach for flood mapping at ungauged areas in Greece. Although working with high-resolution topographic data produces high-quality flood modeling outputs, it may not be sufficient for rapid flood mapping scenarios (Hese and Heyer 2016).

Volunteered Geographic Information

More recently, there has been an emerging interest in using Volunteered Geographic Information (VGI) as a source of data to analyze and model flood and related human behaviors for better situational awareness and real-time support for emergency management (Smith et al. 2015). The major perception of VGI is that it is collaborative geographic data generated by users through crowdsourcing or individual platforms (Goodchild 2007). VGI is a non-authoritative source, as the information is not collected using traditional methods by official authorities or agencies (Schnebele et al. 2014). VGI sources can be derived from crowdsourced projects and/or social media. Crowdsourced data are based on tasks undertaken by a group of volunteers, rather than an individual, as in the case of social media, to contribute mass inputs for a specific project (Howe 2006). Social media data are commonly described as the broadcasting of individual micro-blog in real-time using social platforms. People can share their updated statuses as well as their geographic location, which comes in handy during disaster events (Smith et al. 2015).

In a flood event, both social media and crowdsourced data can provide information in real time with details on water conditions and the damage situation, which is an advantage over traditional methods that might take a long time for data collection and processing (Li et al. 2017). Guan and Chen (2014) observed the spatiotemporal activity of Twitter data before, during, and after Hurricane Sandy to understand how the rhythms of life evolve through time. A new geographic approach has also been developed that combines both authoritative geographic information and social media data for crisis management and disaster response (de Albuquerque et al. 2015). Furthermore, Li et al.

(2017) suggested a new approach to enhance the use of Twitter data to delineate a near real-time inundation map for the 2015 floods in South Carolina. Besides the text message and its location, multimedia content, including pictures and videos, from social media or crowdsourcing, may include important information about a disaster landscape that can be extracted. For example, Fohringer et al. (2015) mapped the inundation depth of flooded areas by using relevant pictures posted on social media. Schnebele et al. (2014) combined both authoritative and non-authoritative data for flood mapping and for generating maps of road damage. Their work included both pictures and videos for Massachusetts and New Jersey. Figure 1 illustrates how such relevant information is contained in multiple VGI forms.

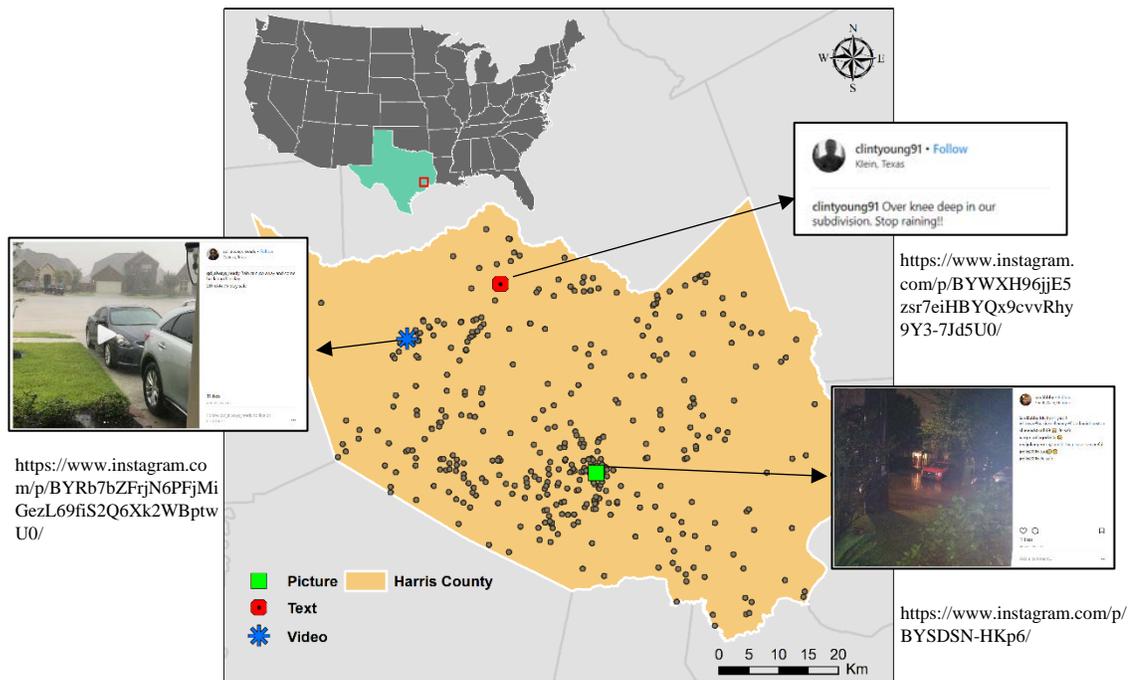


Figure 1. Example of Geotagged Information About Floods and Water Level Shared by Different Users in Multiple Data Modalities Using Social Media Platforms. (Data sources: TNRIS)

Purpose Statement

Rapid flood mapping is essential to disaster response and relief, emergency planning, and future mitigation. In a disaster event when the time is critical, and human lives are at stake, all sources of relevant information about a flood and the associated damage should be considered and collected. Since traditional methods and data sources are limited in terms of the spatial and temporal resolutions, VGI sources can provide an alternative, and perhaps supplementary to conventional sources, insights to be used in flood monitoring. Although VGI data can be acquired from multiple sources, it requires a lot of pre-processing and the data quality is uncertain. Therefore, it is important to evaluate the accuracy and representativeness of such data for flood mapping. The primary objective of this study was to evaluate the quality of multiple VGI data sources against authoritative data for inundation mapping. Hence, the research questions pursued was as follows:

- 1- Are there any significant differences in WD among different VGI data modalities (text/pictures/videos) in terms of:
 - a. Precision?
 - b. Spatial?
 - c. Temporal?
- 2- Are there any significant differences between synthesized VGI WD and RS WD?
- 3- Are there any significant differences between synthesized VGI WD and:
 - a. Authoritative stream gauge data?
 - b. Authoritative modeled depth grids?

II. LITERATURE REVIEW

Flood monitoring can vary by its approaches, data formats and data sources collected at the associated scales. Based on the research objective, this chapter will review three interrelated themes important to flood monitoring. First, the chapter will review the application and usage of RS data to extract and delineate water bodies using water spectral characteristics in order to model the extent of flooded areas. Second, the chapter will examine the contribution of VGI in flood hazards as an alternative source to leverage flood modeling by extracting water-relevant information expressed and shared by the public through different web platforms either in the form of text or multimedia. Finally, this chapter concludes with the findings of previous studies that validated flood models using VGI as a source of data against authoritative models to identify and measure differences and similarities between both models.

Remote Sensing and Flood Disasters

In addition to the in-situ water data collected from stream gauges and field surveys, multiple studies (e.g. Gregg and Casey 2004; Karaska et al. 2004) have deployed RS techniques to obtain and augment water information using the spectral reflectance characteristics of water (Jensen 2009). In general, these previous studies addressed the application of RS for flood analysis in two main directions: (1) developing and comparing water extraction indices (e.g. Amarnath 2014), and (2) using spatial sharpening algorithms to enhance the spatial resolution of moderate spectral bands in RS data (e.g. Du et al. 2016). The development of multiple water indices and comparing them against existing indices provide the users with more insights to select the proper RS data to extract water bodies based on the spectral capabilities used in these indices.

Amarnath (2014) developed an algorithm for the normalized difference surface water index (NDSWI), using short-wave infrared (SWIR), near-infrared (NIR), and green bands of Landsat5 TM for detecting and extracting water bodies using the following equation:

$$NDSWI = G * \frac{\rho_{swir} - \rho_{nir}}{\rho_{swir} + c1\rho_{nir} - c2\rho_{green}} \quad (1)$$

where G is a given gain factor, c1 and c2 are applied to adjust backscattering and absorption coefficients, and ρ is the reflectance of SWIR, NIR, and green bands respectively. Amarnath (2014) compared the proposed algorithm with previous approaches for mapping flooded areas using the normalized difference water index (NDWI) equation:

$$NDWI = \frac{\rho_{nir} - \rho_{swir}}{\rho_{nir} + \rho_{swir}} \quad (2)$$

and the normalized difference vegetation index (NDVI) using the following equation:

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \quad (3)$$

He concluded that the NDSWI results showed better results in inundation areas detection compared to the NDVI and the NDWI because of the ability of SWIR band used in NDSWI to detect water turbidity, compared to the use of NIR in the other indices. He also illustrated that cloud blockage was a source of limitation in optical sensors, which affects the process of detecting surface water body. In addition, the NDSWI was able to detect water turbidity during floods. However, the addition of in-situ turbidity measurements is suggested to improve the accuracy of detecting and assessing turbidity of floodwaters. Feyisa et al. (2014) proposed the use of an automated water extraction

index (AWEI) to classify water based on using the five spectral bands of Landsat 5 by applying the following equations:

$$AWEI_{nsh} = 4 * (\rho_{band2} - \rho_{band5}) - (0.25 * \rho_{band4} + 0.75 * \rho_{band7}) \quad (4)$$

$$AWEI_{sh} = \rho_{band1} + 2.5 * \rho_{band2} - 1.5 * (\rho_{band4} + \rho_{band5}) - 0.25 * \rho_{band7} \quad (5)$$

where ρ is the reflectance value of Landsat5 TM band 1 (blue), band 2 (green), band 4 (NIR), band 5 (SWIR) and band7 (SWIR). $AWEI_{nsh}$ is the index used to eliminate non-water pixels, while $AWEI_{sh}$ was used to improve the water extraction accuracy by removing the effect of shadows in the pixels. In addition, Feyisa et al. (2014) used different coefficients, determined based on the variation in the reflectance properties of different land cover types, to set a threshold that could distinguish water from non-water pixels. The proposed AWEI model was compared with the modified normalized difference water index (MNDWI) using the following equation:

$$MNDWI = \frac{\rho_{green} - \rho_{swir}}{\rho_{green} + \rho_{swir}} \quad (6)$$

And it was compared with the maximum likelihood (ML) classifier for water extraction. Based on kappa coefficients and error matrices, the classification accuracy of the AWEI was significantly higher than that of the MNDWI and ML in detecting water features, even in areas where shadows or dark features existed. Also, the AEWI algorithm was more accurate in classifying edge pixels, a mix of land and water, than the other compared techniques at all tested sites, where the omission and commission error of AWEI at all the sites, in average, was about 50% of that of the MNDWI and 25% of that of the ML. This result could be an advantage in studies seeking to detect changes in water bodies and to extract water pixels at the edges.

Beside developing water indices, some previous studies focused on using different band sharpening algorithms to provide the users with an approach to enhance the spatial resolution of water bodies extraction when using moderate spatial resolution bands (e.g. SWIR). Despite the availability of multiple algorithms for spatial sharpening in the literature, the user should have high-level skills and knowledge to select and perform the suitable sharpening algorithm. Du et al. (2016) used satellite imageries of Sentinel-2 for water body extraction by leveraging different pan-sharpening techniques to downscale the spatial resolution of the SWIR from 20 m to 10 m and compared the results. The pan-sharpening techniques used were principal component analysis (PCA), intensity hue saturation (IHS), high pass filter (HPF), and the À trous wavelet transform (ATWT). Du et al. (2016) used spectral water indices to extract water features including the NDWI and the MNDWI. They applied the pan-sharpening techniques mentioned above for each water index and compared between the outputs of each index. They found that 10 m MNDWI produced by the four algorithms extracted more accurate water features than 10 m NDWI and 20 m MNDWI (NDWI and 20m MNDWI produced 95.71% and 94.25% average accuracy respectively, while the HPF produced a 10 m MNDWI with 95.94%, which was the least of the four). Among the pan-sharpening algorithms, HPF produced a 10 m sharpened MNDWI image with a higher correlation coefficient (0.9991) and a lower Root Mean Square Error (RMSE) (0.0215) compared to other algorithms. However, in water body extraction, HPF did not show higher accuracy results compared to the other algorithms due to its confusion between water and non-water bodies. ATWT showed more accurate results in water body mapping, especially in

detecting linear water features. For future work, they suggest developing more powerful pan-sharpening algorithms that could be applied to enhance the results of flood mapping.

Social Media and Crowdsourcing for Flood Analysis

Social media is a source of information contributed by individuals through public social platforms, such as Twitter, Facebook, Flickr, etc. This information can be in the form of a text message and/or multimedia (e.g. pictures, video). Previous studies (e.g. de Albuquerque et al. 2015; Fohringer et al. 2015) have explored the use of social media information, either separately or by integrating both text and multimedia forms, for flood disaster analysis. On the other hand, crowdsourced data is collected by a group of non-authoritative volunteers that can be shared through different web platforms (e.g. OpenStreetMap). These two sources are valuable for disaster situations and much of the literature in this field has aimed to leverage the use of data from one or both types of sources for flood analysis and inundation mapping. Previous studies (Guan and Chen 2014; Li et al. 2017) focused on one modality of VGI (e.g. text) to extract and map water level after a flood event. Other work focused on the combination of multiple modalities (e.g. pictures and videos) for flood analysis and data extraction (Fohringer et al. 2015). Regarding damage and risk assessment related to flood events, semantic and sentiment analysis were common approach to exploit social media. Another group of studies (Schnebele et al. 2014; Li et al. 2017) assessed the quality of VGI outputs by comparing it against authoritative sources (e.g. FEMA).

Text Messages in Social Media

Most previous studies on using social media data for flood disaster analysis have examined text information obtained from platforms such as Twitter (Guan and Chen

2014; Li et al. 2017). These researchers attempt to quantify and model the spatial relationship, if there's any, between the kernel density of tweets with the hazardous area. Guan and Chen (2014) related disaster-related ratio (DRR), which is the number of disaster-related tweets in an area divided by the total number of general tweets in the same space, with storm surge and wind damage in four cities (New York City, Philadelphia, Baltimore, and Washington DC). They have examined DRR relationship with hurricane damage data in three spatial categories—coastline proximity using multiple buffer distances and counted the DRR at each buffer distance, urban areas proximity by generating buffer zones around the geometric center of the four cities to explore how DRR varies between the cities, and storm surge and wind damage at county level by calculating the DRR at each county in the study area and compare it against the storm surge and wind damage. Similarly, Li et al. (2017) explored the correlation between flood possibility index (FPI) within flooded areas by using geotagged flood-related tweets and USGS water gauges to extract water height points (WHPs) from DEM. Both studies found that urban areas with large populations tended to have higher social media activity (Guan and Chen, 2014) and that people tended to tweet more about the flood when they were closer to the flooded areas (Li et al. 2017). In addition to just the number of tweets, there was also an overall spatiotemporal positive correlation between the relevance of tweets and the damage from the hurricane (Guan and Chen 2014). In fact, about 10% of flood-related tweets are within the inundated areas identified by the USGS (Li et al. 2017). However, the limited availability of text tweets in less populated areas to generate WHPs led the model to underestimate parts of the flooded areas (Li et al. 2017) and further research is needed regarding the spatiotemporal uncertainties (Guan

and Chen 2014). Nevertheless, it is likely that urban areas with large populations tended to have higher social media activity, so there may be a need to normalize such kernel density by population. This finding seems to hold true across natural hazards, including storm surge and wind damage (Guan and Chen, 2014) and flooding (Li et al. 2017).

Besides the density of hazards-related tweets, recent hazard researches also explored the use of flood information by using qualitative classification (De Albuquerque et al. 2015), as well as semantic and sentiment analysis (Deng et al. 2016; Shalunts et al. 2014). In a bi-lingual study focusing on both English and German, De Albuquerque et al. (2015) leveraged the usage of text information to identify on-topic tweets, and classified them to seven-coded themes (volunteer actions, media reports, traffic conditions, first-hand observations, official actions, infrastructure damage and other). Their findings were consistent with Li et al. (2017) and Guan and Chen (2014) where on-topic tweets are more likely to occur near affected (flooded) areas (≤ 10 km away from the affected areas), in catchments with high relative water levels (+0.75m), as defined by the difference between daily maximum water level and average flood water level throughout the study period.

Other researchers aimed to analyze social behavior during disasters to evaluate the risk and damages both spatially (Deng et al. 2016) and temporally (Shalunts et al. 2014). During Typhoon Haiyan in 2013, Deng et al. (2016) used Weibo (a Twitter-equivalent in China) API to extract text information based on pre-defined keywords. They set forth an index-system model that included the following: (1) index selection, (2) the score of the index, (3) the weight of the index, and (4) an evaluation function. Likewise, Shalunts et al. (2014) proposed a new tool (SentiSAIL) for sentiment analysis

during the Central Europe floods in 2013. They obtained and classified the tweets into positive, negative, mixed, and neutral sentiment. After analyzing the spatiotemporal variation of the crowds sentiments, the degree of risk varies from one place to another before the hurricane made landfall (Deng et al. 2016) and the trend of positive, negative, and neutral sentiments was nearly consistent with the temporal pattern of the flood, with negative sentiment dominating (Shalunts et al. 2014).

Using Multimedia in Flood Modeling

Compared to text, multimedia (e.g. pictures and videos) contains more graphical and contextual details regarding the geographic phenomena at stake. Like text, multimedia forms can be geotagged or not. While it is straightforward to extract the location of a geotagged photo, users may also identify the surrounding landmarks or objects to infer the approximate location of the picture/video. Recently, the availability of multimedia in social media or crowdsourced projects has been brought to the attention of researchers on disaster analysis and modeling.

Beside using text from social media for flood analysis, several studies used multimedia (pictures and videos) to extract and analyze flood information (Fohringer et al. 2015) or for damage assessment (Schnebele et al. 2014). One approach followed a step-by-step model (PostDistiller) for flood information collection (PostCrawler), storing (PostStorage), and visualization (PostExplorer) to estimate and map inundated areas (Fohringer et al. 2015). It was also possible to combine authoritative data (e.g. FEMA storm surge) and fused (using kriging) non-authoritative data (e.g. aerial photographs from the Civil Air Patrol (CAP) and YouTube Videos) for road damage assessment during floods (Schnebele et al 2014). The utilization of multimedia from VGI for flood

analysis could be useful in cases of rapid flood modeling compared to traditional methods e.g. RS, (Fohringer et al. 2015) and in cases of rapid assessment of road damage caused by floods (Schnebele et al. 2014). Nevertheless, extracting water information from the multimedia is a time-consuming process, especially when done manually, and the uncertainty of the location of the post could limit the flood results (Fohringer et al. 2015). However, the VGI multimedia have the potential to show a good agreement when compared with authoritative sources (Schnebele et al. 2014) and water level extracted from VGI could vary from the authoritative sources up to decimeter, which is acceptable in rapid flood analysis situations (Fohringer et al. 2015).

Integration Between Multiple Data Modalities

Most of the work on flood analysis and modeling using VGI relies solely on one data modality, whether text, pictures, or video. Few studies have used multiple data modalities of VGI (Schnebele, Cervone, and Waters 2014), combined them with RS data to analyze the event (Cervone et al. 2015) or suggested to use multiple data modalities (Fohringer et al. 2015) for flood analysis and modeling. For example, Schnebele et al. (2014) combined CAP aerial photos with YouTube videos in their approach. While Cervone et al. (2015) integrated RS imagery with social media data during hazard events for damage assessment. They used Twitter data to select suitable satellite images related to the floods by searching for ten tweets related to the event, as a threshold, in a 100 km² area and selecting satellite images that cover the area. Then, they fused social media data, including Twitter and Flickr, with collected images, using kernel interpolation, for damage assessment analysis to predict which roads were more likely to be damaged and impassable. The integration between multiple data helped filling the gaps of

spatiotemporal data availability or coverage during disasters (Schnebele et al. 2014) and show effectiveness in detecting road damage and estimating road closures (Cervone et al. 2015). Despite the integration of multiple VGI forms in several studies, the scarcity of research in this field presents a need to explore the feasibility, as well as uncertainty, of data integration between VGI and conventional geospatial data in flood analysis.

Validation Against Authoritative Data

Authoritative data—which is any form of data collected, modeled, and distributed by official authorities or agencies. In flood modeling, such data include the stream gauge data from USGS or inundation maps from FEMA, which provides reference data for validation to measure the accuracy of non-authoritative data. Multiple studies validated their results with authoritative data for quality assessment. Some results showed an overall 83% match between the modeled inundation map from twitter and the official maps produced by the USGS (Li et al. 2017), while others found good agreement between the non-authoritative data, the FEMA flooded extent and the CAP aerial photos (Schnebele et al. 2014). In general, assuming that the authoritative data are of higher quality, there were some uncertainties regarding the reliability of the non-authoritative data and the accuracy of their geolocation. Table 1 summarizes the literature review.

Table 1. Literature Summary.

Article	Objectives	Data	Methodology	Findings
Guan and Chen 2014	To observe the spatiotemporal activity of social media related to a natural disaster based on Twitter activity pre-, during and post- a disaster	<ul style="list-style-type: none"> • Tweets • Storm surge and wind damage data 	<ul style="list-style-type: none"> • Preprocessing for relevant tweets • Kernel density mapping and multi-scale proximity analysis 	<ul style="list-style-type: none"> • Urban areas with large population tended to have higher social media activity. • The study confirmed the relationship between social disruption magnitude and Hurricane Sandy.
Li et al. 2017	To provide an approach to enhance the use of Twitter data to delineate a near-real-time inundation map	<ul style="list-style-type: none"> • Tweets • Stream gauge • DEM • Official inundation maps 	<ul style="list-style-type: none"> • Spatiotemporal analysis of tweets • Generating a FPI using kernel density • Validating against USGS data 	<ul style="list-style-type: none"> • People tended to tweet more about the flood more when they were spatially close to flood areas and when the flood magnitude increased. • The model validation showed about an 83% match between the modeled inundation map and the official maps. • The limitations of the model involved the availability of tweets to cover the study area.
de Albuquerque et al. 2015	To develop a new geographic approach that combines both authoritative geographic information and social media data for crisis management and disaster response	<ul style="list-style-type: none"> • Georeferenced tweets • Water level data • DEM • HydroSHEDS drainage direction 	<ul style="list-style-type: none"> • The model is divided into geoprocessing analysis for hydrologic data (ArcHydro) and processing of tweets (theme classification). 	<ul style="list-style-type: none"> • On-topic tweets tended to be higher near affected (flooded) areas in catchments with high relevant water levels. • Tweets close to the affected areas were more likely to be on-topic than tweets close to non-affected areas, which agreed with the hypothesis of the study. • The contribution of this study was the clear evidence of a strong spatial relationship between the proximity to flood areas and useful messages (tweets) for disaster management and response.

Article	Objectives	Data	Methodology	Findings
Fohringer et al. 2015	To use multimedia posted on social media to map inundation depth	<ul style="list-style-type: none"> • Water level data • DTM and DEM • Flooded areas • SPOT 5 • Twitter stream • Flickr 	<ul style="list-style-type: none"> • Multi-processing steps to extract, filter, store, analyze, and explore the data 	<ul style="list-style-type: none"> • Quantitative information could be extracted from the social media. • Limitations related to the uncertainty of geolocation of posts and topographic data.
Schnebele et al. 2014	To leverage the use of non-authoritative data for flood mapping and damage assessment	<ul style="list-style-type: none"> • Georeferenced videos • Tweets • Aerial photographs • Storm surge GIS layers • Roads 	<ul style="list-style-type: none"> • A Multi-level approach to generate damage assessment map, assess damage from videos, fuse CAP data and videos using Kriging, and classify damaged roads 	<ul style="list-style-type: none"> • Social media can be used to leverage damage assessment. • There was a good agreement between the non-authoritative data and the FEMA flooded extent and the CAP classified data. • There were some uncertainties regarding the reliability of the non-authoritative data and the accuracy of their geolocation.
Amarnath, G. 2014	To develop a new algorithm for water body extraction using remote sensing	<ul style="list-style-type: none"> • Landsat TM • EO-1 ALI • ALOS/POLSAR 	<ul style="list-style-type: none"> • Preprocessing of images • Applying NDSWI algorithm • Comparing extracted water features pre-, during, and post-floods for damage assessment • Comparing algorithm outputs with NDVI and NDWI 	<ul style="list-style-type: none"> • NDSWI showed better results in flood mapping compared to NDVI and NDWI. • Limitations were associated with the optical sensor capabilities, such as spatial and temporal resolution. • The addition of in-situ turbidity measurements should improve the inundation-mapping model.
Shalunts et al. 2014	To present a modified tool SentiSAIL for sentiment analysis during natural disaster events	<ul style="list-style-type: none"> • Tweets 	<ul style="list-style-type: none"> • A Multi-level approach to classifying tweets as positive, negative, mixed, and neutral 	<ul style="list-style-type: none"> • The trend of the sentiment was nearly parallel with the temporal pattern of the flood. • A combination of social media and other media types should be considered.

Article	Objectives	Data	Methodology	Findings
Deng, Q. 2016	To introduce a new method for risk and damage assessment by extracting sense from the crowd	<ul style="list-style-type: none"> • Weibo 	<ul style="list-style-type: none"> • Multi-index model and social media semantic analysis 	<ul style="list-style-type: none"> • Based on official damage reports, the indices were consistent with losses in actual locations.
Feyisa, G. 2014	To develop a water index, the Automated Water Extraction Index (AWEI), that improves water body extraction with the existence of different environmental noise factors	<ul style="list-style-type: none"> • Landsat TM • Google Earth • True water boundaries digitized manually • ASTER DEM 	<ul style="list-style-type: none"> • Preprocessing of images and multi-model for pixel extraction 	<ul style="list-style-type: none"> • The proposed technique was able to detect water features in areas where shadows or dark features existed significantly. • The AEWI was more accurate in classifying edge pixels than the other compared techniques, which could be an advantage in studies seeking to detect changes in water bodies.
Du, Y. 2016	To use Sentinel-2 products for water features extraction and to compare multiple pan-sharpening techniques	<ul style="list-style-type: none"> • Sentinel-2 data L1-C 	<ul style="list-style-type: none"> • Applying multi-spectral indices for water features extraction and multi-pan-sharpening methods for spatial downscaling 	<ul style="list-style-type: none"> • MNDWI showed better results than NDWI in distinguishing between water bodies and built-up areas, and it enhanced the detected water bodies. • 10m MNDWI extracted more accurate water body maps than 10m NDWI and 20m MNDWI.

Gaps of the Literature

The literature discussed how different forms of VGI data can be implemented, either combined or separately, in flood analysis and modeling. Furthermore, the literature showed how RS data can be used to extract water bodies using different techniques and algorithms. With all the wealth of information available and methodologies presented in the literature in regard to flood analysis, the gaps of the literature can be summarized as follows:

- 1- The use of multiple data modalities from VGI for flood monitoring is limited. The value of output from each data modality should be validated in terms of accuracy, spatial, and temporal differences and similarities.
- 2- It is unclear how VGI can be used in conjunction with other data, i.e. RS, for flood modeling. Both data should be investigated and compared to reveal similarities and differences in terms of synthesizing WD data during flood events.
- 3- In addition, the validation between integrated VGI data modalities (text, pictures, and videos) and authoritative flood data should be explored and measured in terms of quality assessment.

Integrating various sources, including RS, VGI, authoritative and other geospatial data may have the potential to enhance flood analysis and modeling at finer spatial and temporal resolutions. Moreover, exploring the quality of different VGI data modalities for flood analysis and modeling should benefit researchers and first responders during conditions when certain data, such as water level or damage magnitude, are not available for rapid assessment or are difficult to collect during hazardous situations. Thus, measuring the quality of these data modalities and assessing its quality against traditional

data sources (e.g. RS and stream gauges) is essential to empower the community to respond, prepare, mitigate and recover from future hazards.

III. METHODOLOGY

Study Area

In this study, the area investigated was Harris County, located in the eastern part of central Texas within the Houston Metropolitan Area (Figure 2). Harris County is the third-largest populated county in the U.S. with an estimated population of 4.6 million in 2016 (U.S. Census Bureau n.d.) and a land area of about 4400 km². Harris County is within the subtropical humid region of Texas with average annual precipitation between 111.7 and 121.9 cm and an average annual temperature of 20°C (Larkin and Bomar 1983). The county's ecoregion is mainly western gulf coastal plain mixed with a minor amount of south-central plains (TPWD n.d.).

Despite the county's urban growth and development through time, it experienced massive natural disasters over the last decade. In 2011, Storm Allison left the county with 22 lives lost and 20,000 homes damaged. The estimated costs of damage were about \$20 billion (Hwang and Lee 2017). The county experienced another catastrophic flood event in 2016 that left at least seven people dead and more than 6,500 houses damaged (Floodlist 2016). In 2017, Hurricane Harvey made landfall along the Texas coast on August 25th and lasted for five days, until August 29th. It was one of the most destructive hurricanes in the history of the state, regarding population damage and economic losses, with 1.4 m of rain and 209.2 km/h winds at its peak. It was estimated that the Houston area experienced 1.3 m of rainfall, which is the largest amount recorded in a single storm. More than 780,000 Texans were evacuated, and there was no drinking water available for 61 communities during the five days of the hurricane (FEMA 2017a). During the storm, there were about 450,000 people who required disaster assistance and 30,000 needed to

be moved to shelters (Weather 2017). The hurricane left the state with 50 people dead and damage costs estimated to reach \$180 billion (Fortune 2017).

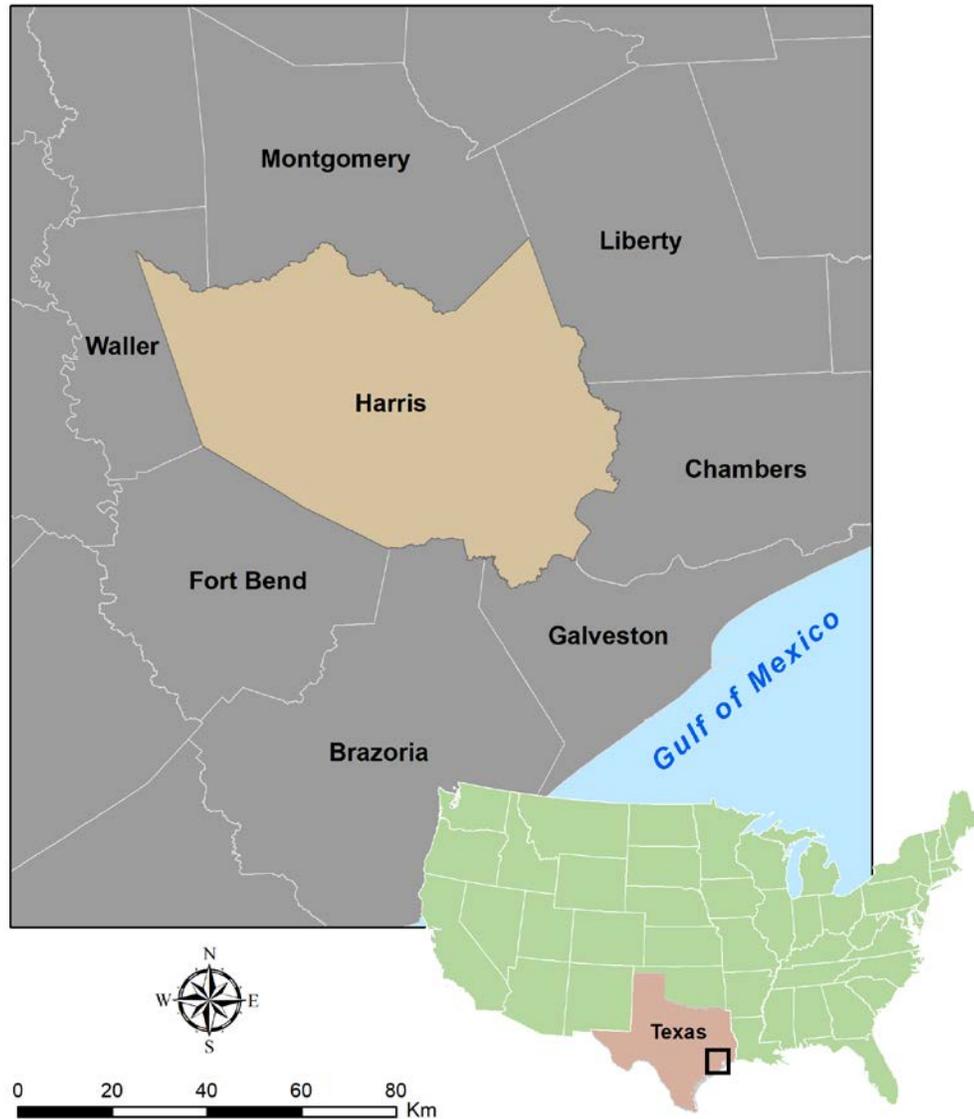


Figure 2. Map of The Study Area. (data source: TNRIS)

Data

As mentioned above, Harvey made landfall on the 25th of August and lasted for five days. Consequently, it is important to understand the social behavior and activity of the impacted population during and after the hurricane using Twitter data. Also, satellite

images to be used should cover during and post the flood event. Hence, the temporal period of the study was from August 25th to September 1st, 2017. The data were obtained from two major sources: non-authoritative (i.e. VGI and RS) and authoritative (i.e. stream gauges and inundation areas).

VGI Data

As described in the introduction, non-authoritative sources include social media content (text, pictures, and videos) and crowdsourced content (pictures and videos). Based on the timeframe of the hurricane, Twitter data were provided by the Geography Department at Texas State University. Then, tweets with geotagged attributes (hereafter tweets) were separated and used for the analysis. The tweets combined both text and multimedia information and part of it was limited to text only. Crowdsourced geotagged pictures were collected from the National Alliance for Public Safety GIS (NAPSG) Harvey picture-sharing platform (NAPSG 2017). This platform allows the users to upload pictures associated with a geographic location regarding the disaster situation in the impacted areas. Finally, crowdsourced geotagged videos were obtained from Homeland Infrastructure Foundation-Level Data (HIFLD 2017).

Authoritative and RS Data

In this study, the inundation maps derived from authoritative sources included the depth grids modeled by FEMA (2017b). Furthermore, water level records collected by the stream gauges from the USGS National Water Information System (USGS 2018b) were used to validate VGI WD data. In addition, satellite images data were collected from the USGS EarthExplorer (USGS 2018a) for VGI WD validation. The obtained satellite images included SWIR band capabilities. The obtained post-disaster image

product with such spectral band was Sentinel-2 collected on September 1st. The selected water index to extract and map the WD was the MNDWI. To be able to synthesize WD from the RS images, a 1.5 m DEM derived from lidar data from Texas Natural Resources Information System (TNRIS 2008) was used. Table (2) illustrates the data used and its sources.

Table 2. Data Summary.

Data	Details	Source
Twitter	<ul style="list-style-type: none"> – Text tweets – Pictures – Videos 	Geography Department – Texas State University
Crowdsourced	<ul style="list-style-type: none"> – Pictures – Videos 	(NAPSG 2017) and (HIFLD 2017)
Stream gauges	<ul style="list-style-type: none"> – Water level 	(USGS 2018b)
Flooded areas	<ul style="list-style-type: none"> – Modeled depth grids 	(FEMA 2017b)
Satellite images	<ul style="list-style-type: none"> – Sentinel-2 	(USGS 2018a)
DEM	<ul style="list-style-type: none"> – 1.5 m DEM derived from lidar 	(TNRIS 2008)

Preprocessing

As discussed above, Twitter data were obtained by the Geography Department, Texas State University. The tweets were collected using Twitter streaming application programming interface (API) in the form of JavaScript Object Notation (JSON) using a bounding box with coordinates that includes Houston area. Then, specific fields (e.g., lat, long, and URL) from the raw data were printed to Comma Separated Values (CSV) file format to be imported into spreadsheets or databases. After obtaining the tweets, selected tweets with information relevant to the hurricane or floods were queried using hashtags and keywords related to the event (Table 3). This study adopted a similar strategy as

previous studies (e.g. Li et al. 2017) by using hashtags or keywords including the name of the event (e.g. Harvey), multiple descriptions of the event (e.g. floods, water, etc.) or using most common hashtags during the event (e.g. based on the trending ones suggested by hashtagify.me) to compile the list of hashtags and keywords used (Table 3). For temporal classification and analysis, the time of the tweets was adjusted from Coordinated Universal Time (UTC) to Central Daylight Time (CDT) UTC-5.

Table 3. List of Hashtags and Keywords Used for Relevant Tweets Classification.

Keywords	usgs, tropical, emergency, high, height, heavy, close, flood, rescue, response, underwater, water, harvey, hurricane, safe, feet, ft, foot, inch, inches, drowned, submerged, overflow, rain, damage, storm, tornado, disaster
Hashtags	#HurricaneHarvey, #Harvey, #Flood, #Floods, #Storm, #Rain, #Damage, #Tropical, #Hurricane, #Water, #Disaster, #Emergency, #Underwater

Then, the relevant tweets were categorized into two groups—those with hyperlinks (multimedia) or those without (text only). Next, the hyperlink group was categorized into two categories— pictures or videos. Afterward, both categories—those with hyperlinks and those without—were classified based on the date of the tweet, after being adjusted to the local time, to generate multiple WD surfaces on a daily basis. Similarly, crowdsourced pictures and videos were grouped based on their dates. The authoritative data from FEMA were in the form of raster depth grids in feet and were converted to meters. The stream gauge data included water level data for the selected

time periods of the tweets, and the water records were used to generate a daily average water level. Then, the daily averages were combined with VGI synthesized WD values for an overall comparison by extracting the interpolated VGI WD values to the gauges points. Regarding RS data, the required spectral bands, b3 and b11, were atmospherically corrected to be used in the MNDWI equation. Then, multiple scenes for each of the two bands were mosaicked and clipped into one image and to match the administrative boundaries of the study area. Finally, the 1.5 m DEM generated and provided by TNIRIS (2008) was used for RS WD extraction.

WD Extraction

Using the text category of tweets (without hyperlinks), text information indicating water level, either explicitly or by description was identified based on the listed hashtags and keywords in Table 2. An example of explicit water level tweets would be those that specify a water depth of 3.5 ft or 40 cm, whereas a tweet with a description would be one indicating the water level relative to an object, e.g. (“my car is flooded!!” or “the sidewalk is covered with water”). An example of water level information indicated in the text is illustrated in Figure 3.

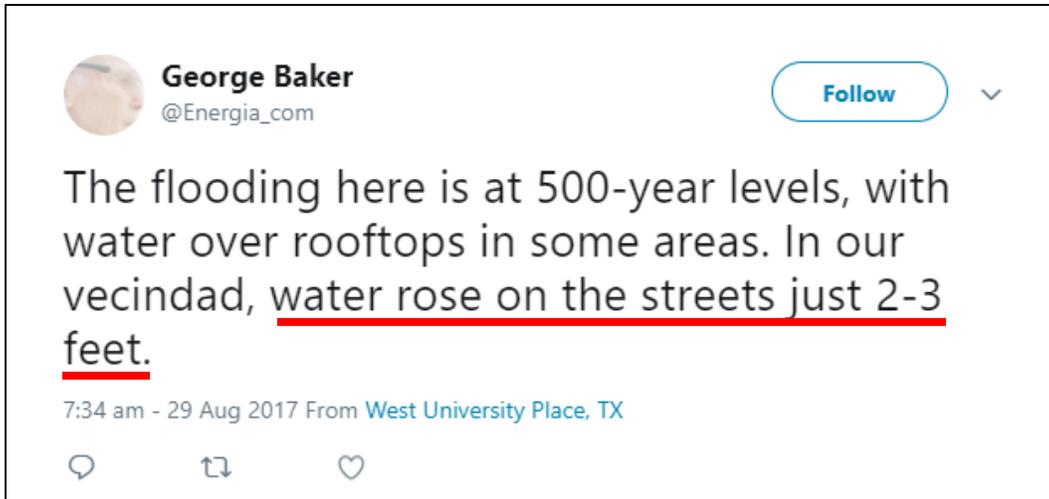


Figure 3. An Example of Water Level Indicated in a Text VGI.

For the pictures and videos in the multimedia category, the water level was extracted visually by observing the surrounding objects, such as trees, sidewalk curbs, buildings, or cars, and estimated the height of water covering these objects. A similar approach will be followed for the crowdsourced pictures and videos. Table 4 lists the approximate height of common objects used for water extraction in this study. An example of water level extraction from multimedia is demonstrated in Figure 4, which shows that the water is slightly covering the sidewalk curbs. Therefore, the water level was estimated to be 0.3 m for this picture.



Figure 4. An Example of Extracting Water Level from a Picture Using the Surrounding Physical Objects. The Red Line Shows the Edge Between the Road and the Sidewalk.

Table 4. Height of Physical Objects for Multimedia Water Level Extraction.

Object	Approximate height in meters
Traffic marks, e.g. stop sign	2 - 2.2
Traffic signals	5.5 - 6
Sidewalk curbs	0.2 – 0.3
Fire hydrants	0.5 - 0.7
Street bin	0.8 – 1
Mailbox	1 – 1.2
One story of a building	2.5 - 3

It's important to note that some tweets might include multimedia and text content, combined in one tweet, indicating water level information. An example of such tweets is shown in Figure 5. Tweets with such content were treated as both text and multimedia as a water level information source from VGI.

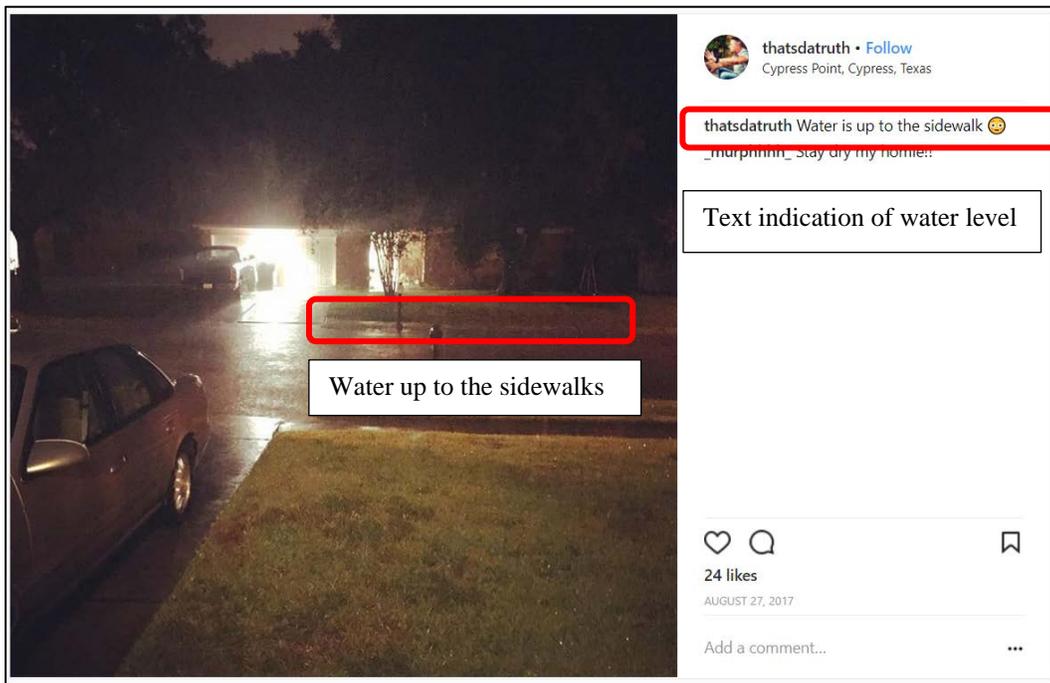


Figure 5. An Example of Water Level Information Included in both Text and Multimedia for a Single Post.

Finally, kriging spatial interpolation was used to generate a WD for each type of VGI (WD_{text} , WD_{pictures} , and WD_{videos}) and generate an overall WD of the VGI (WD_{VGI}). Kriging was used because of its application for surface water level estimation in previous studies (e.g. (Zrinji and Burn 1992)). The interpolation is done because of the difficulties to find multiple coincided VGI modalities points at the exact location to perform the comparison, or to find the same location of both VGI and authoritative data, e.g. stream gauges. Then, the WD_{VGI} was used for validation against authoritative data. The overall VGI workflow is demonstrated in Figure 6.

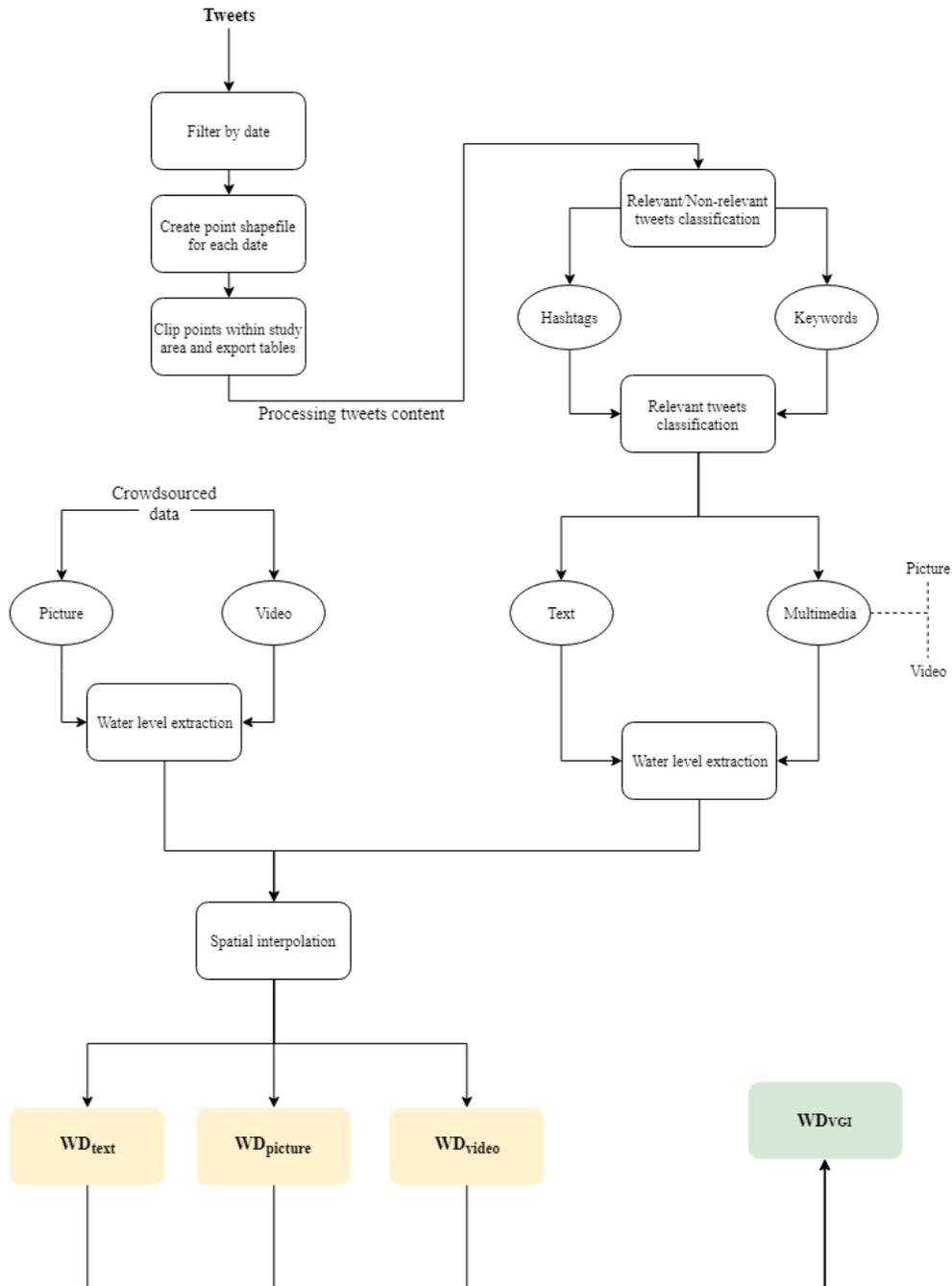


Figure 6. Twitter and Crowdsourced Data Processing and WD Extraction.

To extract water bodies from RS data, the MNDWI model was used to extract water features. Since band3 (green) have a higher spatial resolution (10 m) compared to band11 (SWIR) with 20 m, band3 was resampled to match the spatial resolution of

band11 to be used in the MNDWI index. The output of the MNDWI was a raster surface with values ranging between -1 and 1. To determine the water pixels from non-water pixels, a threshold (≥ 0.10) was selected, based on visual observation, to delineate the water bodies. After extracting and delineating water bodies, the water bodies spatial resolution was resampled to match the DEM resolution (1.5 m) and was used as zones to extract the maximum elevation for each zone from the DEM using zonal statistics. Then, the value of each DEM pixel within the water body zone was subtracted from the maximum value ($MAX - DEM$) extracted from the earlier step to get the water depth at the pixel location using local statistics. The WD extraction and the overall RS workflow are illustrated in Figures 7 and 8 respectively.

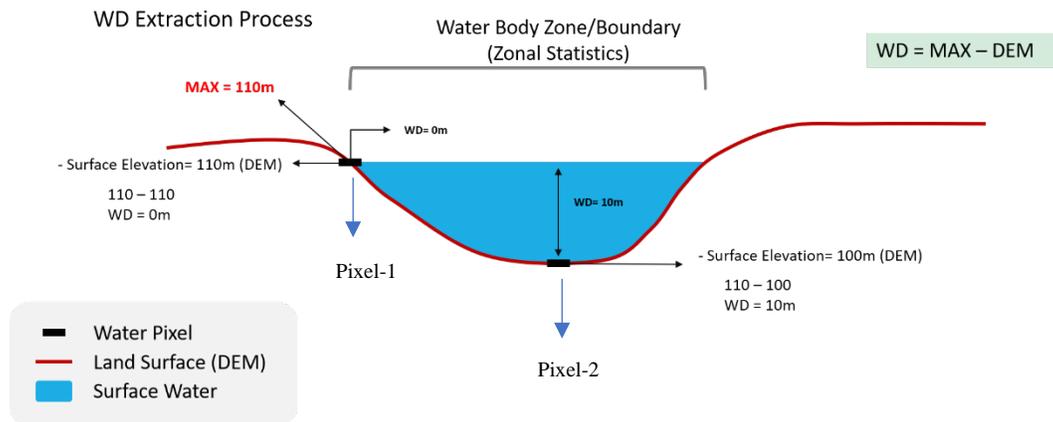


Figure 7. WD Extraction from RS Water Bodies. First, The Maximum Elevation Is Extracted From the Water Body Zone (In This Example, the Maximum Was 110 m). Then, the Value of a DEM Pixel is Subtracted from the Maximum of the Zone (In This Example, Pixel-1 $WD = 0$ Since the Maximum is Equal to the DEM Pixel Value. While Pixel-2 Water Depth = 10 m Since the DEM Value of Pixel-2 = 100 m).

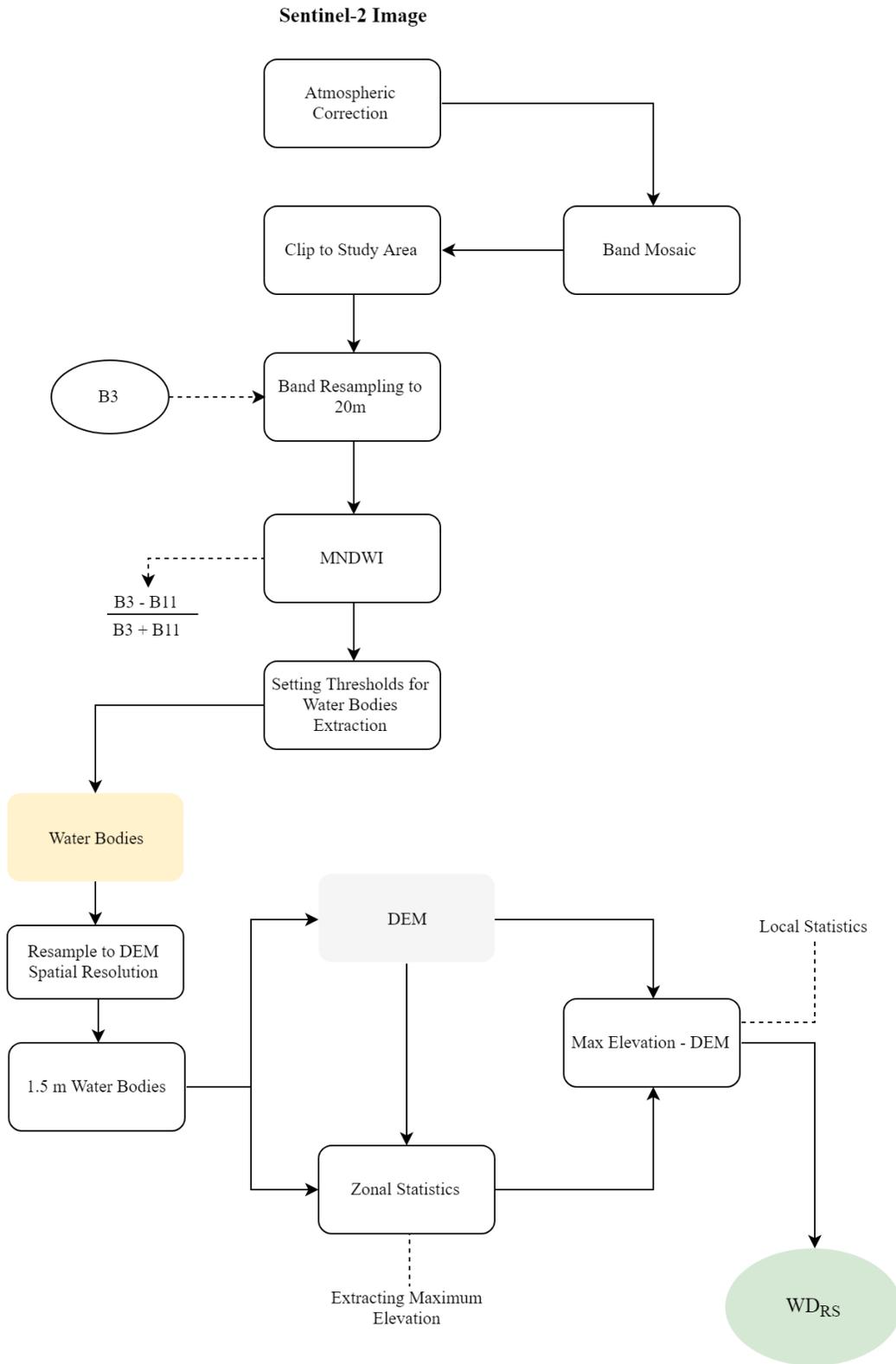


Figure 8. Remote Sensing Processing and WD Extraction.

The overall research design is illustrated in Figure 9.

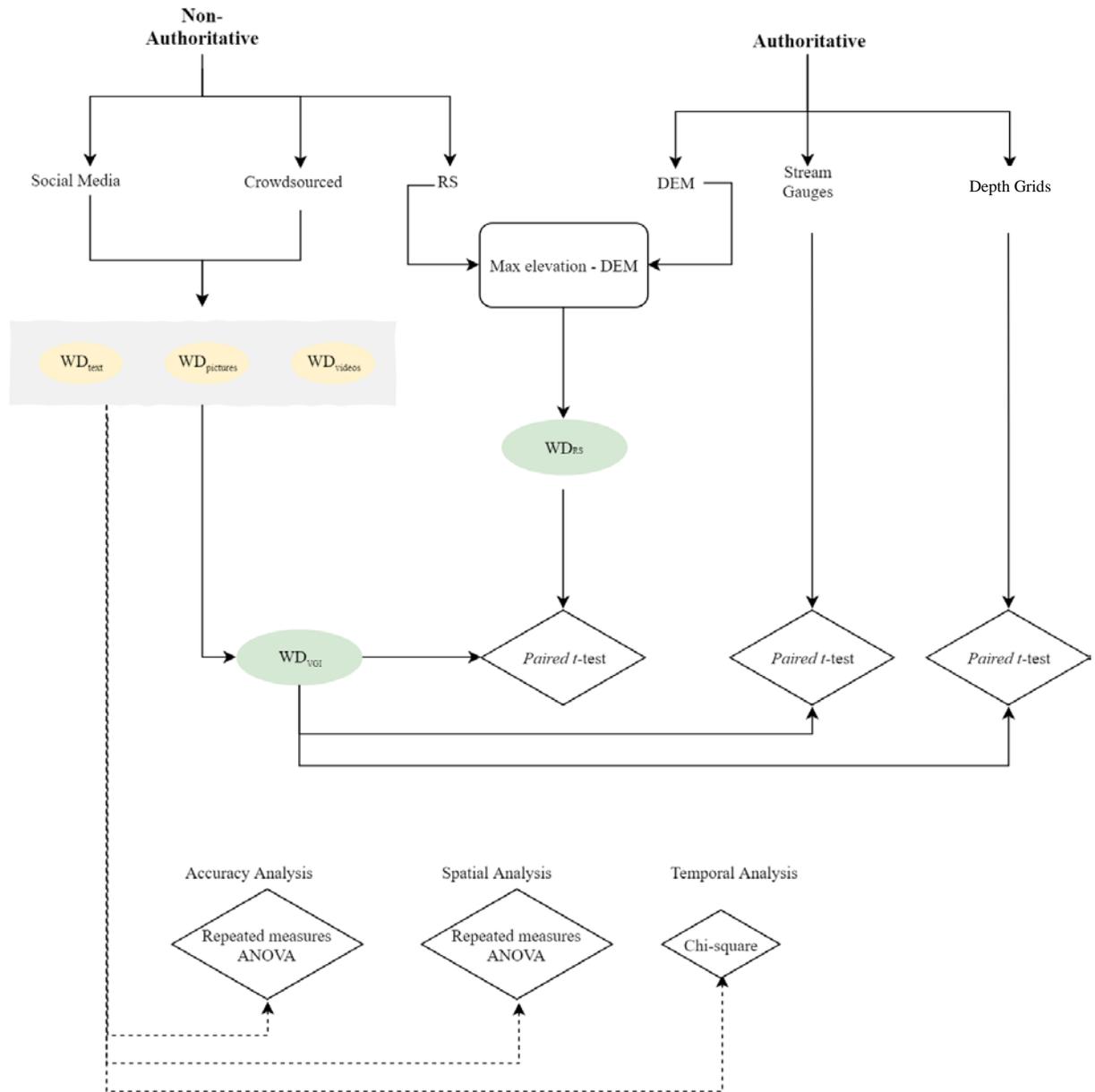


Figure 9. Overall Research Design.

WD Comparison

To compare between the three VGI WDs, a repeated-measures ANOVA test was performed to examine any significant differences among the VGI data modalities regarding precision and spatial differences. For the temporal differences among the VGI WDs, Chi-square test was conducted.

- Null Hypothesis 1a: $H_{1a}: WD_{\text{text}} = WD_{\text{pictures}} = WD_{\text{videos}}$
- Alternative Hypothesis: $WD_{\text{text}} \neq WD_{\text{pictures}} \neq WD_{\text{videos}}$

Where WD = water depth.

- Null Hypothesis 1b: $H_{1b}: KD_{\text{text}} = KD_{\text{pictures}} = KD_{\text{videos}}$
- Alternative Hypothesis: $KD_{\text{text}} \neq KD_{\text{pictures}} \neq KD_{\text{videos}}$

Where KD = kernel density.

- Null Hypothesis 1c: $H_{1c}: \text{COUNT}_{\text{text}.t} = \text{COUNT}_{\text{pictures}.t} = \text{COUNT}_{\text{videos}.t}$
- Alternative Hypothesis: $\text{COUNT}_{\text{text}.t} \neq \text{COUNT}_{\text{pictures}.t} \neq \text{COUNT}_{\text{videos}.t}$

Where COUNT = count of flood-related tweets for a specific data modality per day [t = 8/25-9/1].

Similarly, the synthesized VGI WD was compared with the RS WD with the following null hypothesis by using the paired *t*-test:

- Null Hypothesis 2: $H_2: WD_{\text{VGI}} = WD_{\text{RS}}$
- Alternative hypothesis: $WD_{\text{VGI}} \neq WD_{\text{RS}}$

To validate the quality of VGI WD, null hypotheses 3 investigated any significant difference between the FEMA depth grids and the USGS stream gauge records using a paired *t*-test on a daily basis.

- Null Hypothesis 3a: $H_{3a}: WD_{\text{VGI}} = WD_{\text{USGS}}$

- Alternative hypothesis: $WD_{VGI} \neq WD_{USGS}$
- Null Hypothesis 3b: $H_{3b}: WD_{VGI} = WD_{FEMA}$
- Alternative hypothesis: $WD_{VGI} \neq WD_{FEMA}$

IV. RESULTS

A summary of the VGI data collected and the output of VGI classification and water observations extraction is first presented. Then, the results of the VGI validation and comparison against other WD sources are demonstrated in three parts. First, the differences among VGI data modalities regarding precision, spatial, and temporal characteristics are demonstrated. Second, the differences between synthesized WD from VGI and RS synthesized WD are presented. Finally, the differences between the synthesized WD from VGI and two authoritative data sources, the USGS stream gauges and the depth grids from FEMA are illustrated.

VGI Summary

VGI data collection period was from August 25th to September 1st, 2017. The VGI data collected included both Twitter and crowdsourcing. Figure 10 illustrates the summary of tweets relevant to the hurricane event after filtering the tweets to relevant and non-relevant using the hashtags and keywords mentioned in table 3. The total count of tweets for the entire period in the study area was 32,586, and the total relevant tweets count was 9,578, which represents 29.4% of the total tweets. About 76.5% of relevant tweets occurred between 25th and 29th, which represents the period the hurricane lasted in the study area.

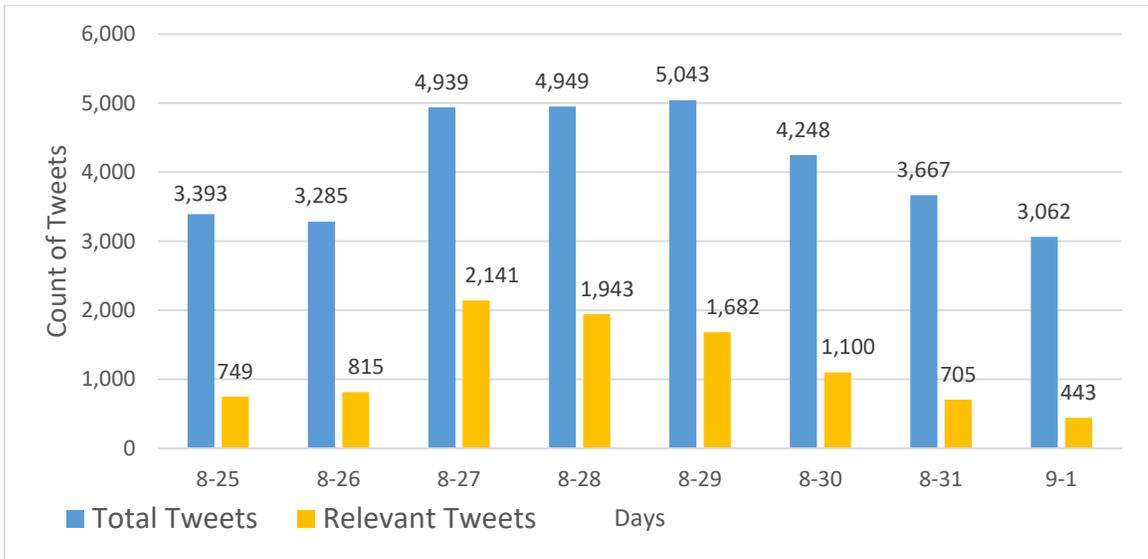


Figure 10. Summary of Relevant and Non-Relevant Tweets in Harris County.

The second step of tweets processing was classifying the tweets to either a text or multimedia (picture or video) type. Figure 11 summarizes the count of relevant tweets for both categories. The majority of tweets were classified as a multimedia type compared to the count of text type.

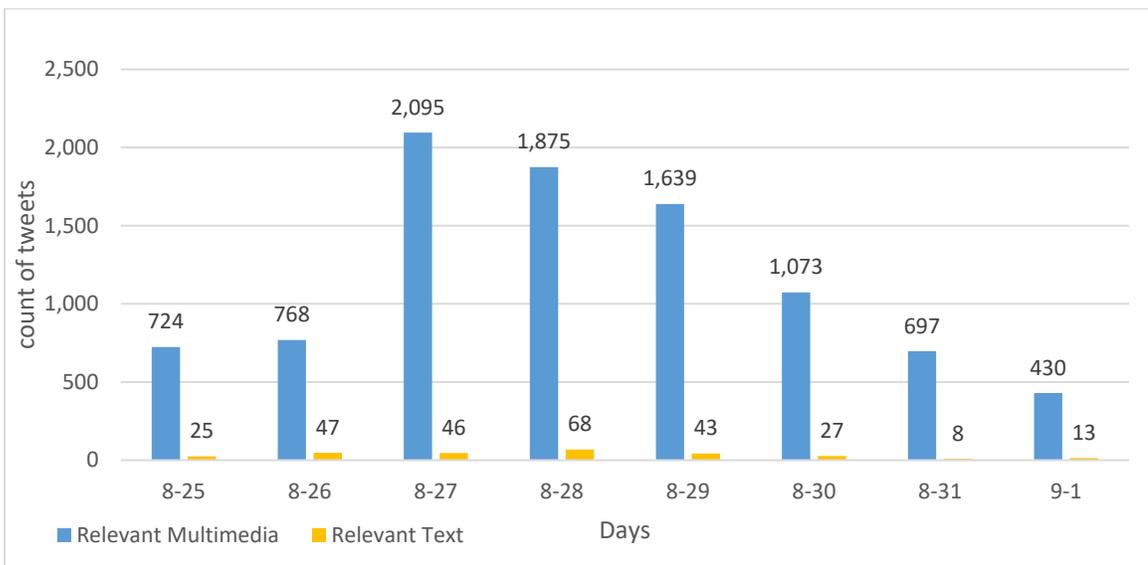


Figure 11. Summary of Relevant Tweets for Text and Multimedia.

After manual observation for each relevant tweet for both categories, the water level indicated by the text content or captured as multimedia was extracted. It's important to note that in some cases, different tweets from different users posted at various times were sharing the exact geographic location. The mean water level of such tweets was calculated and used as the single value to represent the duplicated tweets location, while the duplicated tweets were removed from the analysis. Figure 12 illustrates the summary of tweets with water level information compared to relevant tweets on a daily basis. The total count of water depth information extracted from Twitter, regardless of data modalities, was 434 observations, representing about 4.5% of the total relevant tweets. The count of tweets with water depth information increases during the mid-way of the hurricane, between August 27th and August 29th, and decrease afterwards. The majority of the relevant tweets does not include water depth information. Regarding crowdsourced data, a total of 152 crowdsourced multimedia shared (1 for text, 17 for videos and 134 for pictures) within Harris County.

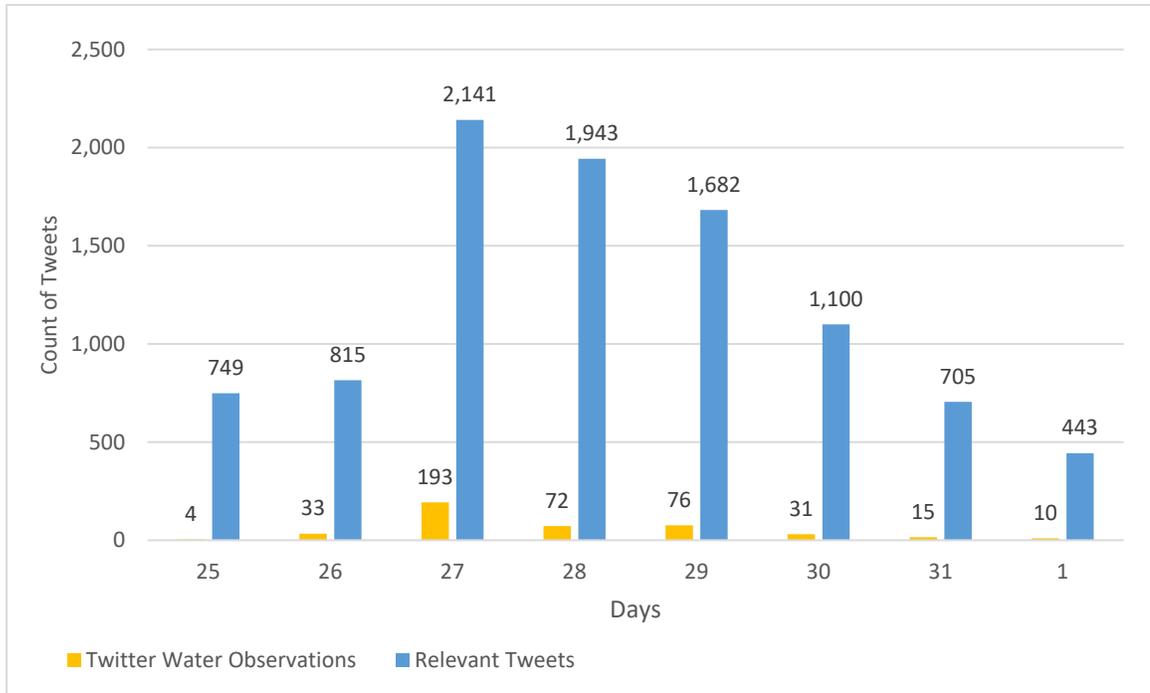


Figure 12. Summary of Tweets with Water Depth Information Compared to Relevant Tweets on a Daily Basis.

After classifying the tweets to text and multimedia (picture or video) and extracting water level from both tweets and crowdsourced data, Figure 13 summarizes the variation in water level extracted from the three VGI data modalities (text, picture, and video) on a daily basis. Out of the total 586 observations (for both tweets and crowdsourced), pictures had the most water level observations contribution with 350 observations representing 59.6% of the total water depth observations, and text shared the least contribution with 30 observations representing 5% of the VGI data. After manual extraction of water level from crowdsourced data, the total observations with water level extracted was 152 (134 for pictures, 18 for videos, and one text). The geographic distribution of both VGI sources, social media and crowdsourced, is displayed on the map in Figure 14.

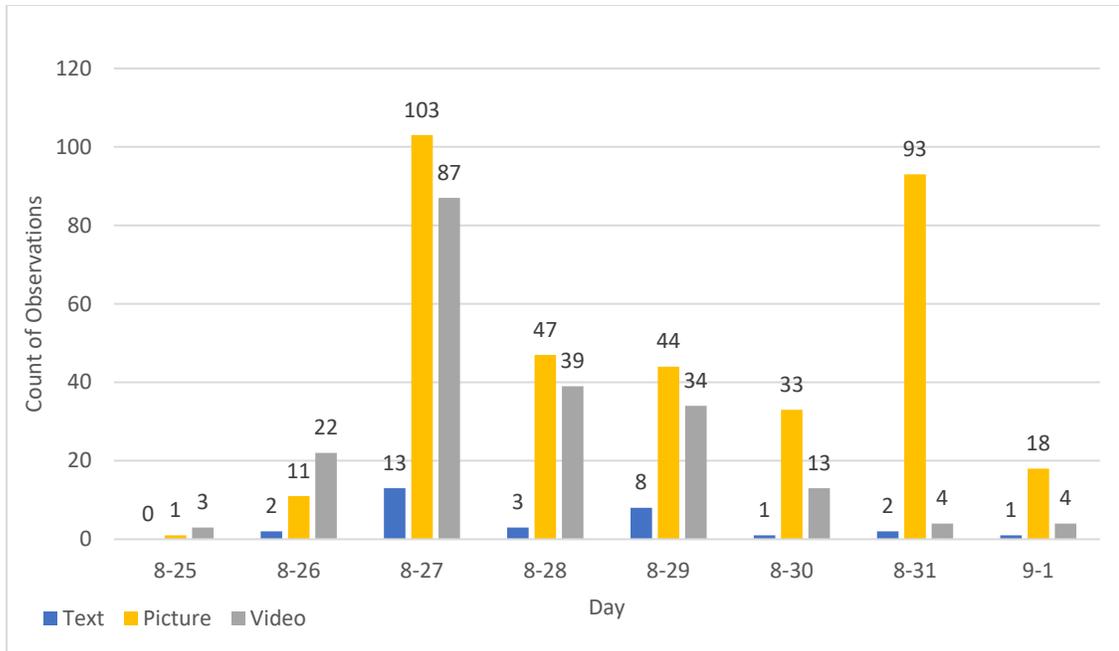


Figure 13. Summary of the Variation in Water Depth Extracted from the Three VGI Data Modalities (Text, Picture, and Video) on a Daily Basis.

The distribution of the points, in general, is scattered and tends to be closer to Downtown Houston, south and north of the center as well, and there is less distribution of VGI points towards the east, while more observations are distributed towards the west compared to the east (Figure 14).

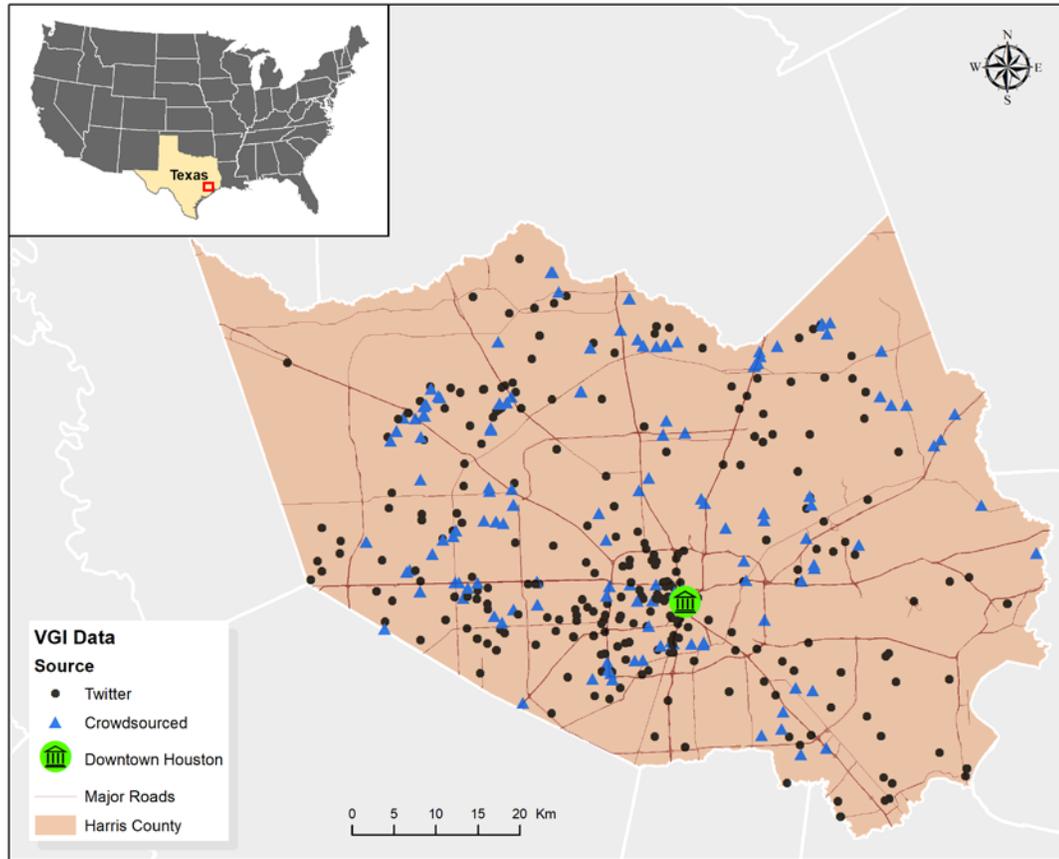


Figure 14. Geographic Distribution of both VGI Sources, Social Media and Crowdsourcing in Harris County.

Regarding the geographic distribution of VGI data modalities, pictures and videos had a similar pattern of the overall VGI points distribution. In general, the text showed more points in the center of the study area and limited presence in the east and west of Harris County (Figure 15).

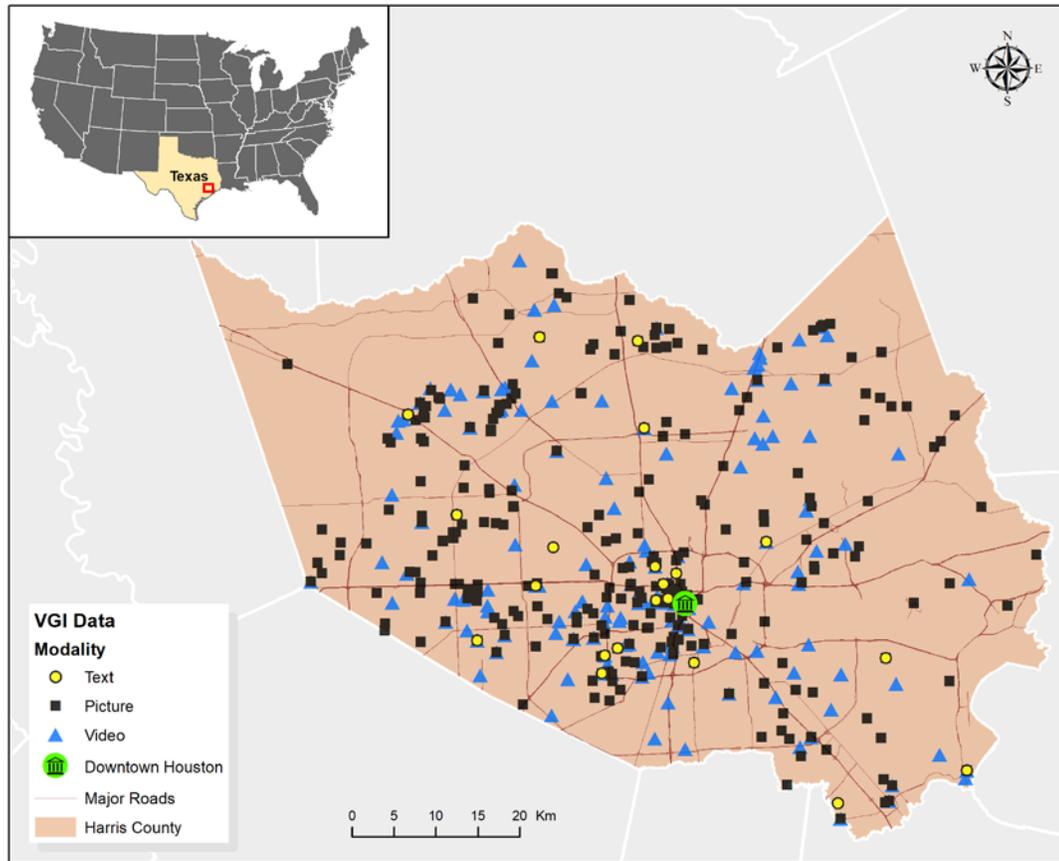


Figure 15. Geographic Distribution of VGI Data Modalities in Harris County.

VGI Validation and Comparison

Precision of VGI Data Modalities

A Shapiro-Wilk test for normality was conducted and showed a significant difference for all the modalities at $p < 0.001$ for text and pictures, and $p = 0.003$ for videos). Since the data are not normally distributed, a non-parametric test was conducted. The results of Friedman test ($\chi^2 = 5.626$; $p = 0.060$; $n = 32$) failed slightly short of a significant difference in precision across VGI data modalities. Therefore, the null hypothesis 1a was accepted. However, the non-parametric Wilcoxon post-hoc test showed a significant difference between two pairs: (1) text and pictures ($Z = -2.565$; $p =$

0.010; $n = 32$) and (2) text and videos ($Z = -2.463$; $p = 0.014$; $n = 32$). There was no significant difference between the third pair (pictures and videos) ($Z = -0.587$; $p = 0.557$, $n = 32$).

Spatial Distribution of VGI Data Modalities

A kernel density surface was derived for each of the three VGI data modalities, and each raster cell was converted into points with kernel density extracted for further statistical analysis. Similarly, a normality test showed that the kernel densities were not normally distributed at $p < 0.001$. The resulting Friedman test ($\chi^2 = 6842186$; $p < 0.001$; $n = 5113697$) revealed a significant difference in spatial distribution across VGI data modalities at. Therefore, the null hypothesis 1b was rejected. A non-parametric Wilcoxon post-hoc test showed that there was a significant difference among all pairs of data modalities: (1) text and pictures ($Z = -1,873.760$; $p < 0.001$; $n = 5,113,697$), (2) text and videos ($Z = -1,704.997$; $p < 0.001$; $n = 5,113,697$), and (3) pictures and videos ($Z = -1317.366$; $p < 0.001$; $n = 5,113,697$).

Temporal Distribution of VGI Data Modalities

A chi-square test was conducted to examine the differences in the temporal pattern of the VGI data modalities throughout the study time frame. Table 9 includes the current and expected distribution of observations. It was found that there is a statistically significant difference between the current and the expected temporal distribution of the data ($\chi^2 = 88.14$; $df = 14$; $p < 0.001$). Therefore, the null hypothesis 1c was rejected.

Table 5. Current and Expected Distribution of VGI Observations for The Chi-Square Test.

Current									
	Day								
	8-25	8-26	8-27	8-28	8-29	8-30	8-31	9-1	Total
Pictures	1	11	103	47	44	33	93	18	350
Text	0	2	13	3	8	1	2	1	30
Video	3	22	87	39	34	13	4	4	206
Total	4	35	203	89	86	47	99	23	586
Expected									
	Day								
	8-25	8-26	8-27	8-28	8-29	8-30	8-31	9-1	Total
Pictures	2	21	121	53	51	28	59	14	350
Text	0	2	10	5	4	2	5	1	30
Video	1	12	71	31	30	17	35	8	206
Total	4	35	203	89	86	47	99	23	586

VGI and RS Validation

After extracting WD from the RS image (RS WD) and interpolating the VGI observations using kriging (VGI WD), the RS WD raster layer was converted to points. Then, the points were used to extract the interpolated WD values from the VGI (Figure 16) for comparison. The VGI WD included observations on August 31st and September 1st, due to the lack of observations on September 1st. due to the large sample size ($n = 37,967,601$), a paired t -test was applied to examine any differences between WD derived from RS and VGI synthesized WD. The results showed that there was a significant difference between VGI and RS water depth ($t = 264.232$; $df = 38067599$; $p < 0.001$). Therefore, the null hypothesis 2 was rejected.

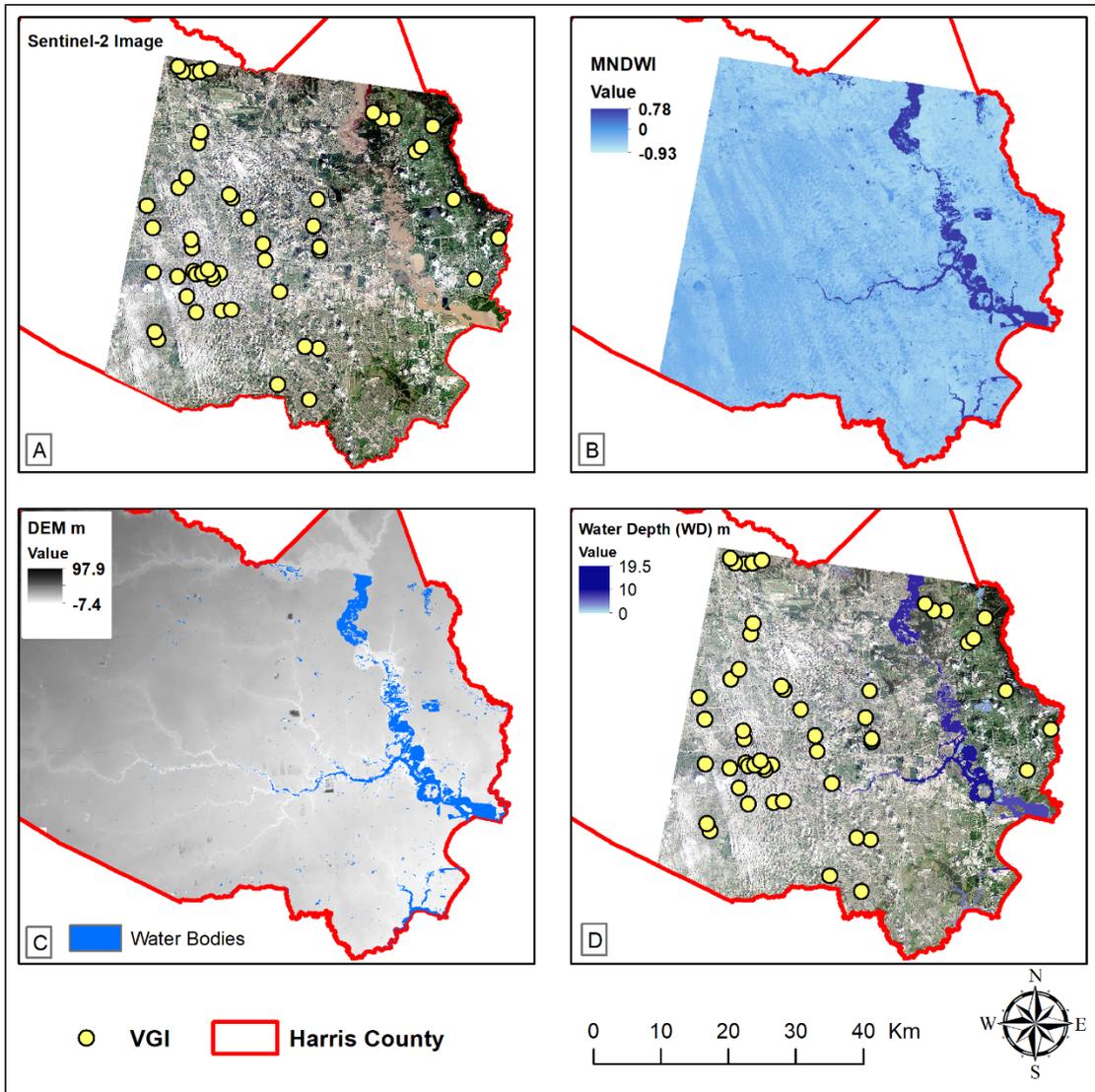


Figure 16. WD Extraction from RS Images. (A) Sentinel-2 True Color Image, (B) The Output of the MNDWI, (C) Water Bodies Delineated after Applying the Threshold (≥ 0.10), and (D) The Final WD Derived from RS Image.

VGI and Authoritative Data Comparison

VGI and USGS Stream Gauges

The USGS stream gauges scatter across the study area (Figure 17). The readings for each gauge are taken at a 15-minute interval with a total of 96 water level readings per day. The total count of gauges within the county was 55, and eight of the total gauges does not have readings for the entire study period. The remaining gauges with water depth observations used in this study were 47 gauges. Some of the gauges had missing readings on a specific day or did not collect any water level readings for the entire day. Thus, such gauges with no observations for the entire day or with few observations to be analyzed were eliminated from the analysis for that specific day. The daily average water level was calculated for each gauge and was used for statistical analysis.

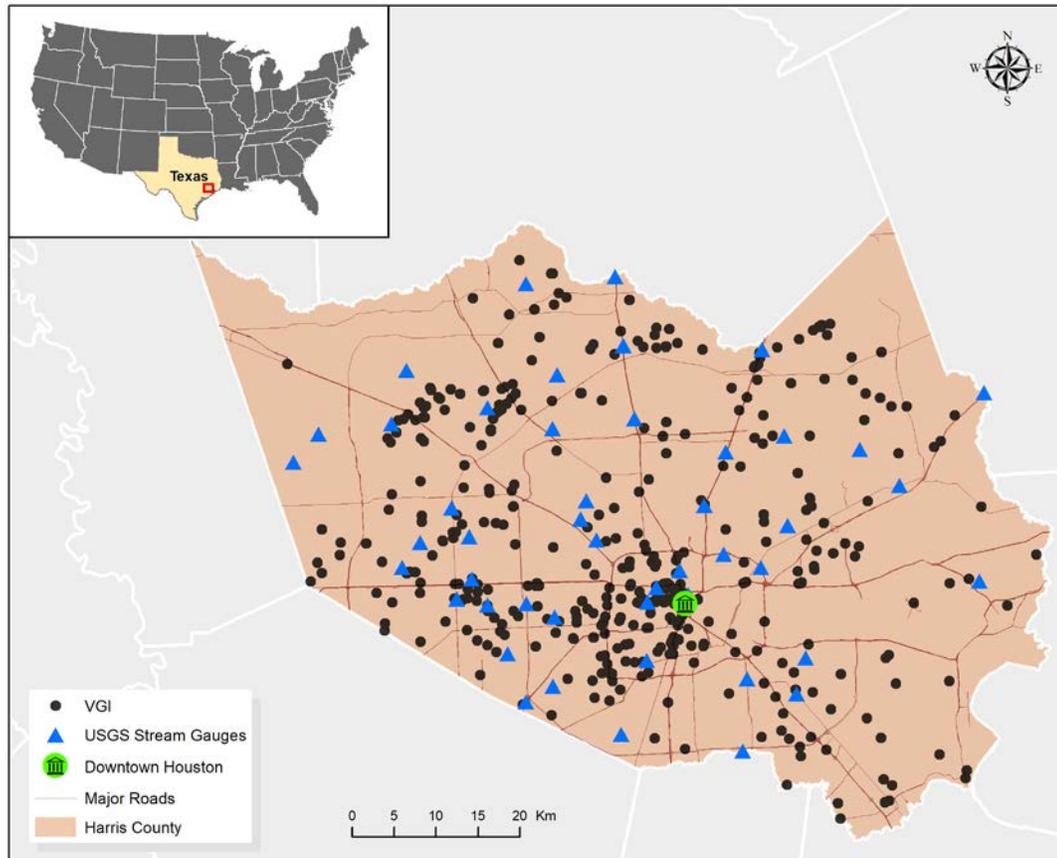


Figure 17. Geographic Distribution of USGS Stream Gauges in Harris County.

In order to compare VGI with stream gauge records, the VGI segregated per day was interpolated into a continuous surface using kriging. Then, the values of the kriging raster were extracted at the USGS gauge points for the same day. After that, a table with both values for the entire days was used to perform the normality test and the paired *t*-test. The analysis did not include August 25th since only four VGI observations were available.

The results showed that the gauges data showed a normal distribution at $p = 0.264$, while the VGI data was not normally distributed at $p < 0.001$. Therefore, a paired *t*-test was conducted along with an equivalent non-parametric test for verification purposes. The results showed no statistically significant difference between VGI and

USGS stream gauge water depth ($t = 1.513$; $df = 275$; $p = 0.131$), and the non-parametric test showed similar results ($Z = -1.652$; $p = 0.099$; $n = 276$). Therefore, the null hypothesis 3a was accepted.

VGI and FEMA Modeled Depth Grids

The modeled depth grids from FEMA were available for five days including 27th, 28th, 29th, 30th, and September 1st (Figure 18). The maximum extent of the inundated areas was during the 30th and September 1st (Figure 18 d and e).

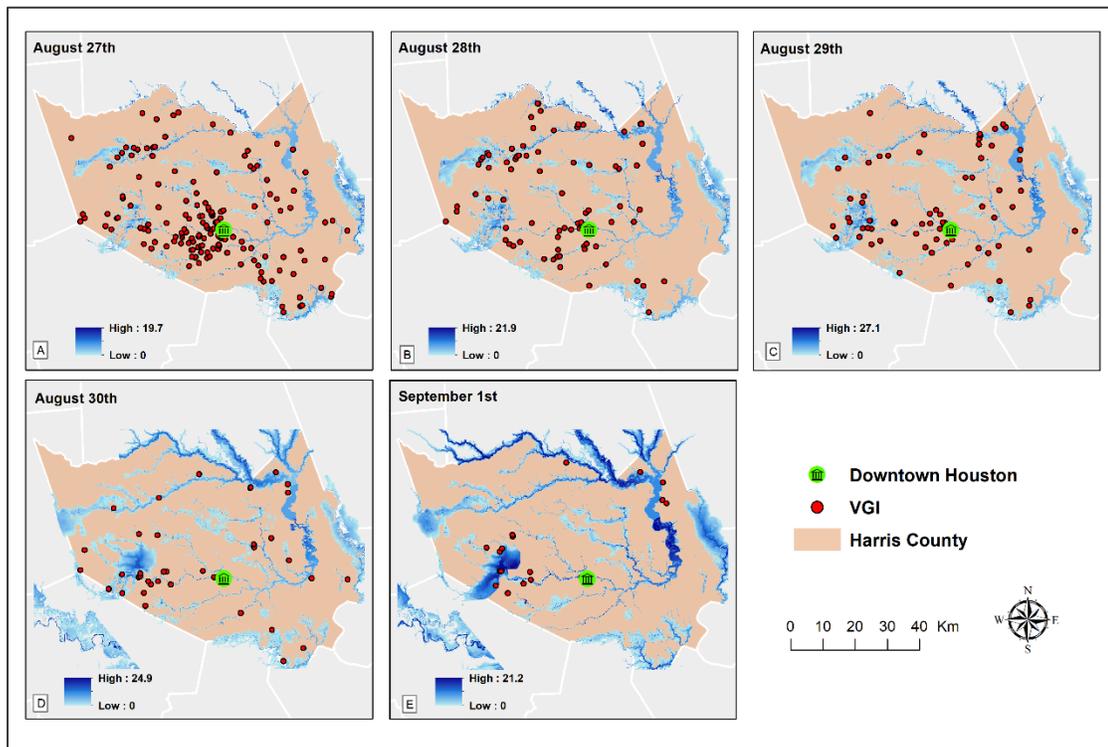


Figure 18. Modeled Depth Grids in Meters from FEMA. (a) August 27th, (b) August 28th, (c) August 29th, (d) August 30th, and (e) September 1st.

The VGI points were used to extract the water depth from FEMA grids for statistical analysis on a daily basis. The normality test showed a significant difference between all pairs during August 27, 28, and 29 at $p < 0.001$, and during September 1 at p

= 0.004 and at $p < 0.001$ for VGI and FEMA depth grids respectively. However, on August 30, VGI had a normal distribution at $p = 0.075$, while FEMA depth grids had no normal distribution at $p = 0.001$ on the same day. A non-parametric equivalent of paired t -test was conducted and the results showed that there was a statistically significant difference between VGI and FEMA depth grids for all the days: August 27 ($Z = -7.255$; $p < 0.001$; $n = 117$), August 28 ($Z = -6.462$; $p < 0.001$; $n = 65$), August 29 ($Z = -6.275$; $p < 0.001$; $n = 56$), August 30 ($Z = -4.782$; $p < 0.001$; $n = 31$), and September 1 ($Z = -4.382$; $p < 0.001$; $n = 25$). For verification purposes, a paired t -test was conducted for the data on August 30 and it showed similar results ($t = -7.824$; $df = 30$; $p < 0.001$). Therefore, the null hypothesis 3b was rejected.

V. DISCUSSION

The summary of VGI processing showed that the count of tweets increases during the mid-way of the event, which can be observed between 27th and 29th of August for both total and relevant tweets with high water depth observations, where the maximum was 5 m, 3 m, and 4.8 m during the 27th, 28th, and 29th respectively. Also, there was a high positive correlation between the relevant tweets and the average water depth from the gauges ($r = 0.86$) which indicates that people tweeted more when the magnitude of the event increased (Figure 19). This pattern agrees with the findings of Li et al. (2017) and Mandel et al. (2012).

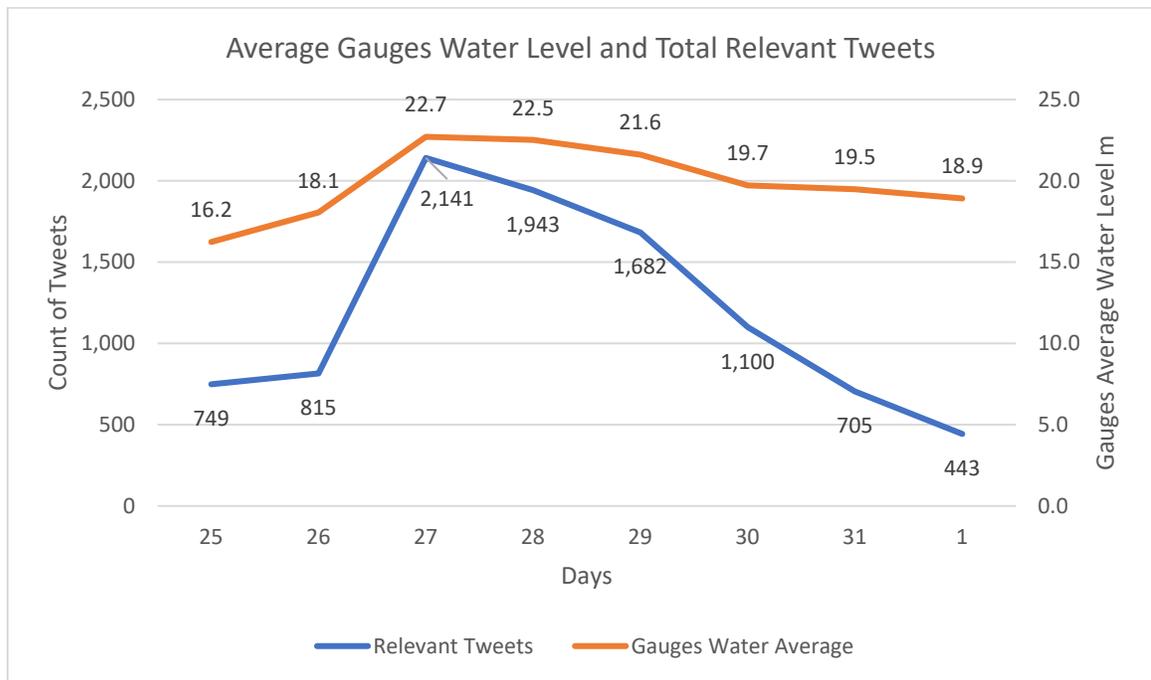


Figure 19. Daily Comparison Between Total Relevant Tweets and Average Water Level from the USGS Gauges.

In addition, 29.4% of all the tweets were classified as relevant, and 76.5% of these relevant tweets occurred during 25th to 29th of August when the hurricane lasted, which

indicates that the majority of tweets relevant to the disaster are more potentially available during the time of the disaster, more than pre or post the disaster.

Research Question Findings

Regarding the analysis of the first research question, the differences among VGI data modalities, the results showed that there were no significant differences of water depth across the three modalities in terms of precision, even though post hoc tests revealed that text modality was significantly different from the pictures and videos. Also, there was a significant difference in terms of spatial and temporal distribution of VGI modalities ($p < 0.001$ for both comparisons). This could be associated with the limited availability of text observations (5%) compared with the multimedia (about 95%), which was demonstrated by Li et al. (2017). This indicates that people tend to share multimedia content regarding their conditions and situation rather than describing it through text messaging.

Since picture and video modalities had no significant differences in terms of precision ($p = 0.557$) and the robust availability of multimedia during disasters, it should be taken under consideration to include VGI multimedia in big data studies because of its large sampling availability and its better context, compared with text. When conducting a damage assessment or rapid flood analysis, the user should consider that text modality requires additional text analysis (semantic and sentiment) to extract damage information or flood conditions. Moreover, the limited text modality sample size could reduce the quality of the assessment outputs and influences the decisions made upon such outputs. On the other hand, the large availability of multimedia could provide a broader insight for

the user in terms of rapid assessment of flood damage or risk with more spatial and temporal coverage.

Regarding the spatial characteristics, the results showed that all VGI data modalities have varying spatial distribution. This could be related to the limited count and distribution of text observations, where it was closer to Downtown Houston and limited towards to the fringe of Harris County, while pictures and videos had more scattered distribution in the study area, with more coverage and distribution of the former (Figure 20).

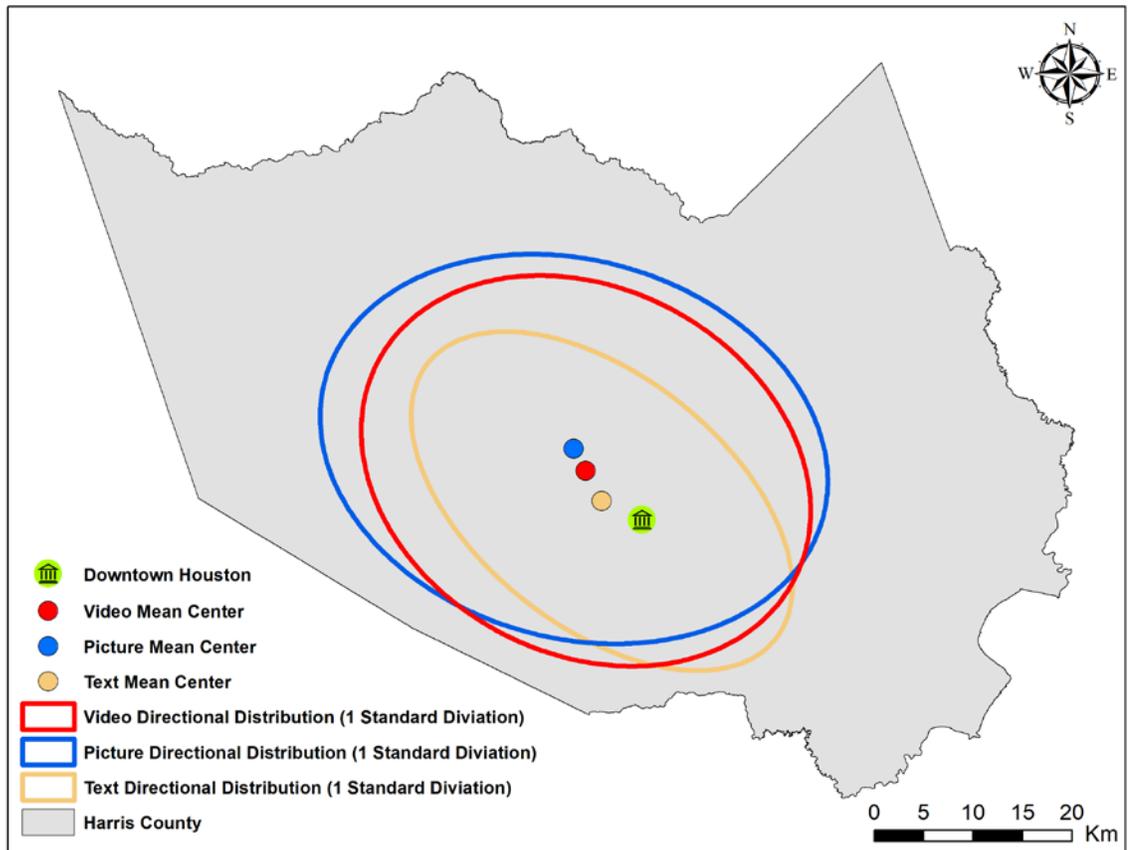


Figure 20. Mean Center and Directional Distribution of VGI Data Modalities in Harris County.

In addition, the water depth visualized in multimedia could be captured at locations with varying distances from the actual flooded area. For example, a multimedia could be taken in the third floor of a building showing a flood at the parking lot in front of the building, while other multimedia could show water depth up to the sidewalk across the street in a small neighborhood with a short distance from the actual location where the user was standing to take the picture or the video. However, text information is often assumed to indicate water depth at the same location of the tweet. Also, zooming could be a factor influencing the distance of measured water depth point from the actual multimedia location. Although pictures and videos had a visually similar distribution, the post-hoc test showed a significant difference between both data modalities. It could be related with the higher availability and the spatial distribution of picture points where it may influence the kernel density outputs, compared with the videos (Figure 21).

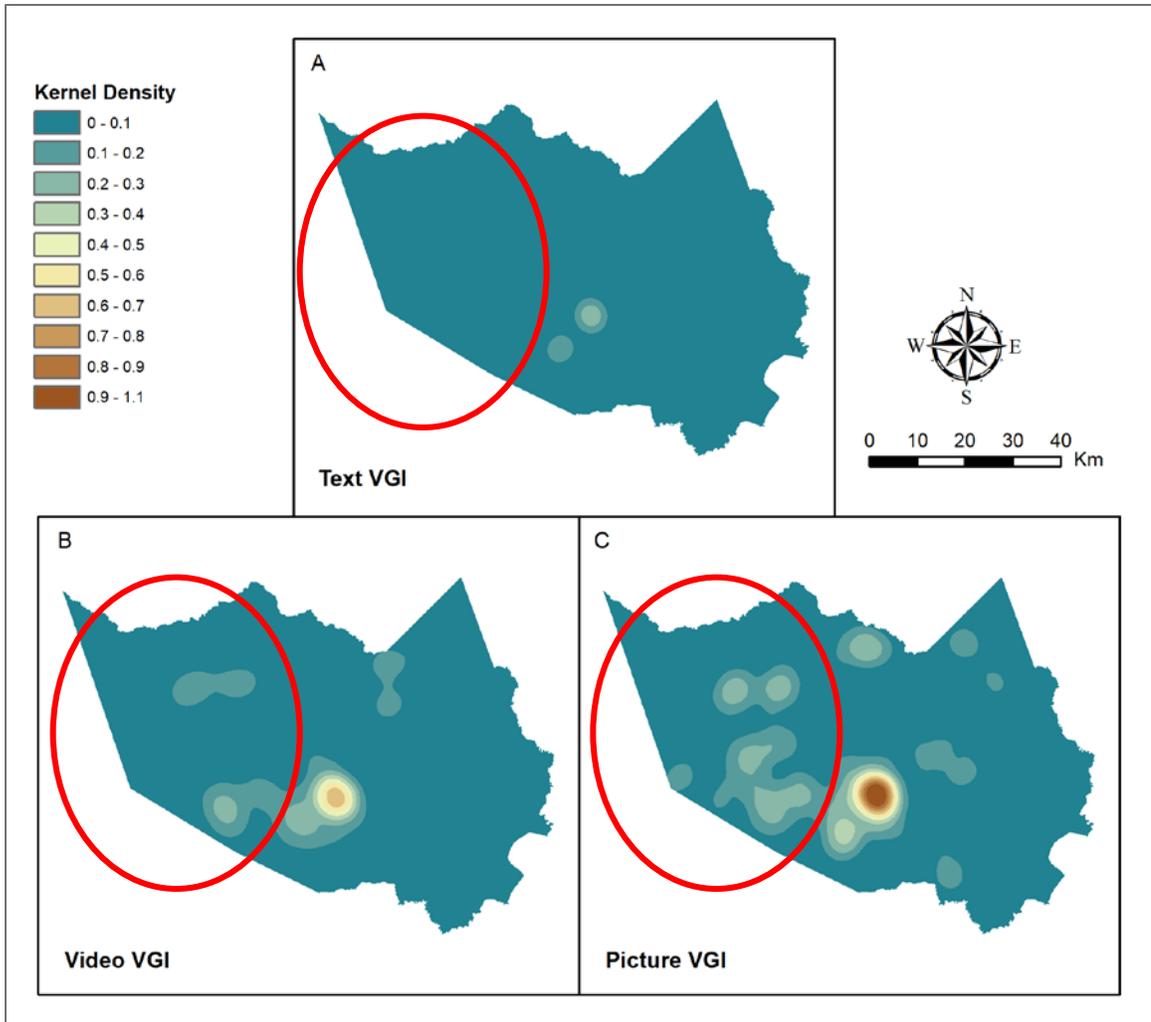


Figure 21. VGI Data Modalities Kernel Density. Due to the variation in sample size and distribution, the western areas of Harris county, highlighted in red, showed more picture modality density (c), compared with the video modality (b). With the small observation.

Another reason to explain the spatial variation of VGI could be associated with the spatial pattern of the digital divide, where population subgroups may have varying access to digital information and communication technology (ICT) (Riggins and Dewan 2005). Based on the digital divide index (DDI) calculated by Gallardo (2018), the DDI score is composed of two main components including infrastructure/adoption and socioeconomic characteristics. The DDI score ranged between 0–100 where high DDI

score indicates high existence of a digital divide and vice versa. Overall, there was a weak to low negative correlation coefficient between the DDI and the count of VGI points at each tract in the study area with -0.09, -0.20, and -0.16 for text, pictures, and videos respectively. Therefore, digital divide did not show significant influence on the spatial distribution of VGI data modalities (Figure 22).

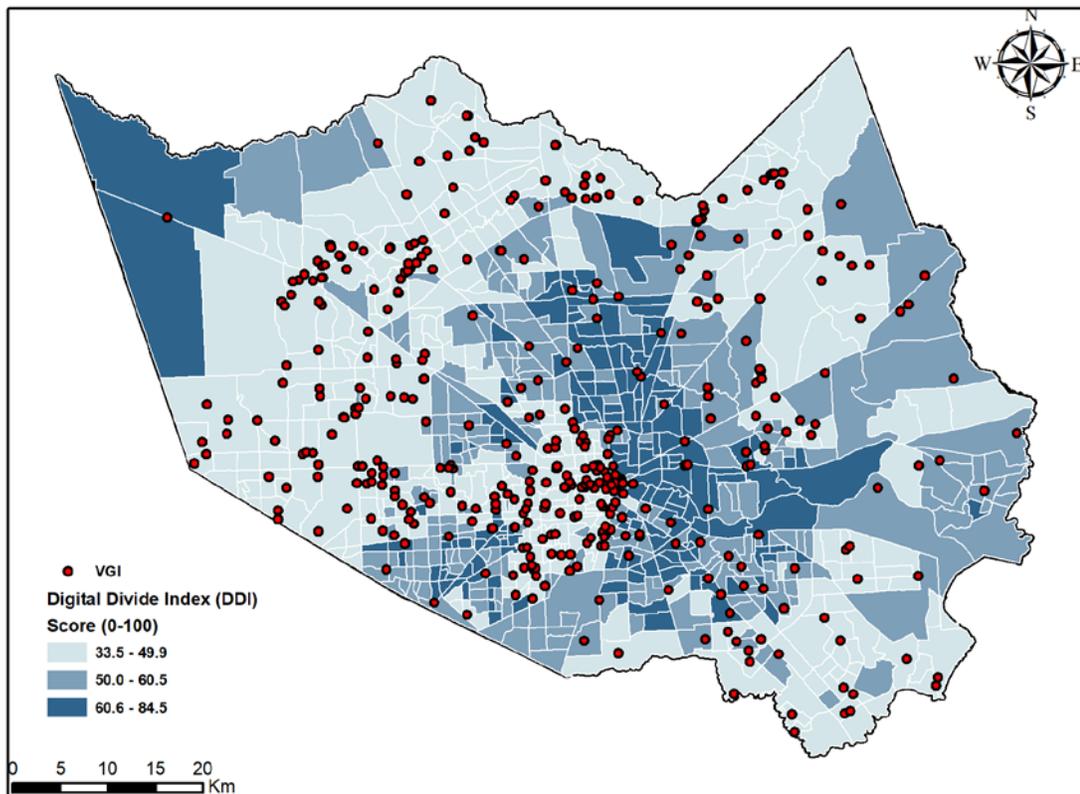


Figure 22. Digital Divide Index (DDI) in Harris County. The DDI Data were Obtained from (Mississippi State University 2018).

Finally, the temporal distribution of the data also varied significantly. Once again, the count of text observations may play a role in this variation. Furthermore, there is a possibility that users in some areas might have experienced technical issues with sharing tweets in real-time, such as power outage or temporarily limited internet access, that

prevented them from sharing their situation during the event. According to power outage reports between August 27th and September 1st, about 71.2% of VGI observations were in areas with up to 5% of power outage, while 22.8% of the observations were in areas between 5% and 20% of power outage, which shows an inverse relationship between the count of tweets and the percentage of power outage in the study area (Figure 23).

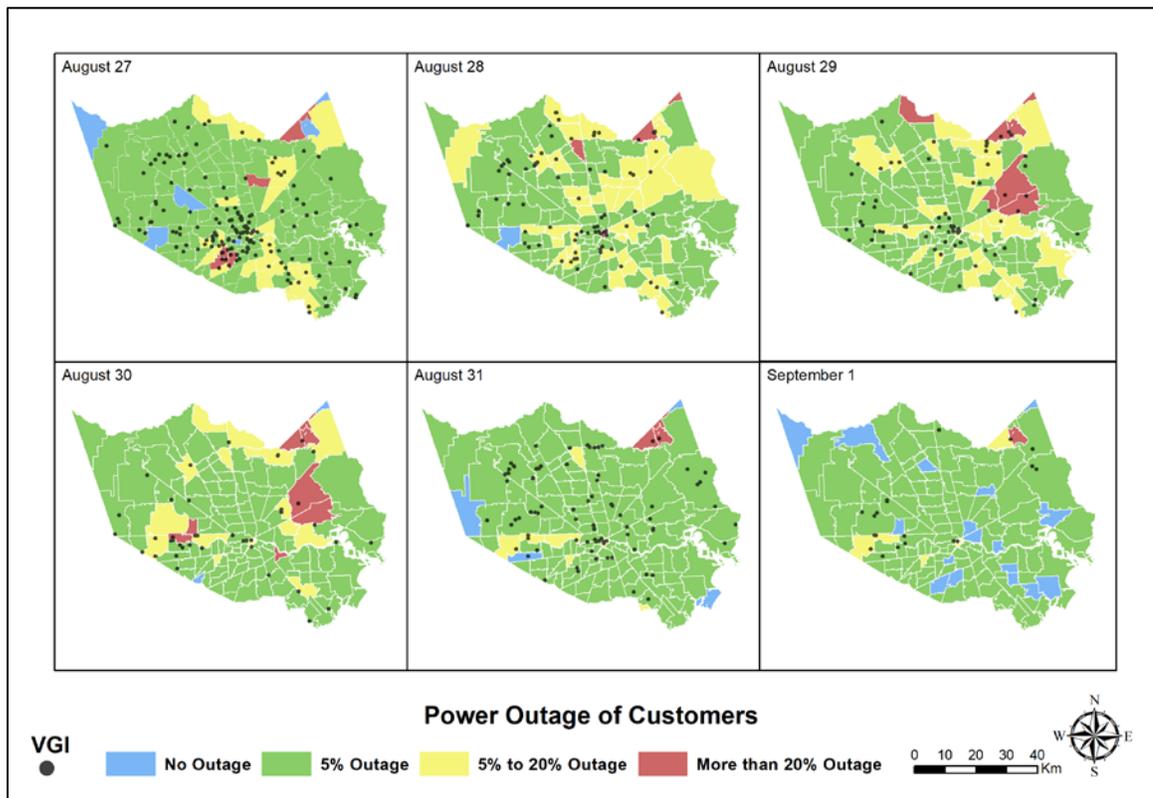


Figure 23. Power Outage Between August 27th and September 1st in Harris County. (data source: ArcGIS 2018).

The overall findings of the VGI data modalities analysis suggests that the limited count of text observations influences the homogeneity of the VGI data. In a previous study, limited counts of text information along with limited spatial coverage could possibly result in underestimated flood extent (Li et al. 2017).

For the comparison between VGI WD and RS WD (i.e. the second research question), the results showed a significant difference between the two WD datasets. It is important to note that the VGI surface used was interpolated, which could have “smoothed” the two datasets and narrowed any significant differences. For example, the interpolation surface has varying water depth values ranging between 0.4-1.8 m, while the VGI data had a maximum water depth of 3 m. Because the only post-disaster RS image available was on September 1st, VGI observations on that day were sampled and aggregated for interpolation, and locations where multiple tweets coincide (e.g. 3m along with varying water depths) were averaged for interpolation. One complication to interpret this finding, however, is the relatively coarse spatial resolution of the RS data being used (20 m after resampling). It is possible to take advantage of pan-sharpening methods to enhance the detection of water pixels more accurately at fine details (Du et al. 2016). However, cloud cover limits the availability of the usable RS image and the ability to detect water leading to possible underestimation of water body (Schnebele et al. 2014; Li et al. 2017).

The third research question examined the validation of VGI WD against authoritative data sources of USGS stream gauges and FEMA depth grids. Using the stream gauges as references, the analysis showed that there is no difference between VGI WD and average water depth from the gauges. In light of water information available at fine temporal resolution for both datasets, this finding of indifferences between the data sources is indicative of the quality of VGI. In addition, the close proximity between the geographic locations of stream gauge and VGI points could also attribute to such agreement.

The second validation was against FEMA depth grids. The analysis showed a significant difference between the two datasets. Despite this finding, one uncertainty in this study is the manual extraction method from the VGI data, where water depth may be subjected to the geographic reference (e.g. height of curbside, hydrant) being used (Table 4). On the other hand, the modeled depth grids from FEMA simulated at a given time would have higher internal consistency. In addition, the spatial variation between FEMA water depth grid and VGI might also influence the comparison between both datasets. The water depth grids from flood simulation are mainly modeled and calibrated at stream gauges, whereas majority of the VGI observations were closer to the urban landscapes and more are found within residential areas (Figure 17). As mentioned, the low spatial accuracy of some VGI points, especially those overlapped at the same location as discussed above, could jeopardize the examination of any significant differences between the two datasets. Nevertheless, the findings of no significant difference between VGI and USGS stream gauge data but its use in modeled water depth grid suggest a myriad effect of different forces in the propagation of uncertainties in flood modeling.

Limitations

Most of the relevant tweets did not include water depth information, where 4.5% of all relevant tweets included water depth information. It could be associated with the selected hashtags and keywords that might filter out tweets that could be counted as relevant, or with the way users might share their experience with the hurricane by using words might not be explicitly related to the event. One of the reasons explaining the large count of relevant tweets is the participation and engagement many Harvey-related hashtags were promoted by individuals and different governmental agencies as a tool for

information dissemination regarding the event (Smith et al. 2015). Nevertheless, these promotion efforts did not target the solicitation of WD information from the public. Another reason is that many tweets might use a popular relevant hashtag (e.g. #HurricaneHarvey) but are not related to water depth (Figure 24). Such VGI content in Twitter increases the time to collect, process and extract relevant water information and may produce false assumptions regarding the availability of relevant water information to be extracted from such data.



Figure 24. An Example of a Tweet with Relevant Hashtag But without Relevant Water Content.

An additional reason that could explain the limited availability of water information in relevant tweets is the missing or broken links of the posted information (Figure 25). Such links could have the potential to increase the observations of water level during the hurricane if they were available.

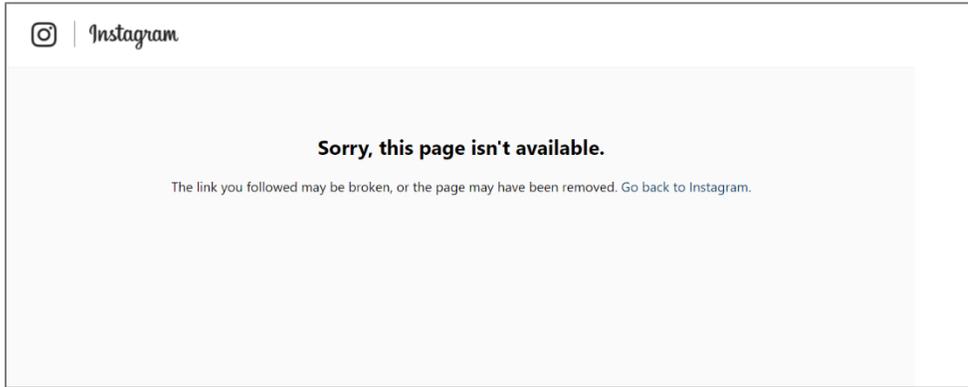


Figure 25. An Example of Relevant Tweets with Missing or Broken Links.

Other explanation could be associated with the re-use of a multimedia by multiple users. Some users might prefer to participate with the affected communities by re-posting or attach a multimedia already shared by other users (Figure 26). These factors influence the time and effort required and consumed to collect, preprocess, classify, and extract relevant information regarding the disaster event from VGI sources. A fifth possible source of limited water observations count is the possibility of location duplication of the geotagged tweets. For example, 44 tweets with water observations from different users shared the same latitude and longitude location on different days. This could be related to the attachment of the same location by the users when posting the tweet. Multiple users might prefer to tag the name of the location, for instance Downtown Houston, when sharing a tweet, which will assign the same coordinates for all the tweets with the same location attached, which is a limitation in Twitter data (Steiger et al. 2015).



Figure 26. An Example of a Re-used Picture During the Event. This Picture was Attached to Multiple Crowdsourced Points in the Study Area.

In overall, the study limitations were associated with the preprocessing and extraction of water depth from the VGI and RS data availability and analysis. The manual extraction from relevant VGI is time-consuming and might be difficult when no obvious physical marks are available, or when they are distant from the position of the user when he/she took the picture or the video. Moreover, the re-use of an image by multiple users makes it difficult to find the duplicate posts with a large amount of data to be observed and may lead to overestimating the count of relevant tweets. In addition, part of the relevant tweets, according to the hashtags and keywords, were sharing non-relevant information (e.g. family pictures or food pictures), and a group of tweets shared the same geographic location, which influences the analysis. Besides, some of the multimedia shared were taken during the night, and it's too dark to observe the level of water.

Regarding RS limitations, the availability of RS data with low or reasonable cloud coverage during or right after the event was limited. In addition, the spatial resolution (20

m) influenced the results of delineated water bodies. The interpolation of VGI data to the spatial resolution of the RS water bodies was time consuming and computational intensive. It took up to 37 hours to interpolate the VGI points to 1.5 m raster using kriging interpolation.

VI. CONCLUSION

This study focused on assessing the quality of VGI data by comparing each VGI data modality (text, picture, and video) against each other and compared to traditional sources of flood data, including remote sensing and authoritative data sources. The VGI data modalities showed significant differences in terms of quality, spatial, and temporal properties. This finding is significant since it was a gap in the literature that required to be assessed.

Regarding the second research question, the study attempted to fill the gap of validating VGI data with RS images. The results showed that limitation of temporal coverage and the existence of clouds in RS data were possible reasons for the disagreement between both data. For future studies, using radar data might increase the probability of leveraging the RS data sources and overcome the cloud problem in the scenes.

Finally, the validation of VGI-derived water depth against authoritative data showed a significant agreement with the USGS gauges. This could be useful for studies using water depth as an input for flood modeling. In addition of stream gauge data, the potential inclusion of VGI may increase the number of observations spatially and temporally and could be used to calibrate and leverage the outputs of physical flood models, e.g. HEC-RAS. The addition of VGI data can also be useful for risk assessment and emergency response to floods. Specifically, the availability of water depth in urban areas, where the coverage of stream gauges is limited, could increase the quality of the flood modeling outputs. However, the modeled depth grids from FEMA did not agree with the synthesized VGI. For future analysis, VGI water depth extraction from

multimedia should be assessed and calibrated to increase the accuracy of the interpolated surfaces to be compared with the modeled depth grids.

Future research should consider the utilization of a pictures and videos VGI modalities as a supplementary source to other data in flood analysis since they shared the most similarities and had the most occurred observations compared with text. For future work, this study suggests measuring the spatial and temporal accuracy of the geotagged tweets for better spatial and temporal analysis. While this study adopted kriging to interpolate WD, the quality of interpolation could be subjected to sample size, spatial autocorrelation among other factors. Hence, future research can consider comparing multiple spatial interpolation methods and suggest suitable technique(s) for a given set VGI sample for flood analysis. Furthermore, this study used visual interpretation for water depth extraction from VGI data modalities, and the same approach could be used to extract damage information and compare it with authoritative damage assessment data for rapid damage assessment analysis. Interestingly, the digital divide did not show a significant association with the spatial distribution and availability of VGI in the study area. Future research may include different socioeconomic status to enhance and leverage the DDI used in this study for better understanding of the impact of the digital divide on sharing relevant information during disasters at multiple locations. Finally, for rapid flood analysis, developing a new method to automate water depth extraction from VGI data using pattern recognition, change detection, and semantic and sentiment analysis should be taken under consideration in the future.

REFERENCES

- Abbott, M. B., J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen. 1986. An introduction to the European Hydrological System—Systeme Hydrologique Europeen,SHE, 1: History and philosophy of a physically-based, distributed modelling system. *Journal of hydrology* 87 (1-2):45-59.
- Ahearn, E. A. 2004. Regression equations for estimating flood flows for the 2-, 10-, 25-, 50-, 100-, and 500-year recurrence intervals in Connecticut.
- Amarnath, G. 2014. An algorithm for rapid flood inundation mapping from optical data using a reflectance differencing technique. *Journal of Flood Risk Management* 7 (3):239-250.
- ArcGIS. 2018. Hurricane Harvey Power Outages by Zip Code.
<http://www.arcgis.com/home/item.html?id=6d251bb5e48e492c90ebac123011ffdb>
(last accessed 22 April 2018).
- Arnaud, P., and J. Lavabre. 2002. Coupled rainfall model and discharge model for flood frequency estimation. *Water Resources Research* 38 (6).
- Brunner, G. W. 2002. Hec-ras (river analysis system). Paper read at North American Water and Environment Congress & Destructive Water.
- Casas, A., G. Benito, V. Thorndycraft, and M. Rico. 2006. The topographic data source of digital terrain models as a key element in the accuracy of hydraulic flood modelling. *Earth Surface Processes and Landforms* 31 (4):444-456.

- Cervone, G., E. Sava, Q. Huang, E. Schnebele, J. Harrison, and N. Waters. 2015. 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.
- Cigler, B. A. 2017. US Floods: The Necessity of Mitigation. *State and Local Government Review*:0160323X17731890.
- de Albuquerque, J. P., B. Herfort, A. Brenning, and A. Zipf. 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.
- Deng, Q., Y. Liu, H. Zhang, X. Deng, and Y. Ma. 2016. A new crowdsourcing model to assess disaster using microblog data in typhoon Haiyan. *Natural Hazards* 84 (2):1241-1256.
- Dilley, M., R. S. Chen, U. Deichmann, A. L. Lerner-Lam, and M. Arnold. 2005. Natural disaster hotspots: a global risk analysis: World Bank Publications.
- Dinu, C., R. Drobot, C. Pricop, and T. V. Blidaru. 2017. Flash-flood modelling with artificial neural networks using radar rainfall estimates. *Mathematical Modeling in Civil Engineering* 13 (3):10-20.
- Du, Y., Y. Zhang, F. Ling, Q. Wang, W. Li, and X. Li. 2016. Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sensing* 8 (4):354.

- Federal Emergency Management Agency (FEMA). 2015. Business Infographics.
<https://www.fema.gov/media-library/assets/documents/108451> (last accessed 17 February 2018).
- FEMA. 2017a. Historic Disaster Response to Hurricane Harvey in Texas.
<https://www.fema.gov/news-release/2017/09/22/historic-disaster-response-hurricane-harvey-texas> (last accessed 3 December 2017).
- . 2017b. National Disasters.
<https://data.femadata.com/NationalDisasters/HurricaneHarvey/> (last accessed 10 April 2018).
- . 2018. Significant Flood Events. <https://www.fema.gov/significant-flood-events> (last accessed 18 February 2018).
- Feyisa, G. L., H. Meilby, R. Fensholt, and S. R. Proud. 2014. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment* 140:23-35.
- Floodlist. 2016. Houston Floods – President Declares Major Disaster.
<http://floodlist.com/america/usa/houston-floods-president-declares-major-disaster> (last accessed 24 March 2018).
- Fohringer, J., D. Dransch, H. Kreibich, and K. Schröter. 2015. Social media as an information source for rapid flood inundation mapping. *Natural Hazards and Earth System Sciences* 15 (12):2725-2738.
- Fortune. 2017. Hurricane Harvey Damages Could Cost up to \$180 Billion.
<http://fortune.com/2017/09/03/hurricane-harvey-damages-cost/> (last accessed 23 February 2018).

- Gallardo, R. 2017. Digital Divide Index. Mississippi State University.
http://ici.msucare.com/sites/ici.msucare.com/files/2015_ddi.pdf (last accessed 22 April 2018).
- Goodchild, M. F. 2007. Citizens as voluntary sensors: Spatial data infrastructure in the world of web 2.0. *International Journal of Spatial Data Infrastructures Research* 2:24-32.
- Gregg, W. W., and N. W. Casey. 2004. Global and regional evaluation of the SeaWiFS chlorophyll data set. *Remote Sensing of Environment* 93 (4):463-479.
- Guan, X., and C. Chen. 2014. Using social media data to understand and assess disasters. *Natural Hazards* 74 (2):837-850.
- Guinot, V., and P. Gourbesville. 2003. Calibration of physically based models: back to basics? *Journal of Hydroinformatics* 5 (4):233-244.
- Hese, S., and T. Heyer. 2016. Earth observation data based rapid flood-extent modelling for tsunami-devastated coastal areas. *International Journal of Applied Earth Observation and Geoinformation* 46 (Supplement C):63-83.
- Homeland Infrastructure Foundation-Level Data (HIFLD). 2017. Hurricane Harvey Video Resources. <https://respond-harvey-geoplatform.opendata.arcgis.com/datasets/hurricane-harvey-video-resources?geometry=-106.583%2C27.242%2C-85.676%2C30.607> (last accessed 21 February 2018).
- Howe, J. 2006. The rise of crowdsourcing. *Wired magazine* 14 (6):1-4.
- Hwang, S. N., and S.-W. Lee. 2017. Does Environmental Risk Affect Human Migration Behavior? *Journal of Geographic Information System* 9 (04):493.

- Jensen, J. R. 2009. Remote sensing of the environment: An earth resource perspective 2/e: Pearson Education India.
- Karaska, M. A., R. L. Huguenin, J. L. Beacham, M.-H. Wang, J. R. Jensen, and R. S. Kaufmann. 2004. AVIRIS measurements of chlorophyll, suspended minerals, dissolved organic carbon, and turbidity in the Neuse River, North Carolina. *Photogrammetric Engineering & Remote Sensing* 70 (1):125-133.
- Kia, M. B., S. Pirasteh, B. Pradhan, A. R. Mahmud, W. N. A. Sulaiman, and A. Moradi. 2012. An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environmental Earth Sciences* 67 (1):251-264.
- Larkin, T. J., and G. W. Bomar. 1983. *Climatic atlas of Texas*: Texas Department of Water Resources.
- Li, Z., C. Wang, C. T. Emrich, and D. Guo. 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*:1-14.
- Mandel, B., A. Culotta, J. Boulahanis, D. Stark, B. Lewis, and J. Rodrigue. 2012. A demographic analysis of online sentiment during hurricane irene. Paper read at Proceedings of the Second Workshop on Language in Social Media.
- Marsh, W. M., and J. Grossa Jr. 2005. *Environmental geography: science, land use, and earth systems*. New York, USA: John Wiley and Sons.
- Meesuk, V., Z. Vojinovic, and A. E. Mynett. 2012. Using multidimensional views of photographs for flood modelling. Paper read at Information and Automation for Sustainability (ICIAfS), 2012 IEEE 6th International Conference on.

Mississippi State University. 2018. DDI Digital Divide Index.

<http://ici.msucare.com/resources/ddi> (last accessed 22 April 2018).

Montz, B. E., G. A. Tobin, and R. R. Hagelman. 2017. *Natural hazards: explanation and integration*: Guilford Publications.

National Alliance for Public Safety GIS (NAPSG). 2017. Hurricane Harvey Photos.

<https://napsg.maps.arcgis.com/apps/StoryMapCrowdsource/index.html?appid=b6ef838e4d26489e8f62102639dc3d91> (last accessed 23 February 2018).

Papioannou, G., A. Loukas, L. Vasiliades, and G. Aronica. 2016. Flood inundation mapping sensitivity to riverine spatial resolution and modelling approach. *Natural Hazards* 83:117-132.

Pradhan, B. 2010. Flood susceptible mapping and risk area delineation using logistic regression, GIS and remote sensing. *Journal of Spatial Hydrology* 9 (2).

Riggins, F. J., and S. Dewan. 2005. The digital divide: Current and future research directions. *Journal of the Association for information systems* 6 (12):13.

Sampson, C. C., T. J. Fewtrell, A. Duncan, K. Shaad, M. S. Horritt, and P. D. Bates. 2012. Use of terrestrial laser scanning data to drive decimetric resolution urban inundation models. *Advances in water resources* 41:1-17.

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

Shalunts, G., G. Backfried, and P. Prinz. 2014. Sentiment analysis of German social media data for natural disasters. Paper read at ISCRAM.

- Smith, L., Q. Liang, P. James, and W. Lin. 2015. 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*.
- Steiger, E., R. Westerholt, B. Resch, and A. Zipf. 2015. Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data. *Computers, environment and urban systems* 54:255-265.
- Strömberg, D. 2007. Natural disasters, economic development, and humanitarian aid. *The Journal of Economic Perspectives* 21 (3):199-222.
- Texas Natural Resources Information System (TNRIS). 2008. Data Search and Download. <https://tnris.org/data-download/#!/county/Harris> (last accessed 4 December 2017).
- Texas Parks and Wildlife Department (TPWD). n.d. Ecological Mapping System. <https://tpwd.texas.gov/landwater/land/programs/landscape-ecology/ems/> (last access 23 February 2018).
- Tsakiris, G. 2014. Flood risk assessment: concepts, modelling, applications. *Natural Hazards and Earth System Sciences* 14 (5):1361.
- U.S. Census Bureau. n.d. Houston city, Texas. <https://www.census.gov/quickfacts/fact/table/houstoncitytexas/PST045216> (last accessed:12 March 2017).
- U.S. Geological Survey (USGS). 2016. Flood inundation mapping (FIM) program. https://water.usgs.gov/osw/flood_inundation/ (last accessed 11 October 2017).
- USGS. 2018a. EarthExplorer. <https://earthexplorer.usgs.gov/> (last accessed 21 February 2018).

———. 2018b. National Water Information System.

<https://nwis.waterdata.usgs.gov/nwis/uv> (last accessed 10 April 2018).

Weather. 2017. Harvey Forces New Evacuations in Texas as Waterways Burst Banks in Houston-Area Flood. <https://weather.com/storms/hurricane/news/harvey-houston-texas-flooding-evacuations> (last accessed 3 December 2017).

Zrinji, Z., and D. H. Burn. 1992. Application of Kriging to surface water level estimation. *Canadian Journal of Civil Engineering* 19 (1):181-185.