

**Cloud-based Federation and Fusion of Distributed Geospatial Data Sources for  
Supporting Hurricane Response: Requirements, Challenges, and Opportunities**

FINAL REPORT  
December 2016

Submitted by:

Jie Gong<sup>1</sup>  
Assistant Professor  
Rutgers University

Manish Parashar<sup>2</sup>  
Professor

Xuan Hu<sup>1</sup>  
Graduate Research Assistant

Zixiang Zhou<sup>1</sup>  
Graduate Research Assistant

<sup>1</sup> CAIT  
Rutgers University

<sup>2</sup>Department of Computer Science  
Rutgers, The State University of New Jersey

<p>In cooperation with Rutgers, The State University of New Jersey And U.S. Department of Transportation Federal Highway Administration</p>
---

### **Disclaimer Statement**

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

1. Report No. CAIT-UTC-047	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Cloud-based Federation and Fusion of Distributed Geospatial Data Sources for Supporting Hurricane Response: Requirements, Challenges, and Opportunities		5. Report Date December 2016	
		6. Performing Organization Code CAIT/Rutgers University/Virginia Tech/ University of Delaware	
7. Author(s) Jie Gong, Manish Parashar, Xuan Hu, and Zixiang Zhou		8. Performing Organization Report No. CAIT-UTC-047	
9. Performing Organization Name and Address Rutgers University, 100 Brett Road, Piscataway, NJ 08854		10. Work Unit No.	
		11. Contract or Grant No. DTRT12-G-UTC16	
12. Sponsoring Agency Name and Address Center for Advanced Infrastructure and Transportation Rutgers, The State University of New Jersey 100 Brett Road Piscataway, NJ 08854		13. Type of Report and Period Covered April 2013 - December 2014	
		14. Sponsoring Agency Code	
15. Supplementary Notes U.S. Department of Transportation/Research and Innovative Technology Administration 1200 New Jersey Avenue, SE Washington, DC 20590-0001			
16. Abstract Geospatial data have been playing an increasingly important role in disaster response and recovery. For large-scale natural disasters such as Hurricanes which often have the capacity to topple a large region within a span of a few days, disaster preparation and response often involves intensive collaborations between multiple agencies and organizations, some of them are geographically distant. However, such collaborations are often hampered by several obstacles: (1) Agencies and utility owners are often reluctant to openly share their own data sets due to their concerns on data security; (2) Sharing and moving of large-scale data sets are difficult, especially over the internet; and (3) Lack of data management solutions that are both scalable and distributed to accommodate various integrative analysis needs. As the result, integrative analyses of system vulnerabilities and system-level planning on response and recovery actions are often not possible. This study focused on Hurricane Sandy as an empirical case to characterize the types of geospatial data sets useful for disaster management, characterize processing patterns and needs with these geospatial data, review types of cloud computing infrastructures that are available for geospatial data analytics during natural disasters, and suggest gaps in technical capabilities and needs in future research efforts. The study provided first-hand knowledge about the type and the volume of geospatial data sets that now can be collected to support disaster management and a synthesis of cloud computing infrastructures for geospatial computing in time-sensitive applications. These outcomes can be used by transportation agencies to better plan their response activities.			
17. Key Words Disaster Response, Cloud Computing, Geospatial Data Processing		18. Distribution Statement	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 38	22. Price

## **Acknowledgments**

This project would not have been possible without the support from the Rutgers Center for Advanced Infrastructure and Transportation. We would like to thank Dr. Ali Maher, Director of CAIT center and Dr. Patrick Szary, Associate Director of CAIT center for their support. Finally, we appreciate the support from many CAIT staff members.

## Table of Contents

Acknowledgments.....	4
List of Figures.....	6
List of Tables.....	7
INTRODUCTION .....	8
Description of the Problem.....	8
Relevance to Strategic Goals .....	8
Background .....	8
Research Goals and Objectives.....	9
Overview of the Report.....	10
APPROACH .....	10
METHODOLOGY .....	10
FINDINGS .....	12
Geospatial data sources for supporting Hurricane Sandy response.....	12
Identify data sharing and integrative analysis needs.....	17
Airborne lidar applications.....	19
Ground-based lidar systems .....	21
Successful Applications of Cloud-based Data Infrastructure for Geospatial Data Management .....	25
Requirement Modeling on Cloud-based Geospatial Data Analytics for Time Sensitive Applications .....	27
Web-based Disaster Data Collection, Visualization, and Processing: Examples and Case Studies.....	29
Case Study 1: Fusion of Geo-Tagged Crowdsourced Post-Storm Damage Photos with Mobile LiDAR Data for Storm Surge Height Measurement.....	29
Case Study 2: Post-Disaster Damage Assessment of Natural Gas Pipeline Systems with Remotely Sensed Data.....	34
CONCLUSIONS.....	36
REFERENCES .....	37

## List of Figures

Figure 1 Ortley Beach Change Detection Performed using Pre- and Post- Airborne lidar .....	19
Figure 2 An Overlay of Pre-Sandy and Post-Sandy Airborne LiDAR Data and Post-Sandy Airborne Imagery .....	20
Figure 3 Volume differences generated by subtracting post and pre cleanup conditions.....	20
Figure 4 Sample debris volume calculation of a specific area.....	21
Figure 5 A LIDAR-based Virtual Disaster Site .....	21
Figure 6 Change Detection Results for Ortley Beach, NJ.....	22
Figure 7 Per Building Level Damage Information Extraction.....	22
Figure 8 Integrating Mobile LiDAR and Pre-Event Airborne Imagery for Displacement Measurement .....	23
Figure 9 The Reconstructed Flooding Scenario for Ortley Beach.....	24
Figure 10 The New ABFE 100-year Flood Plane overlaid on Rockaway Homes.....	25
Figure 12 Information Value vs. Time (adapted from Hodgson et al. 2014).....	29
Figure 13 Proposed Surge Height Estimation Approach .....	30
Figure 14 GPS-based Screening of Images from Different Sources .....	30
Figure 15 SIFT-based Matching between VGI photos and Mobile Lidar Imagery .....	31
Figure 16 Establishing Correspondence Between Point Clouds and VIG Images .....	33
Figure 17 The correspondence between Lidar Point Clouds and VGI Images.....	33
Figure 18 High Water Mark Measurement from VGI Images and Point Cloud Data .....	33
Figure 19 Cloud-based Post Disaster Damage Assessment of Natural Gas Pipeline System ...	35
Figure 20 Proposed Post-disaster Pipeline Risk Assessment Framework .....	36

## List of Tables

Table 1 Data Storage, Sharing, and Processing Mechanisms .....	26
Table 2 Distributed Computing for Geospatial Data Processing .....	27
Table 3 Volume of Hurricane Sandy related 3D disaster data sets .....	28
Table 4 Pipeline threat and related indicators.....	36

## **INTRODUCTION**

### **Description of the Problem**

Unlike most weather events, hurricanes have the capacity to topple a large region within a span of a few days. Increasingly, preparing for and responding to a major hurricane event involves intensive collaborations between multiple agencies and organizations, some of them are geographically distant. To facilitate large-scale collaborations, technology is required that enables data at multiple sites, often geospatially coded, to be interlinked and presented as a unified source, allowing stakeholders to perform integrative analyses. However, such collaborations are often hampered by several obstacles: (1) Agencies and utility owners are often reluctant to openly share their own data sets due to their concerns on data security; (2) Sharing and moving of large-scale data sets are difficult, especially over the internet; and (3) Lack of data management solutions that are both scalable and distributed to accommodate various integrative analysis needs. As the result, integrative analyses of system vulnerabilities and system-level planning on response and recovery actions are often not possible.

### **Relevance to Strategic Goals**

During the last decade, nature disasters have been creating mounting stresses to the nations aging transportation infrastructure. In contrast to the gradual deterioration process in normal operating conditions, transportation infrastructures are often facing the danger of rapid destruction during nature disasters, calling for immediate mitigation and response actions. In this research, we will investigate effective data federation and fusion methods to support decision makings on these rapid mitigation and response actions. The outcome of this research will contribute to improving the security and resilience of the critical infrastructures.

### **Background**

Severe weather events such as hurricanes, ice storms, surge and flooding have been occurring across U.S. and around the world, threatening places where economic and industrial activities are heavily concentrated. Advanced geospatial sensing technologies are playing an increasingly important role for decision support in disaster



preparation, response, and recovery operations as they greatly expand our ability of collecting disaster data. For instance, in recent years, because state and federal agencies have made airborne LiDAR (Light Detection and Ranging) data collection a priority, post-storm LiDAR collection is now routine after large surge event and vast amounts of disaster data are now freely available online (e.g. NOAA's Digital Coast). In another example, emerging high resolution sensing systems such as terrestrial/mobile LiDAR have also been deployed for damage data collection during recent events such as Superstorm Sandy, generating an unprecedented amount of 3D geospatial data. Analysis of these data sets offers tremendous opportunities in improving our understanding, modeling, and prediction of the impacts of coastal hazards on transportation infrastructure systems and coastal communities as well as providing real world education and training experiences in classrooms and offices. Despite this potential, there are fundamental challenges in managing, analyzing, and interpreting the growing size and complexity of spatial disaster data sets. The vast size and complex processing requirements of these new data sets make it challenging to effectively use them in real-life scenarios. For example, during large-scale coastal storm events, crucial information is often hidden in these data sets and is in no way integrated into ongoing decision processes. To fully exploit their potential, we need to revisit data sharing and analytics methods and develop new capabilities to rapidly synthesize information from these large data sets.

### **Research Goals and Objectives**

In this research, we will explore the potential of cloud-based federation and fusion of distributed geospatial data sources for supporting hurricane response. Specific objectives are:

- (1) Understanding how geospatial data sources have been used, shared, and analyzed to support major natural disasters like Hurricane Sandy and what are the limitations
- (2) Identifying opportunities for using cloud computing to enhance data sharing, federation, and integrative analysis at distributed sites.

(3) Developing testing cases to highlight the effectiveness of cloud computing to support future data-intensive disaster response applications.

The product of this research will be a catalog of lessons learned in Hurricane Sandy response related to geospatial data support and a guidance on recommended cloud computing techniques and data infrastructures for enhancing geospatial data support during major disasters.

### **Overview of the Report**

This report documents the research approach, methodology, findings, conclusions and recommendations of this collaborative research project. The following sections outline the approach and methodology. The next section presents the findings, followed by sections documenting the conclusions and making recommendations for future work and application in state Departments of Transportation.

### **APPROACH**

This research involves the collaboration between CAIT and Rutgers' RDI2. Dr. Jie Gong from the CAIT led the project and be responsible for the overall management of the project. Building on his expertise in geospatial sensing and post-disaster damage assessment, he will focus on identifying geospatial data sources, developing data process workflow, gathering data from stakeholders, and implementing data federation and fusion on a cloud-based platform. Dr. Manish Parshar provided guidance on cloud computing techniques and distributed processing implementation. The research involved several student researchers in data gathering and research development.

### **METHODOLOGY**

The following tasks will be undertaken to complete the objectives of this project.

Task 1. Identify geospatial data sources for supporting Hurricane Sandy response (01/2014 - 02/2014)

In this task, we will examine how various data sets relevant to disaster preparation and response were utilized and processed during the Hurricane Sandy to gain understanding on existing data analytics practices.

Task 2. Identify data sharing and integrative analysis needs (03/2014 - 05/2014)

This task involves interviewing relevant state officials to understand the existing workflow used in gathering, distributing, and analyzing geospatial data sources to gain situation awareness and to identify limitations of these existing methods. Through the analysis of existing practices, the study will further explore how these distributed data sets can be shared, federated, and analyzed in a more efficient and effective manner.

Task 3. Review and identify successful applications of cloud-based data platform and distributed computing for mission critical applications (03/2014 - 05/2014)

This task involves reviewing literature and industry case studies to develop a catalog of successful applications of cloud computing in data intensive applications.

Task 4. Requirement modeling on using cloud computing infrastructures to support data intensive disaster response applications (06/2014 - 09/2014)

A series of quantitative analyses on data load, bandwidth requirement, data streaming capacity need, query optimization requirement will be conducted in order to fully understand hardware and software needs for transforming the existing data analytics into a cloud-based one.

Task 5. Demonstration on using cloud computing as data infrastructure for post disaster applications (09/2014 - 11/2014)

In this task, we will develop testing cases to demonstrate the use of Cloud-based data infrastructure for providing a unified data federation and fusion platform for real-time applications.

Task 6. Final Reporting (11/2014 - 12/2014)

The project relied largely on literature review, discussion and expert input. The original proposal involved eight tasks that are summarized here for completeness.

Expected outcomes from this research are:

- (1) List of data sources for supporting hurricane preparation, response, and recovery
- (2) Catalog of implemented methods in sharing, distributing, gathering, and analyzing geospatial data during Hurricane Sandy and their limitations

(3) Recommendations on the required computing platform for data federation and fusion to support more effective disaster response, including cost, functionality, and system configurations

(4) Final reporting documenting the process used to develop these products.

## **FINDINGS**

### **Geospatial data sources for supporting Hurricane Sandy response**

Use of remotely sense data in disaster management is an extensively researched area. These remotely sensed data include data sensed from space, air, and land. More recently, the rising of social media network is another form of sensing that brings data in large geographic regions. This is also coined as social sensing and crowd-sourcing. In the following, a brief overview of remote sensing from different platforms is provided.

**Satellite remote sensing platform:** The use of satellite remote sensing in disaster is becoming increasingly common as the advancement in of state of the art geospatial technology and increased awareness in the demand for disaster need (Joyce et al 2009). Geospatial data comes from both active and passive sensors carried by satellites. From the hurricane prospective, the flooding damage is readily apparent in both optical (imagery) and Synthetic aperture radar (SAR) data. High resolution optical imagery has been widely employed in forecasting, monitoring and assessing natural disasters in general. One of the very first pioneer studies in precutting the tropical intensities can be found in Dvorak (1975). Recently techniques such as Landsat-5 TM and ALOS AVNIR-2 have successfully mapped the flooding event of 2007 and 2008. SAR provide the strong capability to detect flooding area even in closed vegetation area, since the backscatter signature of water is distinctive by the nature of this technique (Lewis et al 1998). Taking advantage of this unique feature, Lymburner et al (2008) and Thankappan et al (2008) extracted flooding inundation area using SAR. Besides flooding damage, hurricane often combined with wind damage. Shen et al (2009) proposed a speed ambiguity removal algorithm to retrieve wind speed in Hurricane Rita (2005).

**Airborne and UAV systems:** Airborne Mapping System (AMS) is suitable for large area scanning or wide regional operations (Yang et al 2015), which naturally meets the timely issue in disaster response. Paired with the LiDAR sensor, it offers the capability to capture three-dimensional (3D) point clouds in different disaster scenarios. As a relative accurate, fast and versatile measurement technique (Wehr and Lohr 1999), AMS survey even become a routine survey practice in some extreme events. In recent years, improvements in Unmanned Aerial Vehicles (UAVs) enable an alternative remote sensing platform for disaster response (Wallace et al 2012). It offers a distinctive combination of high resolution 3D points at a significantly lower cost. While the UAV image system has been widely studied in literatures, only few of the UAV LiDAR systems are actually deployed in real disaster scenarios. Related studies include forest inventory (Wallace, et al 2011) and landslide (Liu et al 2011).

**Land-based remote sensing systems:** Terrestrial lidar and mobile mapping systems are example sensing systems that have been used in large-scale disaster mapping projects. Compared to air-based systems, land-based systems do not have long ranges like ALS does, but it offers much higher resolution and accuracy. In many cases, these systems have been used for city-scale mapping projects, which provide valuable baseline data for disaster impact assessment. For example, several states in the United States such as Utah and Oregon and cities like Indianapolis have already conducted state-wide/city-wide mobile lidar data collection. Recently, static/mobile lidar have also seen increased deployment during natural disasters.

**Crowdsourcing:** With the advancement in the sensor embedded devices as well as the communication technology, data acquisition is no longer the responsibility of government agencies such as DHS or FEMA. Volunteer geography information (VGI) or crowdsourcing is a potentially valuable data income for disasters. The important role of VGI or crowd-sourced in disaster response is highlighted (Poser et al 2010). The availability of handheld device in the market compliments the traditional LiDAR survey platform as a resource of volunteered geographic information or crowd-sourced. The new technology presumably may have altered the emergency response mapping needs, but the reliability and quality of such a derived data are still a concern (Hodgson et al 2014).

One of the growing importance of data used in disaster response and vulnerability assessment is point cloud data from lidar systems. In the following, we provide a systematic overview of lidar technologies.

Lidar (also written LIDAR or LiDAR) is a remote sensing technology that measures distance by illuminating a target with a laser and analyzing the reflected light. A lidar system typically consists of several components:

- Laser transmitter and detector/receiver
- Deflection mechanism of the laser ray
- GPS/INS
- Computer, onboard software and storage devices, including precise timing device that synchronizes all sensors
- Optionally other optical sensors such as digital cameras

For a lidar system, 600–1000 nm lasers are most common used for non-scientific applications. They are inexpensive, but since they can be focused and easily absorbed by the eye, the maximum power is limited by the need to make them eye-safe. Eye-safety is often a requirement for most applications. A common alternative, 1550 nm lasers, are eye-safe at much higher power levels since this wavelength is not focused by the eye, but the detector technology is less advanced and so these wavelengths are generally used at longer ranges and lower accuracies. They are also used for military applications as 1550 nm is not visible in night vision goggles, unlike the shorter 1000 nm infrared laser. Other than the type of lasers used, a lidar system is also characterized by the following mechanical and performance factors:

- Pulse repetition frequency (PRF) or pulse rate: number of pulses sent per second
- Echoes: number of received pulse reflections recorded for one sent pulse
- Minimum vertical object separation: minimum distance between 2 separate echoes
- Scan rate: number of scan patterns (e.g. scan lines) per second
- Field of View (FOV) or scan angle: across-flight angle that laser beam can cover

- Beam divergence: the angle showing the deviation of the laser beam from parallelity
- Wavelength: important for measuring certain objects
- GPS/INS measurement frequency and accuracy
- Range resolution and accuracy

Lidar systems can be installed on a variety of platforms such as airborne or ground-based platforms. While there are occasionally other platforms such as satellites or waterborne vehicles which can be used, airborne and ground-based Lidar systems are the most commonly systems.

### **Airborne lidar systems**

The typical platforms used for airborne lidar are fixed-wing airplanes, helicopters, and more recently unmanned airborne vehicles (UAVs). According to their applications, airborne lidar systems can be further divided into airborne topographic mapping and bathymetric mapping lidars.

#### ***Airborne topographic mapping lidars***

Airborne topographic mapping lidars generally use 1064 nm diode pumped YAG lasers. Airborne lidar systems typically use scanning lasers at pulse rates that can exceed 100k/second to produce dense ( $>1/\text{square meter}$ ), high accuracy ( $\sim 0.1\text{m}$  vertical) point clouds along 300-600m-wide swaths at forward speeds around 100knots. Returns will include both canopy (trees, houses) and ground, often with multiple returns from a single lidar pulse, and often have co-located aerial photography.

#### ***Bathymetric mapping lidars***

Designed for accurate sea-depth determination, bathymetric lidar systems are composed of two beams, one green (532 nm) and one infrared (1064 nm). The green beam traverses the air-water interface and propagates in the water until the sea bottom with the least attenuation. The infrared beam is reflected by the water and gives the range from the plane to the sea surface. Low-flying aircraft equipped with GPS/IMU and a pulsed laser scanner is the platform of choice for this application. The data are used to support navigation, military operations, and environmental and recreational needs. Bathymetric lidar systems are often flown simultaneously with digital cameras and

hyperspectral sensors to gather additional information about water quality and bottom composition.

### **Ground-based lidar systems**

Ground-based lidar systems are either static (on a stationary platform such as a tripod or mast) or dynamic (on a moving vehicle). It has been incorporated into surveying and metrology instruments and is often employed in mobile lidar systems. In a static implementation, a GPS/INS geo-referencing system is not needed. The lidar is set up over a known point, and the scan angles for each point are recorded in the data set. Reference points on the target surface can also be surveyed to provide additional geo-referencing control. In a dynamic implementation of ground-based lidar, GPS/IMU is utilized to provide geo-referencing, just as it would be on an airborne platform. Using an infrared or green wavelength laser, ground-based lidars pulse at rates up to 1000 Hz, and can map objects from about 1 meter up to 1000 meters away with accuracies on the order of millimeters to a few centimeters. The accuracy of individual points can be affected by atmospheric conditions, distance to the target, angle of incidence of the laser pulse upon the target, and the reflectivity of the target surface. Very shiny, polished surfaces and very black surfaces that absorb nearly all incident light are difficult to image. Three types of scanning systems are employed in ground-based lidars:

- Panoramic scanners rotate 360 degrees around the mounting axis, and scan 180 degrees vertically to provide seamless and total coverage of the surroundings.
- Single axis scanners also rotate 360 degrees but are limited to a 50-60 vertical swath.
- Camera scanners point in a fixed direction with limited angular range both horizontally and vertically.

Like during many other extreme events, geospatial products and tools are an essential part of every stage of disaster management during Hurricane Sandy, from planning through response, and recovery to mitigation of future events. But unlike many other extreme events where the available spatial data are often limited in size and type, Hurricane Sandy has seen a surge of massive remotely sensed data sets. These data sets are generally imagery and point cloud data. These data can be more broadly defined as low-dimensional, spatio-temporal datasets, in which data elements are



defined at points in a 2D/3D spatial coordinate system and over time. The specific data sets considered in this study include various airborne lidar data sets collected at different points of time before and after the landing of Hurricane Sandy. First, airborne lidar data dated back to 2010 exist for most part of the New York-New Jersey metropolitan area and are archived in data repositories including Digital Coast and USGS Click. Second, on October 29, 2016, the day before Hurricane Sandy landed in New Jersey, the USGS Coastal and Marine Geology Program collected airborne lidar data along the New Jersey Coast using its Experimental Advanced Airborne Research LiDAR-B (EAARL-B) system. Immediately after the landing of Hurricane Sandy, NOAA collected airborne imagery followed by USGS EARL-B airborne lidar data collection. During the period of November 11-24, 2012, USACE conducted another wave of airborne lidar data collection along the New Jersey and New York coastal line. During the period of December 5-9, 2012, Rutgers conducted mobile lidar scanning of severely impacted coastal communities in the state of New Jersey and New York City. Throughout the disaster response period, street-level images of storm damage have also been collected by various damage assessment teams and citizens. Some of them were distributed through social media channels such as Facebook and Twitter. During the disaster recovery stage, more geospatial data sets have also been collected for the purpose of assessing recovery progress and future vulnerability. These data sets include 2014 USGS airborne lidar data collection along the coastal lines in the northeast region and mobile lidar data collection in Ocean County, New Jersey in 2016. Collectively, these data sets are too massive to be efficiently managed and processed to derive scientific insights into ways of improving coastal resilience.

### **Identify data sharing and integrative analysis needs**

For years, federal government has tried painstakingly to reduce duplicative and improve capability of geospatial data. Even though large infrastructures such as National Spatial Data Infrastructure (NSDI) has been built, data sharing remains a formidable challenge and desperate need (Koontz 2003). Data sharing and integration is an essential step in data aggregation for centralized decision making. It is not possible to obtain the useful information from a single data source or processing solution that will satisfied the need for all decision makings. In emergency cases, decision makers might be reluctant to use

geospatial information due to a lacking of familiarity or insufficient information (Alexander 2002; Brucewicz 2003; Wang and Yuan 2010). To better make decision-making information available, data sharing is the key step. One example is the visitation for disaster response. Experts need enrich details of visions to develop and validate their experience-based judgements. The importance of information sharing and integration in disaster situation cannot be overstated. It addresses to the need for the decision makers to have a thorough understanding of the disaster area Visualization. It also helps maximize the response in the human visual system and increase the saliency of the object (Quartulli et al 2013).

Sufficient or even abundant information are needed to compare, cross-validate, fuse together so that the information details and accuracy can be improved. Comparison analysis or change detection often require revisit of the same sensor or combination data from different sensors. In terms hurricane damage, flooding depth can be estimated using geo-registered LiDAR data from different time frame or paired SAR and other geospatial data such as DEM. In addition, data captured from different platforms need to fuse so that different resolution can be achieving; the shortcoming of certain platforms can be overcome. For example, data collected from ground based platforms are often with higher resolution, enrich vertical details while data from air-based platforms are featured with fast collection time and larger scales. Another example is the fusion of image with LiDAR: 3D geospatial information provided by LiDAR can add the third dimension to the colorized images using algorithms such as SFMS. Moreover, even crowdsourcing data are potentially helpful in understanding the disaster situation, the correctness are highly doubtful (Gupta et al 2013). Similar situation is confronted when using other geospatial data. Data sharing and integration can provide good reference for the data to cross-validate so that the false message can be filtered out. Most of the current studies use in situ geospatial information, there is few real time or near real time operations portal for hurricane disaster analysis or natural disaster in general. Thanks to the state of art sensing technology, massive data can be generated in disaster environment. At the same time, extracting meaningful information from this data are computation prohibited from one spot within a narrow time window. Especially when multi-tasks are required, the conflict between limited computing resources and the

urgent demand for information is highly obvious. To solve this problem, data sharing is essential so that data can be distributed to nearby available in field or in house computing resources.

Data analysis needs vary significantly from disaster to disaster. This is partially due to the availability of data and resources. In the following, we provide a list of analytic applications with the data sets identified in the Task 1.

### ***Airborne lidar applications***

#### ***Wide-area change detection for damage assessment***

One important analytics application of airborne lidar data is damage assessment based on wide-area change detection between pre- and post-event lidar data. As long as the pre- and post-event are aligned in precision, the differences between pre- and post-event data often indicate changes in terrain conditions and man-made structures, which are often associated with storm damages. For example, Figure 1 clearly shows eroded dunes (in blue and cyan color) and deposited debris on the streets (in red color).

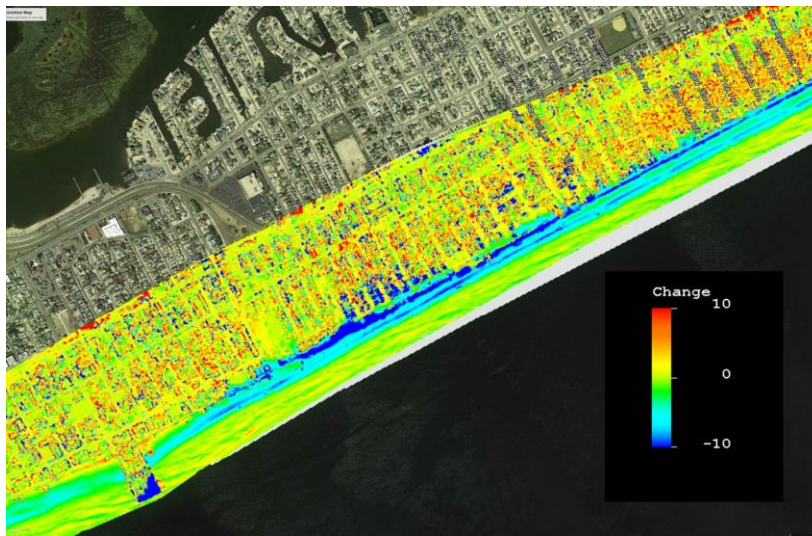


Figure 1 Ortley Beach Change Detection Performed using Pre- and Post- Airborne lidar

#### ***Debris volume quantification***

Large-scale natural disasters such as major hurricanes often generate a tremendous amount of debris that overwhelms the capacity of disaster response organizations and creates roadblocks on the path to recovery. For example, Superstorm Sandy destroyed over 650,000 homes and leaving 8.5 million people without power, while generating over 10 million cubic yards of debris in New Jersey alone. The challenge of debris removal following a natural disaster lies in the difficulty of accurately estimating the volume of debris and executing debris removal in an effective and efficient manner while minimizing impacts on natural environment and complying with legal

requirements. Airborne lidar can be used gather information to help the process of debris removal and recycling, and monitoring after major disasters (Figures 2 and 4).

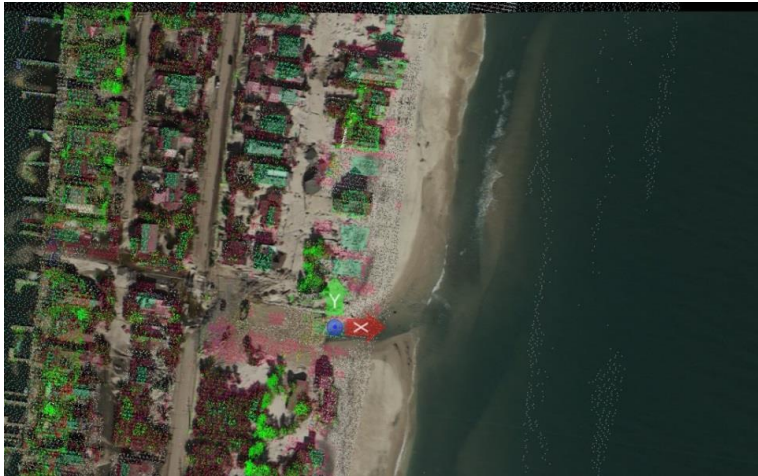


Figure 2 An Overlay of Pre-Sandy and Post-Sandy Airborne LiDAR Data and Post-Sandy Airborne Imagery

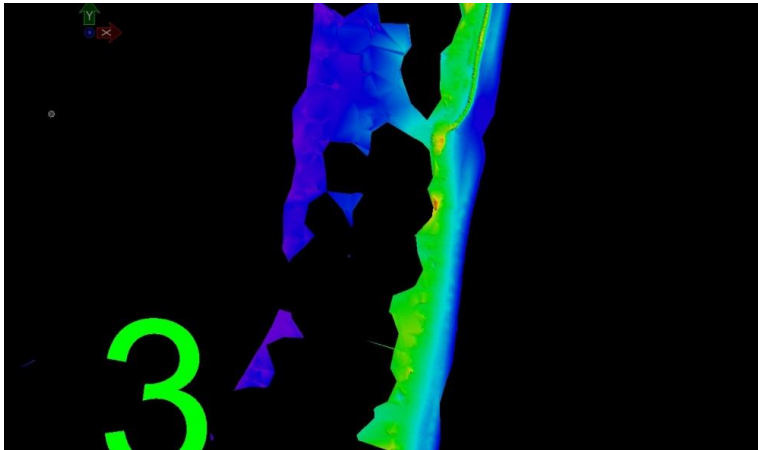


Figure 3 Volume differences generated by subtracting post and pre cleanup conditions

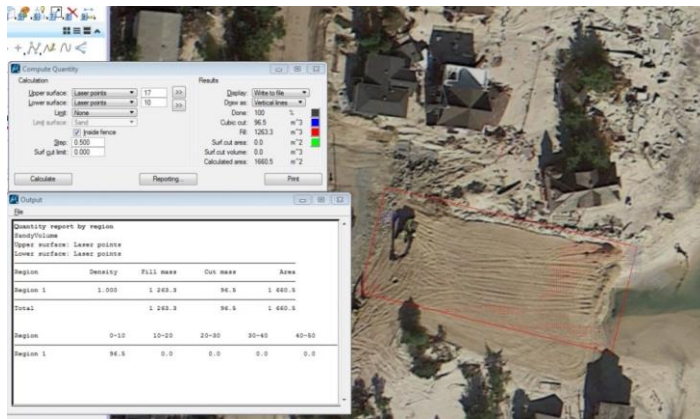


Figure 4 Sample debris volume calculation of a specific area

## Ground-based lidar systems

### Virtual Disaster Site

The high-resolution, three-dimensional images generated by mobile LiDAR can be used to create a virtual reality of the damaged communities (Figure 5). A similar approach has been used with airborne LiDAR data. Traditional post-hurricane damage assessment is a foot-on-ground house by house approach according guidelines such as ATC-45 Field Manual: Safety Evaluation of Buildings after Windstorms and Floods. Several types of evaluations including rapid safety evaluation, detailed safety evaluation, and engineering evaluation will be performed on buildings and residential homes to determine whether the structure is safe for re-occupancy. These evaluations are typically performed in different time periods after a disaster strikes. In many cases, the decisions on structure safety are difficult to make since little quantitative information about the extent of damages is available. Most of the damage data are data merely done by visual estimation without any precise measurement. The availability of 3D mobile LiDAR data can potentially change these common ways of damage assessment.



Figure 5 A LIDAR-based Virtual Disaster Site



### *Change Detection*

In general, damage assessment with 3D mobile LiDAR data can be done using change detection at different scales. At a larger scale, change detection can be performed on large blocks of LiDAR data to identify damaged structures. Using New Jersey's Ortley Beach as an example, we used pre-disaster airborne LiDAR data as the baseline and post-disaster mobile LiDAR data as the post-event data. We generated Digital Terrain Models (DTM) from the pre- and post-event data, then performed change detection on the generated DTMs. Figure 6 shows the change detection results. By examining the change detection results, houses that were 100% damaged can be readily identified (the red parts in Figure 6).

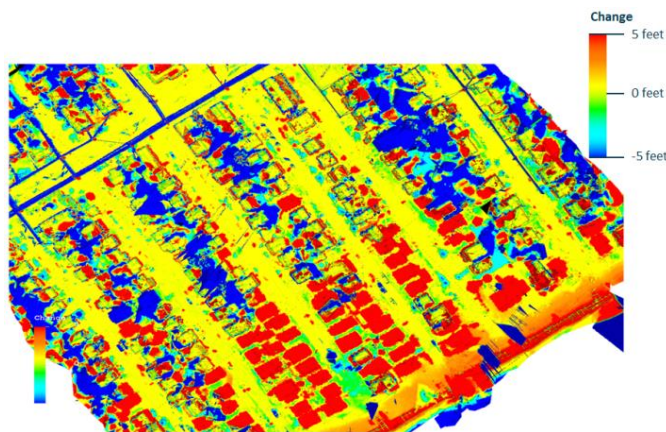


Figure 6 Change Detection Results for Ortley Beach, NJ

### *Damage Assessment*

Beyond visualization, the mobile LiDAR point clouds and imagery provide rich information for residential building damage assessment. First of all, it can be used to conduct accurate measurement on the damaged buildings to extract desired damage information (Figure 7). In addition, it can be integrated with other geospatial data sources to reveal what has changed due to hurricane events. An example of integrating mobile LiDAR point clouds with pre-disaster airborne imagery is shown in Figure 8.

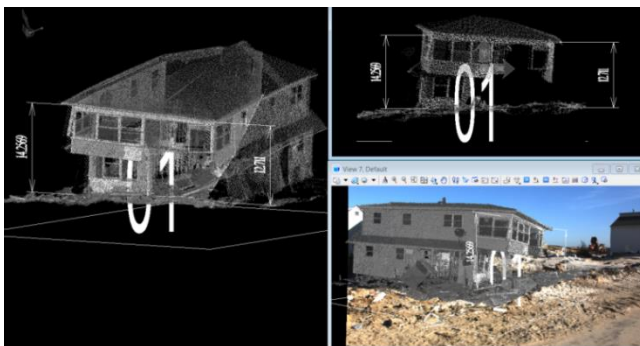


Figure 7 Per Building Level Damage Information Extraction

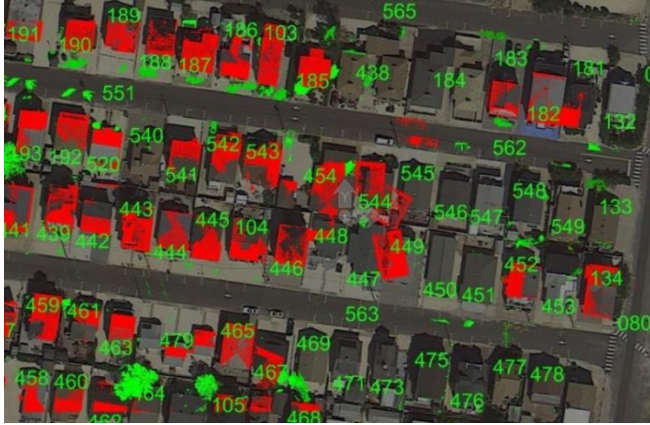


Figure 8 Integrating Mobile LiDAR and Pre-Event Airborne Imagery for Displacement Measurement

### Flooding Scenario Reconstruction

Determining storm surge heights is an important task after a major hurricane event. This information is not only a crucial piece of knowledge that can be used by coastal engineers to improve storm surge prediction models but also reveals the extent of flooding damages to assets. Building inspectors often determine storm surge heights through examining debris trace and water markers on structures and trees. Figure 9 (left), for example, shows the debris traces on the fence. For a typical mobile LiDAR system, photo imagery was geo-referenced and calibrated to the center of IMU. Also the projection between imagery and point clouds is fixed and determined through a bore-sight procedure. As the result, each pixel in an image has known geospatial coordinates. Therefore, geospatial coordinates of the debris traces as shown in Figure 6a can be quickly determined. In this way, the flooding scenario can be reconstructed for visualizing the extent of flooding and its impact on assets (Figure 9).

The three-dimensional images of the flooded community can also be examined in close details by coastal engineers to study wave, surge, and wind damage mechanisms. Wave and surge models have been shown to result in increasing errors as they move overland arising from two major factors: (A) Increased dissipation overland is not fully accounted for in standard models, particularly for water waves; and (B) Overland wind stress will be partially absorbed by canopy elements like trees and houses and will not entirely reach the water surface, with strong implications for surge and waves. The mobile LiDAR data provides a data set with unprecedented detail and accuracy to support the study of the impact of canopy factors on wave dissipation process. These will provide the increased predictive accuracy and detail during surge events needed to make decisions for sustainability.

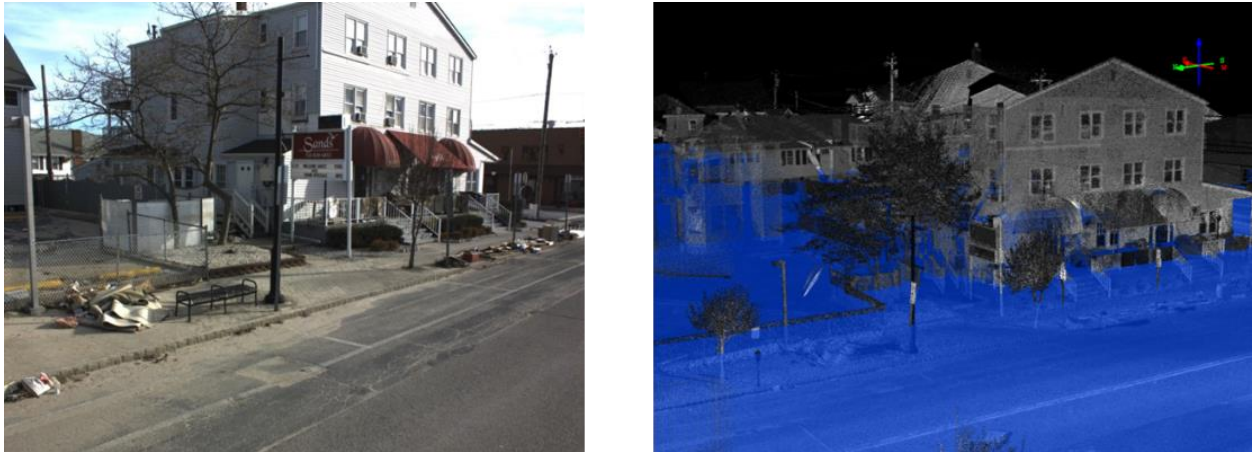


Figure 9 The Reconstructed Flooding Scenario for Ortley Beach

#### *Visualization of Resiliency Rebuilding Requirements*

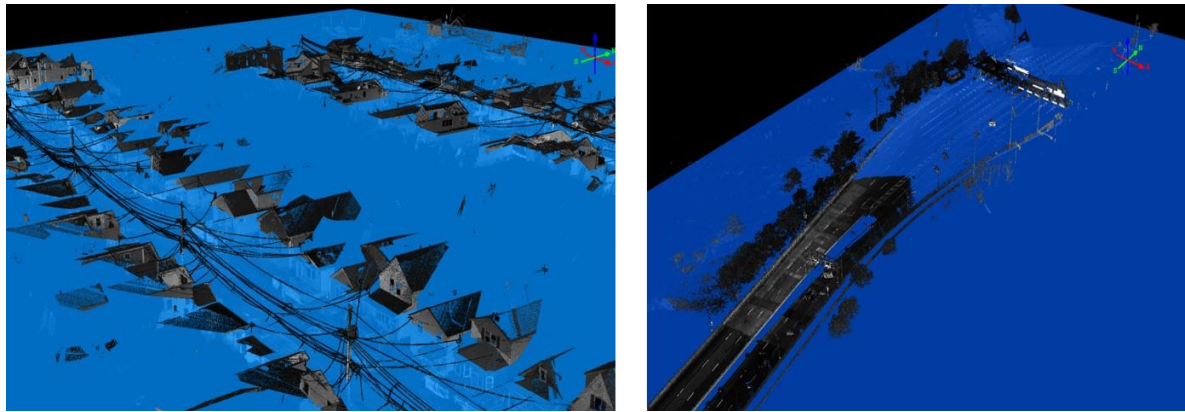
In the aftermath of Hurricane Sandy, FEMA released new Advisory Base Flood Elevation (ABFE) maps for many parts of Jersey Shoreline Communities. These new ABFEs are developed based on sound scientific studies, and are released in the hope of assisting shoreline communities to build stronger and more resilient structures in the face of future storm events. These ABFEs will have profound impact on the cost of flood insurance for residents living in the flood zone. However, to meet the new ABFE, houses have to be elevated; roads have to be rebuilt; and utilities have to be adjusted. The adoption of the new ABFE incurs significant but unknown cost. To many shoreline residents, the risk and the extent of future flooding to their homes are not clear; the benefits of elevating their houses cannot be fully appreciated. Also, the tradeoff between flood insurance cost and mitigation measures is uncertain.

Mobile LiDAR offers an opportunity to scan buildings and other infrastructure at ground level, and to overlay the new Advisory Base Flood Elevation (ABFE) maps on the converted visual digital data to better understand the extent to which they should be raised to prevent future flooding. Incorporating effective resiliency measures into the rebuilding of critical infrastructure and housing is a high priority, as will choosing the most cost-efficient investment strategy across the impacted area.

In this study, we used canopy classification methods include, but are not limited to classifying points by height from the ground and depicting them visually by color, to overlay geospatially accurate digital copies of the recently released FEMA-NJ State Advisory Base Floor Elevation (ABFE) maps onto to the local streets and homes, businesses and infrastructure that were captured in the mobile scans. The results visually projected the location of the advisory minimum first floor elevation on the exterior of existing buildings and over other types of infrastructures. For example,



Figures 10 displayed LiDAR data for a residential community in Rockaway, NYC and one of its vital transportation links (Crossbay Bridge). The new 100 year flood elevation plane was overlaid on these LiDAR data for visual identification of the future risk of flooding to individual homes (Figure 10), the community, and its transportation link (Figure 10). These visual representations are a clear and powerful risk communication means. This new approach of visualizing and analyzing mobile LiDAR data provides city planners and transportation agencies with a great visual tool to understand the vulnerability of communities and transportation infrastructures and the resiliency rebuilding needs.



**Figure 10 The New ABFE 100-year Flood Plane overlaid on Rockaway Homes**

### **Successful Applications of Cloud-based Data Infrastructure for Geospatial Data Management**

In this study, one primary data set of concern is point cloud data set. Currently, management of large point cloud data sets and their derived data products relies on file-based solutions: point cloud data are divided into tiles based on a predetermined spatial boundary template, and each tile is stored into a file in a common file format such as the ASPRS LAZ format. For large geospatial point cloud data sets, this could create a large number of files. For example, AHN2, which contains 640 billion points, is stored and distributed in more than 60000 LAZ files. The file-based solutions provide a rigid way of managing point cloud data sets, which often have varied resolution and coverage. Often the data have to be accessed in the native projection and discretization of the sensors (e.g. tiles), causing unnecessary computational overhead. The rigidity also creates significant challenges for query and runtime optimizations in a unified framework.

Transferring and processing of large point cloud data sets can be accomplished in many different ways. The following table shows three common ways how large geospatial data sets can be stored, shared, and processed.

Table 1 Data Storage, Sharing, and Processing Mechanisms

	Advantages	limitations
Portable storage device Delivery to local centralized computing resource	It is the traditional method for data backup and transfer. It is free from the concerns of cyber security. For the localized data sharing, portable storage device often offers a higher speed.	1 Not efficient in terms of long distance data sharing or among multiple users. 2 Requires large amount of Portable storage device when the data is volumetric.
Cloud computing on nearby distributed mobile computing nodes	Mobile devices are widely available and the computation and communication capabilities of mobile devices improve tremendously. It enriches the underlying resource pool and enable use mobile devices as a source of computing power and storage.	i) the insufficient computing capabilities and unavailability of complete data on individual mobile devices and ii) the prohibitive communication cost and response time involved in offloading data to remote computing resources such as cloud datacenters for centralized computation (Viswanathan et al 2015)
Cloud computing on big computation infrastructures	It has capabilities that enable the scalable, extensible and interoperable utilization of regional, national and international production distributed cyberinfrastructure.	1 High width band internet are required to connect to these cyberinfrastructures; 2 Special network interface are required to connect to these facilities.

Large collections of 3D geospatial data sets such as point cloud or airborne imagery are often partitioned into tiles, and each tile is stored in a file. Data processing tasks in a given workflow can run in parallel on each individual tiles, leading to the so-called “embarrassingly parallel” problem. Distributed computing frameworks such as Hadoop and Spark have been recently studied to explore their capability in managing large

geospatial data sets and carrying out certain processing tasks. For example, some of these studies focus on generation of Digital Elevation Models (DEM) (Krishnan, et al 2011; Jian et al 2015; Růžicka et al 2017; Hegeman et al 2014). Others implemented the concept of parallel or distributed computing into other applications such as Generate Octree for visualization (You et al 2014), Change Detection (Liu et al 2016), Feature Extraction (Guo et al 2015) and Geographic Data Storage (Hanusniak et al 2016). A more complete list of these applications is shown Table 2.

Table 2 Distributed Computing for Geospatial Data Processing

Application	Archetype	Reference
Generation of digital elevation model (DEM)	Hadoop	Krishnan, et al 2011; Jian et al 2015; Růžicka et al 2017; Hegeman et al 2014
Generate Octree for visualization	Hadoop	You et al 2014
Change Detection	Spark	Liu et al 2016
Feature Extraction (polygon retrieval)	Hadoop	Guo et al 2015
Geographic Data Storage	Hadoop	Hanusniak et al 2016

### **Requirement Modeling on Cloud-based Geospatial Data Analytics for Time Sensitive Applications**

In this task, we approach the problem from two perspectives. First, we characterized the amount of data that need to be processed in order to derive valuable insights into better disaster response strategies. Second, we conducted a comprehensive review on data processing needs in post-disaster situations. In the following, we detail our findings in each effort.

Although big data do not purely mean the large volume of data, data volume remains a major concern in disaster response and recovery missions, where how large the amount of data generated often determines what kind of protocols to be used for storage and transferring and how much computation resources is required to process it. We systematically analyzed the volume of the above data sets. Table 3 provides a quick summary about the volume of some typical data sets used in disaster response and recovery phases during Hurricane Sandy.

Table 3 Volume of Hurricane Sandy related 3D disaster data sets

	<b>Data collection date</b>	<b>Data Volume</b>
Archived airborne lidar	Archived	29.6GB
USGS EARL-B lidar	10/29/12	2.1GB
USGS EARL-B lidar	10/31/12	2.1GB
USACE lidar	11/19/14	21.2 GB
Rutgers mobile lidar	12/01/12	575GB
USGS CMGP lidar	2014	105GB
Photos for SFM reconstruction	Streaming	20GB
Rutgers mobile lidar	2015-2016	15TB

While the potential of large geospatial data sets in disaster management is well known, these large data sets often remain peripheral information during natural disasters (Brucewicz 2003; Alexander 2002). These value-added products that can be derived out of these geospatial data sets are not used in decision making due to the fact that minimum consideration has been placed on the time requirements, i.e., how fast these data need to be processed (Lippitt et al 2014). Given the fact that decision making in disaster response is time sensitive, information derived from remote sensing data sources is also time sensitive, and therefore remote sensing in disaster scenarios is time sensitive (Lippitt et al 2014; Hodgson et al). Based on this, information extraction from large geospatial data sets has a limited temporal utility windows (Lippitt et al 2014; Hodgson et al). The timeframe for collecting and analyzing the information needed to make decisions differs from one organization to another and can vary according to the type of disaster considered (slow onset vs sudden onset). On a regional and field level, the decisions need to be taken faster, in the majority of the cases within 48 hours. Globally, and if there is no presence on the ground, the respondents claim that the timeframe ranges from 6 hours to 7 days. In another paper, Hodgson et al. surveyed many emergency response organizations in the United States, and developed several very important curves (2014) regarding how the value of information regarding various damage impacts degrades as the time goes on (Figure 12). These charts are extremely useful in terms of providing a realistic bounds for processing requirement.

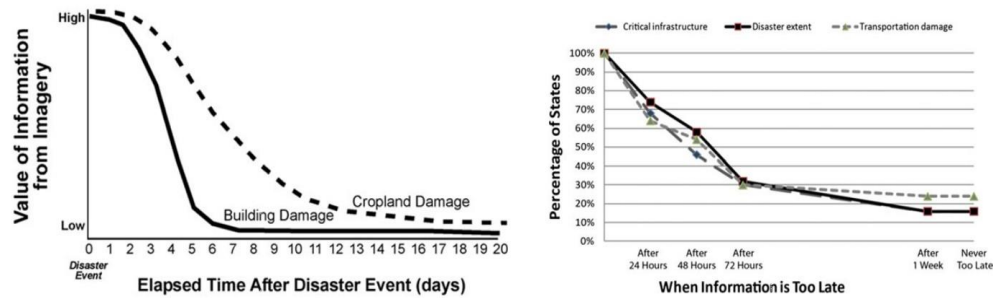


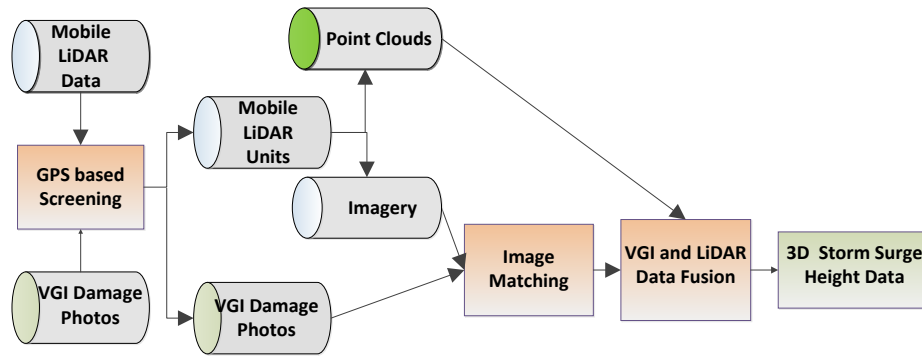
Figure 11 Information Value vs. Time (adapted from Hodgson et al. 2014)

## Web-based Disaster Data Collection, Visualization, and Processing: Examples and Case Studies

### ***Case Study 1: Fusion of Geo-Tagged Crowdsourced Post-Storm Damage Photos with Mobile LiDAR Data for Storm Surge Height Measurement***

We introduce a new data fusion method that fuses disaster photos from a variety of sources (field teams, volunteers, and Internet) with mobile lidar data for accurately obtaining high water mark information without performing field measurement. Because of the pervasive use of mobile devices, disaster photos are now often widely shared and available on the Internet. Many of these photos provide first-hand information on the extent of disaster, some of them way before the arrival of field teams. The new method will promote a crowd source-based approach for collecting critical hurricane impact parameters

The technical rationale behind our proposed approach is local feature-based image matching and 3D alignment of point clouds with photos from heterogeneous sources, such as disaster assessment teams, social media, and etc. We assume the photos taken from sources other than the mobile lidar system itself are geo-tagged as the pervasive use of GPS capable mobile devices. Figure 13 outlines the detailed workflow in our proposed approach. The major components are GPS based screening, image matching, and VGI and lidar data fusion. To develop and validate all the methods involved in this workflow model, we used several data sets acquired during Hurricane Sandy.

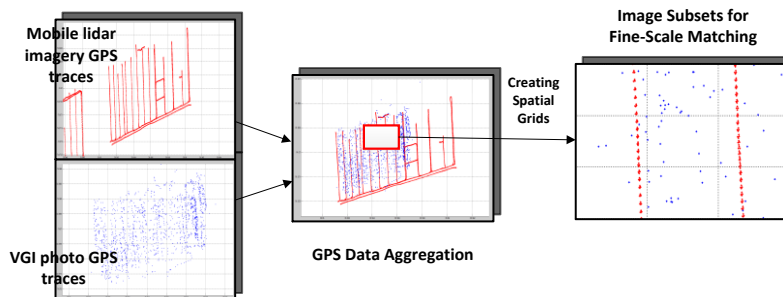


**Figure 12 Proposed Surge Height Estimation Approach**

### GPS-based Screening

Running local feature-based image matching between the entire set of mobile lidar imagery and survey photos is prohibitively expensive, even if it is only for individual communities. Considering both mobile lidar imagery and survey photos have embedded GPS information, albeit at different positional accuracies, one logic step is to group them into subsets - a step we referred to as GPS-based screening. Another motivation behind the screening is that it is well-known that building structures often have similar local features that tend to confuse the local feature based image matching methods. By dividing the photo sets into smaller subsets, this source of confusion can be greatly reduced.

The mobile lidar imagery in this research is organized according to the vehicle trajectories. Thus, the trajectory information was used as one source of information to pair with field survey photos. The detailed steps for GPS screening are shown in Figure 3. An essential step in this approach is projecting the GPS traces of VGI and mobile lidar photos onto one map frame which has grid sizes spanning 0.002 latitude and 0.001 longitude. The subsequent photo matching will be only carried out within each grid. In this way, computational effort and complexity can be greatly reduced.



**Figure 13 GPS-based Screening of Images from Different Sources**

## Image Matching

This step involves registration or alignment of VGI photos with mobile lidar imagery. The mobile lidar imagery has known projection properties to lidar point clouds. Therefore, once the relationship between mobile lidar imagery and VGI photos can be derived, 3D information of water marks in a common chosen coordinate system can be calculated. However, manually finding the paired mobile lidar imagery for each VGI photo is not an easy work. To automatically detect images which captured damages of the same structures, a SIFT (Scale-invariant feature transform) based method for automated image matching was employed (Figure 15). This method is capable of finding images which display similar scenes based on local image features. Figure 4 shows some example matching of damage images from two different data sets used in this research. It can be seen that the method is robust to view angle and illumination variances.



Figure 14 SIFT-based Matching between VGI photos and Mobile Lidar Imagery

## VGI and Lidar Data Fusion

This step involves 3D alignment of VGI images with mobile lidar imagery. Once pairs of mobile lidar imagery and VGI images are identified, it is straightforward to relate VGI photos to point cloud data since there is known correspondence between mobile lidar imagery and lidar point clouds. Once this is accomplished, there are three steps involved in recovering the coordinates of high water marks.

The first step is to compute the Projection Matrix  $\mathbf{P}$  between each point in a lidar point cloud  $\mathbf{P}_w$  and each pixel in a VGI image  $\mathbf{P}_c$ . Currently, this projection is solved with the assistance of manual inputs of several correspondences between VGI image pixels and lidar point clouds. We used the camera calibration with 3D objects method developed by Zhang (2004). More specifically, the feature points of the building in both lidar point cloud and images are manually detected in exactly the same order as shown in Figure 5, and are saved  $P_c$  and  $P_w$ . The 2D pixel is denoted by  $P_c = [x, y]^T$ , and 3D point is denoted by  $P_w = [X, Y, Z]^T$ . Write the point coordinate as homogeneous coordinate format as  $P_c = [x, y, 1]^T$  and  $P_w = [X, Y, Z, 1]^T$ . The relationship between  $P_w$  and  $P_c$  could be expressed as

$$sP_c = \mathbf{A}[\mathbf{R} \ \mathbf{t}]P_w \quad (1)$$

where

$$\mathbf{A} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

is the intrinsic matrix and  $\mathbf{P} = \mathbf{A}[\mathbf{R} \ \mathbf{t}]$  is the projection Matrix. Based on (1), the correspondence between  $P_w$  and  $P_c$  could be written as:

$$\begin{bmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & x_i X_i & x_i Y_i & x_i Z_i & x_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & y_i X_i & y_i Y_i & y_i Z_i & y_i \end{bmatrix} \mathbf{P} = \mathbf{0}$$

where

$$\mathbf{P} = [p_{11}, p_{12}, \dots, p_{34}]^T$$

For the  $n$  selected feature points, stack all equations as:

$$\mathbf{G}\mathbf{P} = \mathbf{0}$$

$$\mathbf{G} = [\mathbf{G}_1^T, \dots, \mathbf{G}_n^T]^T$$

The Projection Matrix  $\mathbf{P}$  is the eigenvector of  $\mathbf{G}^T \mathbf{G}$  associated with the smallest eigenvalue.

The next step is to obtain the correspondence between LiDAR Point Cloud and Image. Since the Projection Matrix is a 3 by 4 matrix, which is non-invertible, the image cannot be projected onto the LiDAR directly. Instead, the LiDAR Point Cloud is projected onto the image using the computed Projection Matrix  $\mathbf{P}$ . For each point of the LiDAR Point Cloud  $P_{w_i}$ , we find out where it is projected in the image  $P_{c_j}$  and construct



the correspondence between  $i$  to  $j$ . The correspondence is illustrated in Figures 16 and 17.

The last step is to assign each point  $P_{w_i}$  a RGB value from the image  $P_{c_j}$ . The lidar point clouds then have information on which points are projected onto the high water marks shown in the VGI images. These water marks can be directly measured in the lidar point clouds (Figure 18).

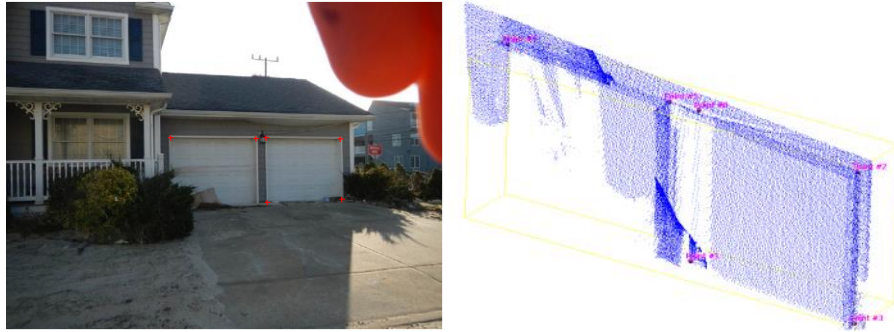


Figure 15 Establishing Correspondence Between Point Clouds and VGI Images

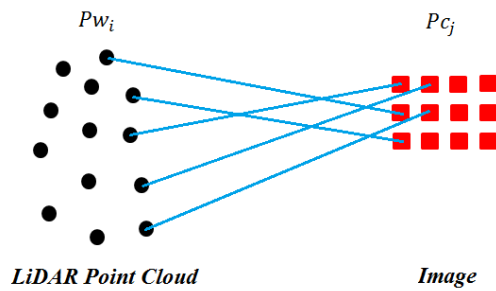


Figure 16 The correspondence between Lidar Point Clouds and VGI Images



Figure 17 High Water Mark Measurement from VGI Images and Point Cloud Data

### ***Case Study 2: Post-Disaster Damage Assessment of Natural Gas Pipeline Systems with Remotely Sensed Data***

In this case study, we demonstrate the use of cloud-based data sources for assessing gas line damages after major hurricane events. The entire framework is shown in Figure 19. To facilitate remote sensing based risk assessment, it is important to realize that a distributed approach would be necessary. This is due to several reasons: (1) most gas operators do not collect remote sensing data by their own; instead, they use publicly available data or hire contractors to do so; (2) most gas operators are reluctant to share data about the location and conditions of their assets as they are deemed sensitive; and (3) natural disasters are rare, meaning it is not economic for them to own software packages that can integrate remote sensing data and risk assessment models. Based on these observations, what we proposed is a distributed and cloud based business model. The workflow we proposed is: the software packages are divided into two components: Web-based risk assessment model and a standalone software package that deals with processing collected remote sensing data and detect hazardous conditions posing threat to the natural gas pipeline system. Once a natural disaster strikes, the gas operator chooses the region of impact for analysis. The threat detection software gathers and processes available remote sensing data and detects salient threats. The geospatially referenced threat data are then extracted and sent back to gas operators. This step does not need detailed information about the locations and conditions of gas infrastructure assets. Then the gas operators upload encrypted gas facility data (through shuffling the data) and their relevant geospatially referenced threat information to the web-based risk assessment program to estimate spatially registered risk on their pipeline segments. This framework would not require the gas operators to purchase the risk assessment program and the threat detection program but only pay as you use. In the same time, it avoids the issue of exposing sensitive pipeline data to the third party.

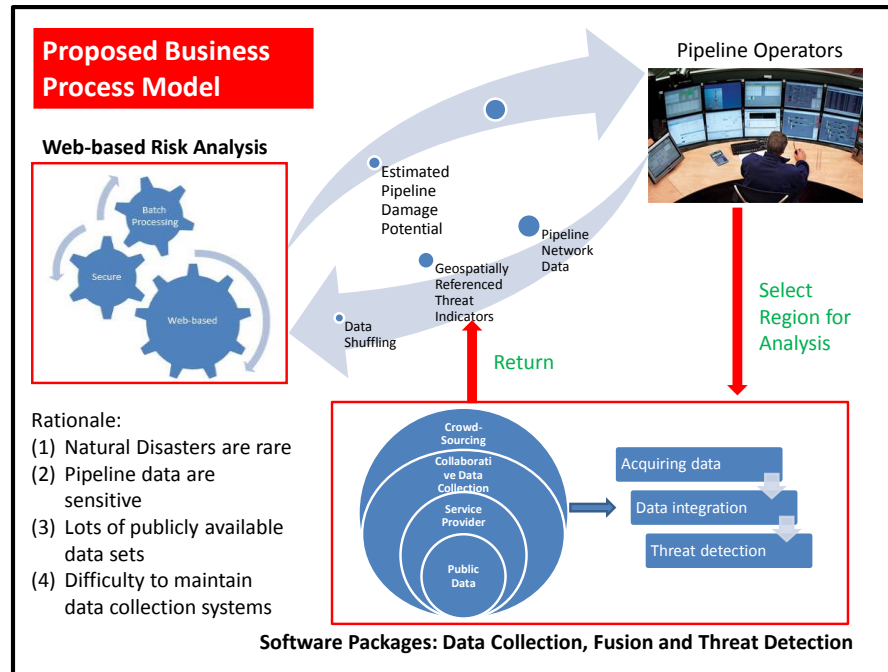


Figure 18 Cloud-based Post Disaster Damage Assessment of Natural Gas Pipeline System

With the cloud-based remotely sensed data, our approach is built on top of a data-driven risk assessment method. Natural gas pipeline failures during a natural disaster are often related to changes in the built environment (structures, terrain, etc.,) adjacent to pipelines. 4 lists common pipeline threats and related indicators caused by natural disasters. The remotely sensed data are used to detect and quantify these indicators. A framework for using these indicators to assess the risk of natural gas pipeline networks is shown in Figure 20. The rationales of this framework include the following: 1) for aboveground pipelines and gas meters, the assessment is conducted based on the assessment of building changes and damage; and 2) for buried pipelines, the main threats indicators are soil movement and flooding height. There are four types of building conditions considered in this framework. They are “no-damage”, “minor-damage”, “major-damage”, and “total-damage”. The first two conditions lead to a decision to inspect the aboveground pipeline segments, while the latter two lead to a decision to replace them. Regarding underground pipeline facilities, the framework requires information including soil settlement, vertical soil movement, horizontal soil movement, and flooding heights to estimate pipeline strain in order to draw conclusions about the probability of failure. A Finite Element Analysis is then used to estimate the

potential pipeline strains once soil movement and flooding height are quantified. Calculated pipeline strains were further imported into a pipeline risk analysis program to estimate probability of failure.

**Table 4 Pipeline threat and related indicators**

Threat	Cause	Phenomenon	Indicator
Water Infiltration	Pressure Head	Water Level Resulting from Flood	Water Elevation Above Ground Surface
Underground Pipe Break	Strain	Soil Deformation Resulting From Flood, Landslide, Hurricane, and Earthquake	Soil Displacement
Above Ground Pipe Break	Strain	External Force from Flood, Landslide, Hurricane, Earthquake, Tornado	Asset Displacement
Exposed Pipe	Soil Erosion	Soil Erosion Resulting From Flood, Landslide, Hurricane, Earthquake	Soil Displacement

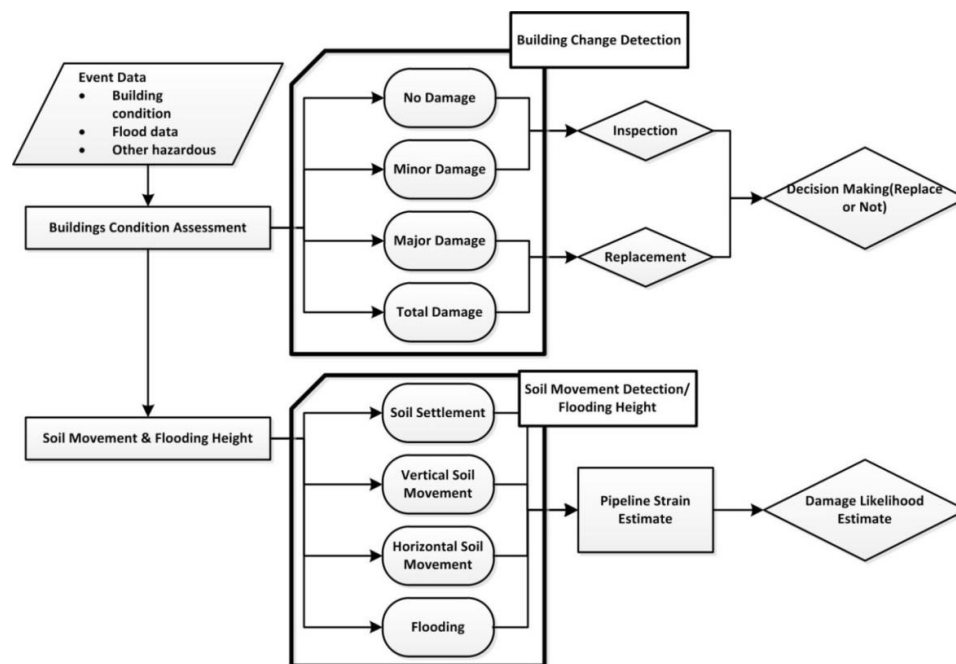


Figure 19 Proposed Post-disaster Pipeline Risk Assessment Framework

## CONCLUSIONS

Preparing for and responding to natural disasters are involving an increasing amount of geospatial data collection and processing. These data are stored in a variety of

platforms such as widely distributed repositories, large data centers, and social media networks. Data volume, variability, and variety pose significant challenges in using these data sets effectively in time-sensitive applications. This study systematically investigated some fundamental aspects of geospatial disaster data including the composition of these data, the characteristics of these data, processing requirements, and available data infrastructure for managing, sharing, and processing these data sets.

The result of this research indicates that despite the great potential of using geospatial data sets to support disaster response, most geospatial data have remained peripheral information due to the complexity and difficulty in processing them in time. Our analysis of geospatial data use in Hurricane Sandy supports this conclusion. Our empirical study also indicates there are a variety of processing patterns related to these data sets. Some of these processing have to be done in real-time, while some others can be done with more time but with higher accuracy and resolution. There is apparent trade-off between the accuracy and the time required for processing the growing volume of geospatial data. It also appears that there is a lack of data infrastructure platforms for managing, sharing, and processing the growing volume and variety of geospatial data sets. The rising of big data processing platforms such as SPARK and Hadoop provides great opportunities to address these pressing data processing problems. Research along this frontier just start emerging, and will likely provide feasible solutions. On the other hand, massive point cloud visualization software such as POTREE is already shaping the workflow in managing, sharing, and analyzing large point cloud data sets. It is reasonable to expect that these kinds of software programs will soon integrate capabilities to visualization other types of geospatial data and provide capabilities in collaboratively working on distributed geospatial data sets. Given the rapid development in this existing field, cloud computing and distributed processing will become the dominant ways of processing future large geospatial data sets.

## REFERENCES

Alexander, D.E., *Principles of emergency planning and management*. 2002: Oxford University Press on Demand.

- Brédif, M., B. Vallet, and B. Ferrand, *Distributed Dimensionality-Based Rendering of LIDAR Point Clouds*. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2015. **40**(3): p. 559.
- Bruzewicz, A., *Remote sensing imagery for emergency management*. The Geographical Dimensions of Terrorism. New York: Routledge, 2003: p. 87-97.
- Cura, R., J. Perret, and N. Paparoditis, *Point Cloud Server (pcs): Point Clouds In-Base Management and Processing*. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2015. **1**: p. 531-539.
- Hegeman, J.W., et al., *Distributed LiDAR data processing in a high-memory cloud-computing environment*. Annals of GIS, 2014. **20**(4): p. 255-264.
- Hodgson, M.E., et al., Geospatial data collection/use in disaster response: a united states nationwide survey of state agencies, in *Cartography from Pole to Pole*. 2014, Springer. p. 407-419.
- Jian, X., et al., *a Hadoop-Based Algorithm of Generating dem Grid from Point Cloud Data*. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2015. **40**(7): p. 1209.
- Krishnan, Sriram, et al. "OpenTopography: a services oriented architecture for community access to LIDAR topography." *Proceedings of the 2nd International Conference on Computing for Geospatial Research & Applications*. ACM, 2011.
- Růžicka, Jan, et al. "Processing LIDAR Data with Apache Hadoop." *The Rise of Big Spatial Data*. Springer International Publishing, 2017. 351-358.
- Sugumaran, R., J. Burnett, and A. Blinkmann. *Big 3d spatial data processing using cloud computing environment*. in *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*. 2012. ACM.
- Lippitt, C.D. and D.A. Stow, *Remote Sensing Theory and Time-Sensitive Information*, in *Time-Sensitive Remote Sensing*. 2015, Springer. p. 1-10.
- Lippitt, C.D., D.A. Stow, and K.C. Clarke, *On the nature of models for time-sensitive remote sensing*. International Journal of Remote Sensing, 2014. **35**(18): p. 6815-6841.
- Liu, Kun, Jan Boehm, and Christian Alis. "Change detection of mobile Lidar data using cloud computing." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives*. Vol. 41. 2016.
- Wang, F. and H. Yuan, Challenges of the sensor web for disaster management. *International Journal of Digital Earth*, 2010. **3**(3): p. 260-279.